Building the Adolescent Life Skills Assessment for Girls (ALSA for Girls): Synthesis Report on Lessons Learned

March 2022

How to Cite This Report

Room to Read (2022). "Building the Adolescent Life Skills Assessment for Girls (ALSA for Girls): Synthesis Report on Lessons Learned." Available at https://www.roomtoread.org/alsa-synthesis-report.

Table of Contents

Table of Tables

Table of Figures

Abbreviations

AIR	American Institutes for Research
ALSA	Adolescent Life Skills Assessment
AS	Acquiescence Score
CASEL	Collaborative for Academic, Social, and Emotional Learning
CST	Challenging Situations Task
DESSA	Devereux Student Strengths Assessment
GEP	Girls' Education Program
IIHMR	International Institute of Health Management Research
IRB	Institutional Review Board
OECD	The Organisation for Economic Co-operation and Development
PISA	Programme for International Student Assessment
RES	Response Extremeness Score
SDQ	Strength and Difficulties Questionnaire

Executive Summary

This report synthesizes Room to Read's efforts from 2015 to 2020 to develop, test and utilize an Adolescent Life Skills Assessment (ALSA) for girls as an outcome measure for its Girls' Education Program (GEP). Since 2001, GEP has helped more than 140,000 girls to complete secondary school with the skills needed to make key life decisions through its in-school programming, and millions more through TV and radio broadcasts of life skills content. Life skills are critical building blocks supporting the wellbeing of children, youth and adults. They are transferable and include skills such as self-confidence, creative problem-solving and expressing and managing emotions.[1] Each life skill includes multiple attributes, which makes measuring them challenging.

This report details the ALSA development process, progress and lessons learned throughout this six-year experience across seven countries: India, Tanzania, Cambodia, Bangladesh, Nepal, Sri Lanka and Vietnam. We hope that this synthesis report will help other development organizations, researchers and evaluators in their efforts to develop ALSAs for their programs and contribute to the broader knowledge base about measuring life skills.

Figure 1. GEP Life Skills Framework
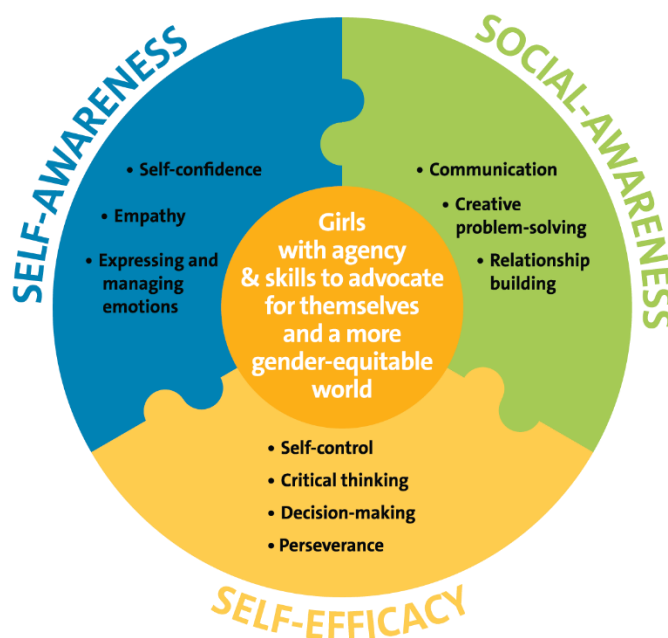
Given the centrality of life skills to GEP and its theory of change, we prioritized the development of a set of measures to assess girls' development of life skills.

---

[1] What we refer to as life skills, other organizations may label as socio-emotional, 21st century, soft or transferable skills. We use "life skills" throughout this report.

ALSA Development Process

The development of the ALSA at Room to Read began in 2015 with a tool review to identify existing life skills measures that might be useful in a tool that would align with the needs of GEP (see Figure 2). We built on existing measures that had been tested and validated, often in the United States. As a result, we adapted tools to fit GEP students where they live.

Figure 2. ALSA Development Process
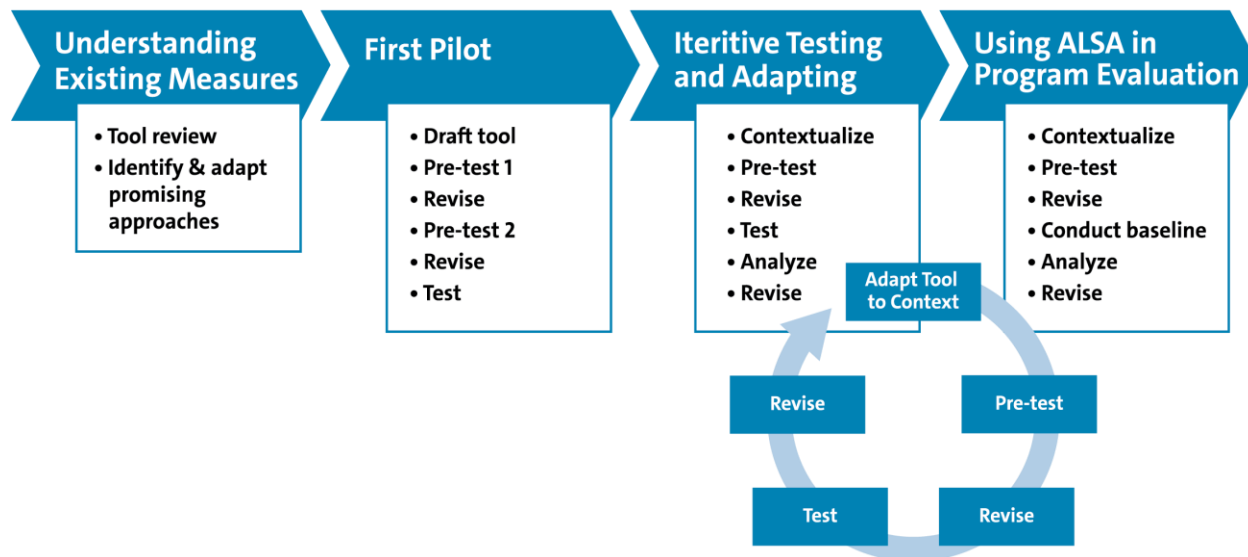
The core of the ALSA is a self-report questionnaire for girls. Each of the pilots in the seven countries includes the questionnaire and at least one validation method outside of the girl questionnaire, such as social mobilizer[2] ranking or teacher's survey (see *Annex 3: Validation Instruments*). The girl questionnaire consists of multiple modules including:

- Demographics
- Introduction to scales used in the tool: warm-up for respondents
- Life skills modules for:
  - ✓ Decision-Making
  - ✓ Perseverance and Self-Control
  - ✓ Emotions and Empathy
  - ✓ Relationship-Building
  - ✓ Communication
  - ✓ Self-Confidence
  - ✓ Creative Problem-Solving
- Knowledge and attitude questions related to gender

---

[2] "Social mobilizers," also known as mentors, are program staff who work most directly with participants.

Within the life skills modules, we sought to identify and address bias using free listing exercises, forced choice questions and, for Creative Problem-Solving, some hypothetical conflicts in which the respondent is asked to list what the person in the conflict would do, and by comparison, what she herself would do. To complement the ALSA, we included questions about voice, and school and career aspirations in some of our GEP evaluations. See *Annex 1: Sample ALSA Tool* for details.

Learning from Pilot Studies

The largest leaps in the development of the girl questionnaire resulted from the first three pilots (India, Tanzania and Cambodia). In the subsequent four pilots (Bangladesh, Nepal, Sri Lanka and Vietnam), we focused on our validation methods, further refining life skills constructs and testing a self-administered questionnaire. Table 1 below summarizes the challenges in the first three pilots and how we addressed them. Despite these efforts, contextualization for additional countries and languages remains challenging and requires explicit attention and additional piloting.

Table 1. Challenges and Mitigation Strategies in Early Pilots

Challenges	Mitigation Strategies / Actions
Some respondents found the questionnaire long, challenging, and/or tiring	• Dropped items and modules that did not perform well • Trained enumerators for four days so they thoroughly understood each item and practiced using good interviewing techniques in roleplaying with each other • Identified synonyms that were approved for use if a respondent did not understand a key word in an item
Respondents did not all understand items or understood them the same way	• Adapted items to local context while maintaining intent of each item • Simplified words/phrases • Paid close attention to translation/back-translation • Avoided long items, complex grammar, idioms (unless local) • Conducted pre-tests that included cognitive interviews and group discussions with girls to "get the words right" • Trained enumerators to use approved synonyms only
Some respondents had difficulties with scales	• Added an introduction to scales with practice questions after the demographic module and before the life skills modules to accustom respondents to scales • Used showcards (handouts) illustrating answer options
Forced choice questions were challenging as they required respondents to remember four options	• Used showcards (handouts) for forced choice answer options
Heavy right-tailed data suggested social desirability bias	• Reversed scales to start with "not true" or "strongly disagree"

	• Added introductions reminding respondents that there are no right or wrong answers and that their honesty is important
Low reliability of life skills constructs — The above strategies increased the reliability of the constructs but unevenly across contexts and between pilots and baselines in the same country (although not necessarily the same locales)	• Tested new items (also from the tool review), used expert review and our conceptual understanding of the life skills to interpret results, and dropped poor performing items

Following the first three pilots, the girl questionnaire's performance became somewhat stable. Some of the life skills measures perform better than others when assessing through the lens of Cronbach's alpha (see Figure 3), but we recognize that this measure is limited by its sensitivity to the number of items and the distribution of scores, and as such, is only one measure we consider when analyzing the quality of our assessments.

Figure 3. Range of Alpha Scores by Skill Area across All Pilots (excluding India)
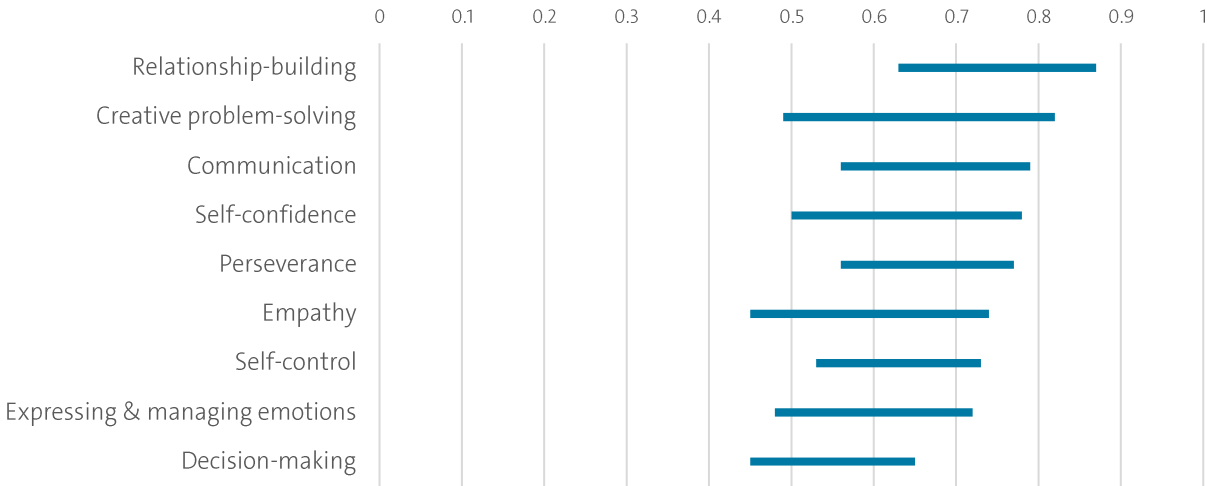
Improving the questionnaire involved a range of methods, some of which were more effective than others toward achieving a reliable and valid girl questionnaire, the core of the ALSA. Table 2 provides a snapshot view of our piloting efforts. The blue-shaded rows across the girl questionnaire were piloted as part of our efforts to identify and address bias, while the gray-shaded rows were used to address validity.

Table 2. ALSA Development over Time

| | Country | India | Tanzania | Cambodia | Bangladesh | Nepal | Sri Lanka | Vietnam |
| | Year | 2016 | 2016 | 2016 | 2017 | 2018 | 2018 | 2018 |
	Type	Pilot	Pilot	Pilot	Pilot	Pilot	Pilot	Pilot
Girl Questionnaire	Demographic / Student Profile	●	●	●	●	●	●	●
	Life Skills Modules	●	●	●	●	●	●	●
	Listing*: Emotions	●	●	●	●	●	●	●
	Listing: Peer Nomination	●	●	●	●	●		
	Listing: Solutions to Problems	●	●	●	●	●	●	●
	Listing: Ways to Calm Down						●	●
	Listing: Qualities of a Good Communicator						●	●
	Listing: Respondents Good Qualities						●	●
	Listing: Qualities of a Leader						●	●
	Forced Choice	●		●	●	●		
	Anchoring Vignettes	●	●	●	●	●		
	Map Exercise	●						
	School & Career Aspirations						●	●
	Voice						●	●
	Mirror Drawing	●	●	●	●			
	Strength & Difficulties		●	●	●	●		
	Roleplay Exercise							●
Third Parties	Social Mobilizer Ranking of Girls				●	●	●	●
	Parent Survey	●	●	●	●			
	Teacher Survey				●	●	●	
	Interviewer Questions					●	●	●

*Free listing exercises ask girls to brainstorm by naming, for example, all the emotions they felt in the past week.

Testing Self-Administration

In Pilot 7 (Vietnam), we tested a self-administered paper questionnaire with girls to compare with the face-to-face interview. Self-administration included the interviewer introducing the questionnaire to a girl and walking her through the practice questions, then leaving her to complete the questionnaire by herself. Our main objectives with self-administration were to explore a potentially efficient administration process and reducing social desirability bias, which can result in overly positive answers and scores skewed to the right. However, we found very little difference between the mean life skills scores of respondents who were interviewed and those who self-administered the questionnaire. Additional testing with self-administration is needed, especially with technology (e.g., tablets) where we can build in support for girls.

Learning from Multiple Analytical Approaches

In addition to the above within-questionnaire methods and third-party methods, we addressed reliability (items holding together well and addressing the underlying life skills construct), validity (measuring what we intend to measure), and bias (sources of systematic prejudice toward particular outcomes) in multiple ways.

Reliability

In terms of *reliability*, the internal consistency of life skills constructs has generally improved over time. Examples of how we made decisions using factor analysis are illustrated in *Section 1.4.1: Addressing Reliability*. We conducted a small sample test-retest in Pilot 5 (Nepal) with results suggesting a larger sample for the test/re-test is needed (see *Section 4.1.2: Test-Retest*).

Validity

Regarding *validity*, in addition to the third-party tools named in Table 2, we have used external data and task-based measures (see *Section 1.4.2: Addressing Validity* for descriptions, analysis and results). After the first four pilots, we dropped the parent survey, as we found that some parents may not be aware of the full range of their daughters' emotions and opinions, or some of their behaviors outside the home. See Table 13. Results of Using Third-Party Assessments for further details.

We have a number of task-based measures, most notably: 1) a mirror drawing exercise that proved too cumbersome and showed no correlation with the Perseverance life skill construct; and 2) a roleplay that evaluates girls' life skills in real life situations. We dropped the mirror exercise and intend to further test the roleplay, despite its design being challenging and its administration being resource intensive.

We collected academic performance data in several subjects and overall grades and found almost no correlation with life skills scores. This could have been a function of our sample size for that pilot and further studies to measure the relationship between academic performance and life skills, while controlling for student variables, are needed. The next steps to improve the questionnaire's

validity include better aligning our life skills constructs with the skills definitions using resources like the Harvard Explore SEL Taxonomy Project[3] and further testing the roleplay exercise.

Bias

We address potential *bias* in many ways within the questionnaire including using positive and negative items, anchoring vignettes, free listing and forced choice questions. Including *positive and negative items* may reduce acquiescence bias, sometimes called agreement bias, where respondents tend to disproportionately select positive response options. Our findings show that transforming data to remove acquiescence bias using the girls' responses to the positive/negative items improves the properties of some of the scales. We have not drawn a conclusion at this point, as we are still working on validating the scales. See *Section 4.3.1: Positive and Negative Items* for details and results.

Anchoring vignettes are scenario-type questions in which girls are asked to respond to short stories about the behavior of other hypothetical girls their own age. These questions are meant to adjust for reference bias so that girls' answers about themselves can be more accurately interpreted. None of our pilot analyses allowed us to conclude that rescaled items based on vignettes were more reliable than non-rescaled items. They were time consuming in a survey that takes on average 40 minutes for a baseline and up to one hour for a pilot. We included the vignettes in the Tanzania baseline but dropped them from the Sri Lanka pilot in 2018, and have not used them since (see *Section 4.3.2: Anchoring Vignettes* for more detail).

Free list exercises assess girls' familiarity and comfort levels within a given topic (see *Section 4.3.3* for details). Since the first pilot, we have asked girls to think about and name the emotions they experienced in the past seven days. We use the results to check for girls' conceptual understanding. We learned, for example, that older girls tend to name more emotions than younger girls and that older girls also name emotions that younger girls do not. We expect the number of: 1) emotions a girl says she experienced in the past week; and 2) ways to calm down when angry to correlate with the life skills of Empathy and Expressing & Managing Emotions. We find weak but statistically significant correlations with some of the measures in Pilots 6 and 7 (see Table 24).

Forced choice questions can be challenging to respondents yet have the potential to identify bias. Each forced choice question consists of four statements. The respondent identifies which statement is "most like her" and which is "least like her." Each statement relates to a life skills question. See *Section 4.3.4: Forced Choice Questions* for a description of the questions and how they are used to identify bias.

Critical Reflection

We continue to refine and test the girl questionnaire in pilots and GEP evaluations. The girl survey became sufficiently stable after the first three pilots to use as a baseline instrument while continually being improved upon. In six years, we learned many lessons about building an ALSA tool. Developing a standardized ALSA for Girls is a great challenge. Engaging in this work has taught everyone who worked/works on it to be creative, open and humble. We have not resolved every

---

[3] See more at https://easel.gse.harvard.edu/taxonomy-project

challenge, but we have made progress in many areas and are convinced of the value of continued development. See *6.1: Critical Reflection* for more of our reflection on the iterative testing, learning, revising and testing over the six years prior to the COVID-19 pandemic.

We see the benefit of:

- Testing internal consistency using multiple methods;
- Testing validity using within-tool measures and exercises, third-party assessment and external data;
- Looking for and mitigating bias in the data;
- Contextualizing the girl questionnaire: attending to how skills manifest relative to location, culture, age and privilege within a context; structural factors affecting how life skills translate into life outcomes; and local stakeholders' views on which skills are important as they vary across contexts; and
- Using an iterative approach to ALSA development.

Some critical issues remain challenging, including

- Internal consistency;
- Measurement time horizon to measure life skills rather than moods or traits;
- Relying on self-report from the girl questionnaire; and
- Adequately contextualizing the girl questionnaire for all locations.

In addition to addressing these issues, going forward, we are adjusting assessment timelines to allow for more reflection after each administration, testing new scales and testing self-administration with a larger sample. We seek to better align measurement constructs to the life skills definitions, which may help us with the overlap of life skills found in the domain analysis (Annex 5: Domain Analysis). We expect to eventually identify a smaller number of dimensions for each life skill.

While we have been measuring at the individual level, we are looking at the enabling environment level and interactions between them. Assessing the conditions in which adolescents learn life skills may in some cases be more appropriate, meaningful and actionable than assessing learning outcomes at the individual level.

Lastly, we want to be careful how we make meaning out of the data, and how we analyze and interpret findings. We do this by either comparing treatment and comparison groups or pre-post evaluations. We look at the time horizon issues and distributions while being aware of a higher desirability bias at baseline than two to three years later at endline when girls better understand the concepts they are being asked about and their position vis-à-vis those concepts and also have more comfort in sharing their growing self-awareness. We are considering how to interpret findings when we see a decrease or no change in life skills scores, and which domains are most affected by these phenomena.

1. Introduction

This report synthesizes Room to Read's efforts from 2015 to 2020 to develop, test and utilize an adolescent life skills assessment (ALSA) as an outcome measure for its Girls' Education Program (GEP). Life skills are critical building blocks supporting the wellbeing of children, youth and adults. Life skills are transferable and include skills such as self-confidence, creative problem-solving, and expressing and managing emotions.[4] Each life skill includes multiple attributes, which makes measuring them challenging.

Over six years, we have shared our experience and results in conferences, workshops and webinars. Building on this collaborative spirit, we have developed this synthesis report to detail the ALSA development process, progress and lessons learned throughout this six-year experience across seven countries: India, Tanzania, Cambodia, Bangladesh, Nepal, Sri Lanka and Vietnam (see Figure 4). We hope that this synthesis will help other development organizations, researchers, and evaluators in their efforts to develop ALSAs for their programs and contribute to the broader knowledge base about measuring life skills.

Figure 4. ALSA Piloting Timeline
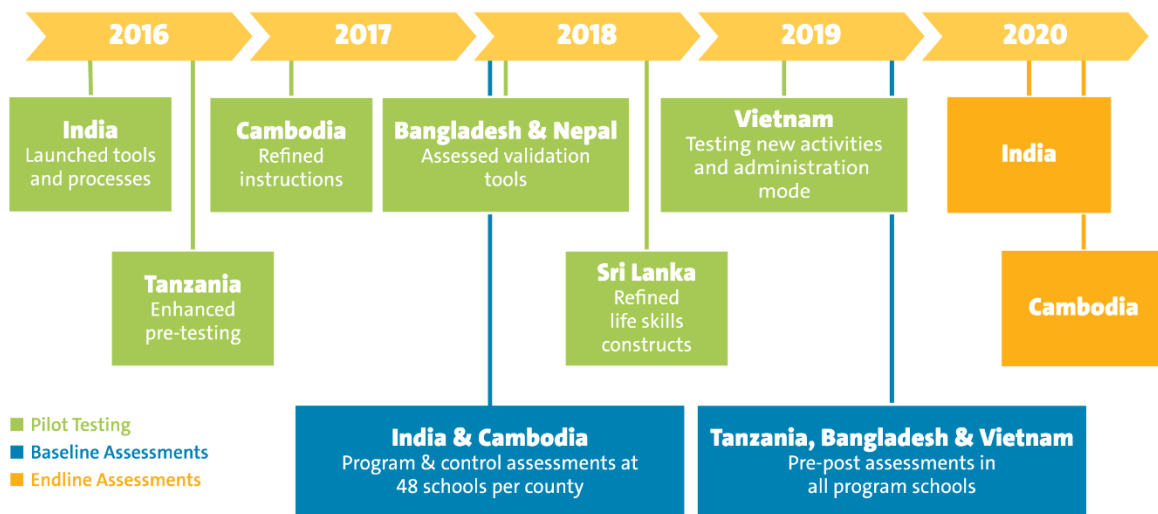
Since 2001, Room to Read's GEP has supported more than 140,000 secondary-school girls in nine countries through its school-based programming, and millions more via radio and TV broadcasts of life skills content, with the aim of helping them to complete secondary school with the skills needed to make key life decisions. In addition to life skills education, GEP provides mentoring, family, school and community engagement support, and need-based financial assistance. The program has established a standard curriculum to support girls in developing a set of critical life skills that we

---

[4] What we refer to as life skills, other organizations may label as socio-emotional, 21st century, soft or transferable skills. We use "life skills" throughout this report.

believe will enable them to thrive in their adult lives. Working with girls to enhance their cognitive, social and emotional skills supports their social progress and wellbeing in the present and throughout their lives.

Following work by the World Health Organization and others, the skills included in the curriculum are Self-Confidence, Expressing & Managing Emotions, Empathy, Self-Control, Critical Thinking, Decision-Making, Perseverance, Communication, Creative Problem-Solving and Relationship-Building (see Figure 5 below). Given the centrality of life skills to the program and its theory of change, we prioritized the development of a set of measures to assess girls' development of these skills as they move through the program.

Figure 5. GEP Life Skills Framework
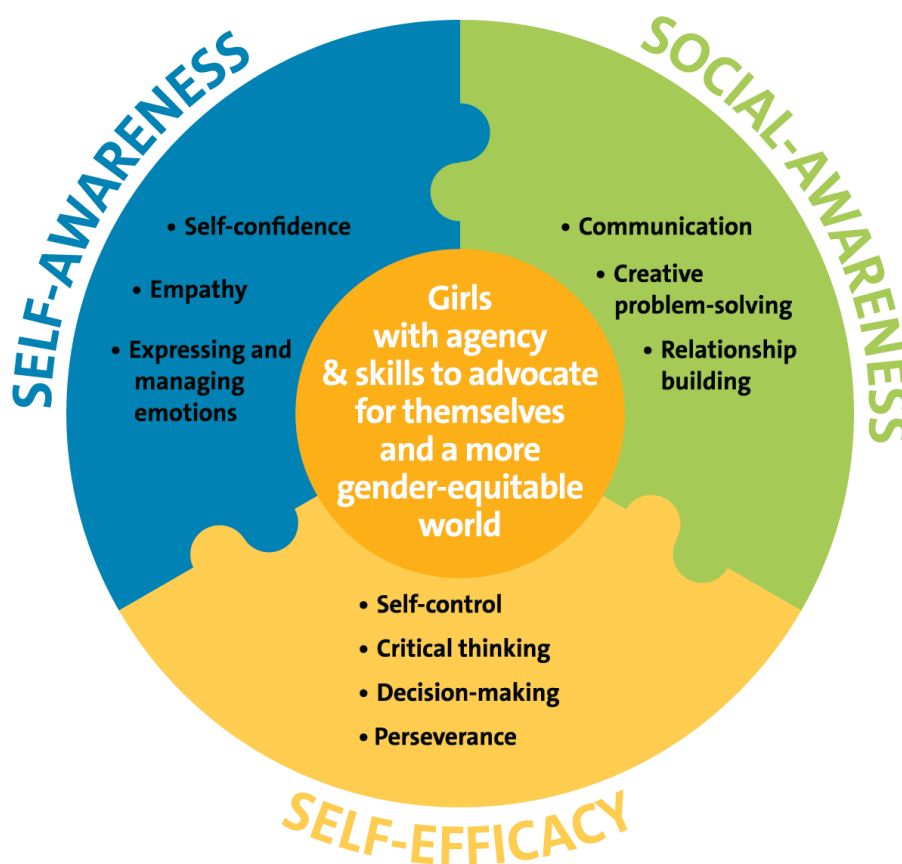
The report is organized to follow the ALSA tool development process (Figure 2 below) and includes the following key sections.

- *ALSA Development Process* describes our initial criteria for the ALSA, a summary of our tool review to identify existing life skills measures, our experience creating and piloting the first ALSA, and lessons learned from Pilot 1 in India.
- *Continuous Improvement through Pilots* describes how pilots influenced the adaption of the girl questionnaire, pre-testing, interviewer training and questionnaire implementation.
- *Continuous Improvement through Analysis* describes how we address reliability, validity, bias, administration mode and using the ALSA in GEP.
- *Room to Read's Pathway Forward* briefly describes our next steps in ALSA development and use.
- *Conclusion* summarizes focal lessons with the results of critical reflection on what we are happy about and what we wish we had done differently.
- *Annexes* provide supporting materials including the current girl questionnaire (Annex 1: Sample ALSA Tool) and first pilot girl questionnaire (Annex 2: Pilot 1 India), validation tools (Annex 3: Validation Instruments) and results from selected analyses (Annex 4: Additional Analyses and Annex 5: Domain Analysis).

2. ALSA Development Process

The development of the ALSA began in 2015 with a tool review to identify existing life skills measures that could be useful in a tool that would align with the needs of GEP. This chapter focuses on the first pilot in India, including a description of the first girl questionnaire, and is followed by Chapter 3 which focuses on lessons learned from subsequent pilots in six other countries (Tanzania, Cambodia, Bangladesh, Nepal, Sri Lanka and Vietnam).

Figure 6. ALSA Development Process
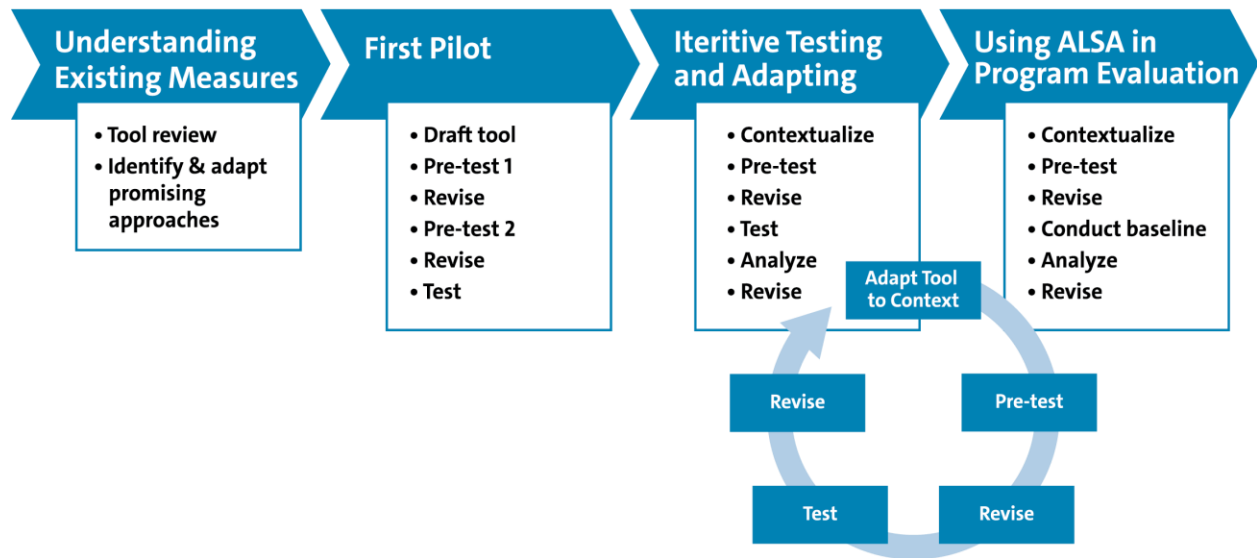
2.1 Establishing Criteria for the ALSA

To guide our tool development, we established a set of parameters for ALSA development:

- It should produce data that can be quantitatively analyzed
- It should be comprehensive and aligned with curriculum and skills building
- It should be valid and reliable
- It should preclude "teaching to the test"
- Its measures should be common skills across contexts
- It should be sensitive to changes in life skills over time
- It should be feasible to adapt across country contexts, with some adaptation
- It should be streamlined and not resource-intensive
- It should be appropriate for Grades 6 – 12
- Its administration should entail limited displacement of instructional time

The ALSA development team included a research leader, methodological advisors and life skills curriculum advisor in addition to a project leader and Room to Read's Senior Director of Research, Monitoring and Evaluation.

2.2 Understanding Existing Life Skills Measures

Understanding Existing Measures	First Pilot	Iteritive Testing and Adapting	Using ALSA in Program Evaluation
• Tool review • Identify & adapt promising approaches	• Draft tool • Pre-test 1 • Revise • Pre-test 2 • Revise • Test	• Contextualize • Pre-test • Revise • Test • Analyze • Revise	• Contextualize • Pre-test • Revise • Conduct baseline • Analyze • Revise
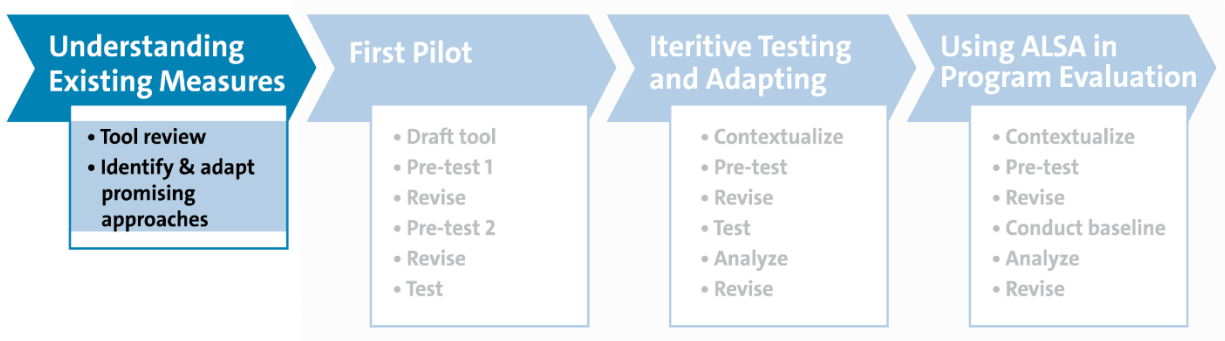
We began the ALSA development process with a review of existing measures that had been tested for reliability and validity in contexts that were as close as possible to GEP contexts. Measures also needed to correspond to the skills taught by GEP.

Table 3 describes the skills of interest and sources of measures that most closely met the criteria. The sources are publicly available tools covering related topic areas. For some life skills such as self-confidence and empathy, we found many measures. However, we found few self-confidence measures designed for children and youth. Almost all the measures found that met most design parameters were tested and validated in the U.S. and other wealthy country contexts, rather than in developing countries. As a result of this finding, we adapted tools for pilot settings with consideration for respondents' age, education, culture, environment and language. We found no viable measures for critical thinking, as most existing tools were for adults and computer based, so the ALSA to date contains no questions specifically designed and tested for that life skill.

Table 3. Description of Life Skills Measurement Sources

Room to Read Skill Area	Source Skill Area(s)	Source(s)	Testing
Self-Confidence	Self-Worth	Positive Youth Development	US, adolescents ages 10+
	Sense of Self	Jamaica Youth Survey	Jamaica, ages 12-18
Expressing & Managing Emotions	Self-Awareness/Self-Management	AIR & CASEL	US, youth
Empathy	Empathy	Child Trends-Flourishing Children	US, teenagers
	Empathy	Positive Youth Development	US, adolescents ages 10+
Self-Control	Domain-specific Impulsivity	Tsukayama, Duckworth, Kim	US, students in Grades 5 & 6
Decision-Making	Responsible Decision-Making	AIR & CASEL	US, teenagers
Perseverance	Grit Scale (child & adolescent)	Duckworth & Quinn	US, adults & adolescents
Communication	Communication	AIR & CASEL	US, youth
	Communication	Positive Youth Development	US, adolescents ages 10+
Relationship-Building	Friendships	Child Trends-Flourishing Children	US, adolescents
Creative Problem-Solving	Challenging Situation Task (inspired)	Denham, Ji, Hamre	Challenging situation task used in pre-school found valid with only 3 items. Inspired RtR to adapt it for GEP.
	Problem-Solving Task	Programme for International Student Assessment	Multiple countries, 15-year-olds

2.3 Creating and Piloting the First ALSA

Understanding Existing Measures	First Pilot	Iteritive Testing and Adapting	Using ALSA in Program Evaluation
• Tool review • Identify & adapt promising approaches	• Draft tool • Pre-test 1 • Revise • Pre-test 2 • Revise • Test	• Contextualize • Pre-test • Revise • Test • Analyze • Revise	• Contextualize • Pre-test • Revise • Conduct baseline • Analyze • Revise
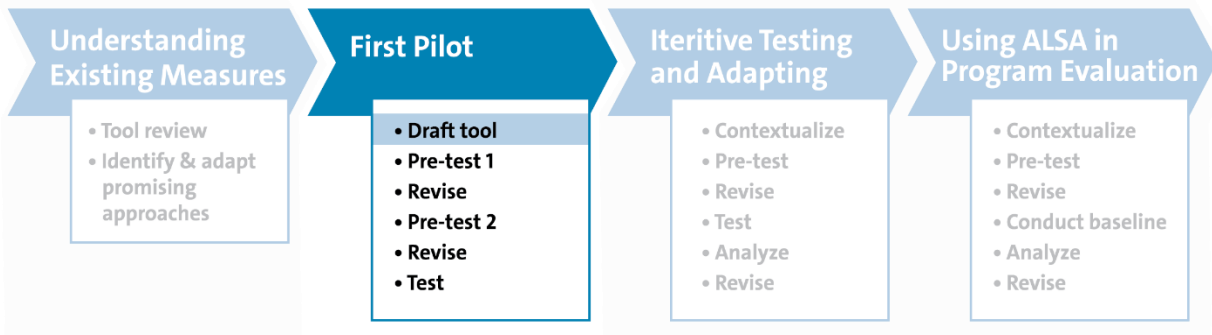
The ALSA is a work in progress. Each pilot includes a self-report questionnaire for girls and at least one validation method outside of the girl questionnaire, such as social mobilizer ranking or teacher's survey (see *Chapter 4: Continuous Improvement through Analysis*). The focus of this section is the first piloted girl questionnaire.

The first pilot was preceded by: 1) discussions among experts on how to adapt measures identified in the tool review into an assessment that would fit GEP and girls in Grades 6 to 12; 2) discussions to adapt the tool for the India context; 3) translation, review and revision; and 4) two pre-tests with revision after each. We created the first girl questionnaire by adding explanations prior to life skills sections, newly-developed scales and exercises, new items and demographic questions.
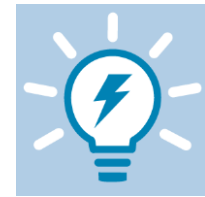
Build on others' work, adapt & test

The next section will further discuss the girl questionnaire and how early challenges affected questionnaire development.

2.3.1   The First Girl Questionnaire

The first piloted girl questionnaire included demographic questions plus about 125 mostly closed-ended items with a Likert scale, free listing and other exercises, forced choice questions and anchoring vignettes (see Annex 2: Pilot 1 India). Due to the Likert scales used in the source tools ranging widely, for consistency, we changed all answer options to a four-point scale. Thus, the response options to statements about which girls could choose as like (or not like) them were *completely true, very true, slightly true* and *not true*, while response options to statements about their perceptions were *strongly agree, agree, disagree* and *strongly disagree*.

The first girl questionnaire consisted of the following modules, most of which provide "I" statements with Likert scales. These are referred to below as "I" statements.

- Demographic module
- Life skills modules
    - *Self-Confidence*: 22 items covering: Measure 1 which provides two descriptions of kids and asks which the respondents is like and then asks if that is very true or sort of true for the respondent; and Measure 2 which contains "I" statements.
    - *Expressing & Managing Emotions*: 13 items covering: Measure 1 which contains "I" statements; and Measure 2 which asks respondents to identify their emotions over different time periods (listing exercise).
    - *Empathy*: 8 items including "I" statements.
    - *Self-Control*: 14 items with "I" statements.
    - *Decision-Making*: 14 items including "I" statements.
    - *Perseverance*: 13 items including "I" statements and a mirror drawing exercise. The mirror exercise was designed to assess the respondent's perseverance by asking her to draw simple to complex images without looking directly at her hand, pencil or paper while drawing; instead, she could see them reflected in a mirror.
    - *Communication*: 10 items with "I" statements.
    - *Relationship-Building*: 23 items with "I" statements and asking for peer nominations in four different scenarios.
    - *Creative Problem-Solving*: 8 items including three hypothetical situations asking girls to suggest solutions (listing exercise) and one mapping exercise, which asked respondents to determine a route between two points and distance (in kilometers).

Each of the life skills scales included multiple questions to be tested. Most but not all of the questions for each life skill are in a single module. Even prior to piloting, we recognized that life skills may conceptually overlap with one another, such as Perseverance with Self-Control and Empathy with Expressing and Managing Emotions. Given our conceptual understanding, we continually examine the inter-relationships between life skills constructs and the items within those constructs (see, for example, *Annex 4: Additional Analyses* and *Annex 5: Domain Analysis*).

Identifying and Addressing Bias

Bias may include tendencies to give extreme, mid-point or socially desirable responses or be subject to acquiescence or the halo effect, etc. We attempted to identify and address bias by using items within the life skills modules such as **free listing exercises**, which asked girls to: 1) think about and name the strong emotions they had experienced in the previous seven days; 2) suggest solutions to problems; and 3) name up to five classmates based on group activities or one classmate based on one-to-one activities. The latter, called a peer nomination exercise, asks questions such as: "The teacher is asking you to work on a project with a team of students, whom would you pick?" for which the respondent can name up to five peers, and "The teacher is asking you to work on a project with one student, whom would you pick?" and the respondent can name one peer.

The Creative Problem-Solving measure included three hypothetical conflicts: 1) between sisters; 2) between students in class; and 3) between a girl and her father. In each case, the girl is asked to answer the question, "What could [person's name in the situation] do?". After the respondent mentions all the solutions she can think of, she is asked which of the solutions she would carry out.

We included anchoring vignettes as another way to address response biases and reference group effects that delimit comparability of rating scales as understood by different respondents.[5] When the life skills constructs are reliable and valid, item values can be re-scaled by applying non-parametric anchoring vignette scoring. The vignettes, addressing each of the following skills, were grouped into their own modules rather than with the life skills that they address, as recommended by Hopkins and King.[6]

- Expressing & Managing Emotions
- Empathy
- Self-Control
- Decision-Making
- Perseverance

Forced Choice Module
Similarly, seven forced choice questions were used to address response bias[7] resulting, for example, from social desirability bias or acquiescence. The questions force respondents to choose from one of four desirable options. We intended to use them in the first pilot to examine response bias in some of the life skills scales. Since the answer options corresponded to the life skills, the questions were asked in their own module, separate from the life skills modules (see *Section 4.3.4: Forced Choice Questions*).

---

[5] King, G. and Wand, J. 2007. "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes," *Political Analysis*, Volume 15, Issue 1, Winter 2007, pp. 46 – 66.
DOI: https://doi.org/10.1093/pan/mpl011; Kyllonen, PC., and Bertling, JP. 2013. "Innovative Questionnaire Assessment Methods to Increase Cross-Country Comparability" in Rutkowski, L. (Ed.), von Davier, M. (Ed.), Rutkowski, D. (Ed.). (2014). Handbook of International Large-Scale Assessment. New York: Chapman and Hall/CRC, https://doi.org/10.1201/b16061.
[6] Daniel Hopkins and Gary King. 2010. "Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability." Public Opinion Quarterly, Pp. 1 – 22. Copy at https://j.mp/2nRql51.
[7] See Brown, A., Maydeu-Olivares, A., 2011 (https://doi.org/10.1177/0013164410375112) for the use of forced choice questions in identifying response bias using a multidimensional item-response theory model based on Thurstone's framework for comparative data.

2.3.2   First Pilot Administration
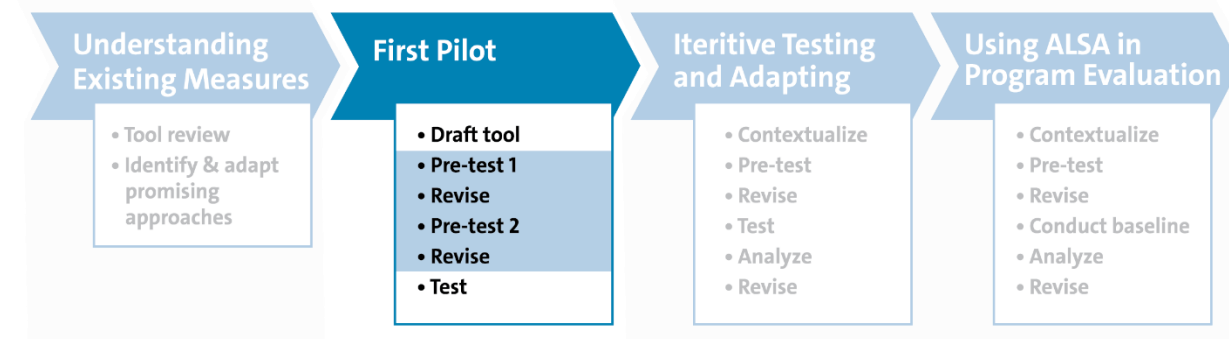
Besides drafting the girl questionnaire, the first pilot was preceded by translating the questionnaire into Hindi, training interviewers, administering two pre-tests and obtaining ethical clearance, permission from school authorities and parental consent. Ethical clearance for conducting the study was obtained by the International Institute of Health Management Research (IIHMR) from the Institutional Review Board (IRB) for Research Activities. IIHMR managed the survey process in India. Following permission from school authorities and principals, and obtaining parental consent, interviewers explained the survey to selected students, and after obtaining assent from each girl, conducted the face-to-face interviews on paper. Figure 7 illustrates the operational approach used for the first pilot.

Figure 7. Operational Approach to the Study
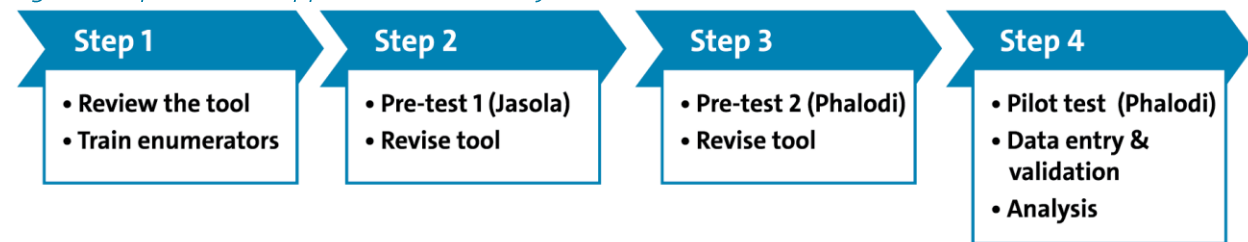
After discussing the tool—including appropriate Hindi words and simplification of some items—12 interviewers were trained over two days. We determined that all interviewers should be women since the girl respondents would likely be more comfortable being interviewed by a woman. Training (in Hindi) consisted of explaining study objectives, how to conduct the interviews, do's and don'ts during data collection, and research ethics. This was followed by detailed discussions of each question, role play and a pre-test at a government school in Jasola, Delhi.[8] After the pre-test, nine interviewers were selected to continue. Based on the pre-test and interviewer feedback, some words or phrases were simplified while others were reworded for better comprehension.

---

[8] The pilot data were collected in Rajasthan, but for logistical reasons, we selected a GEP school in Delhi for the first pre-test. This location did differ from the pilot data collection site in some important respects related to language and setting. This is one reason it was important to pre-test in Phalodi as well.

A second pre-test was conducted in Phalodi, Rajasthan, also with nine randomly selected respondents. During the pre-test, interviewers asked girls to rephrase a question in her own words, which helped us to know that for this statement: "I feel my life has a purpose," younger girls did not understand the concept of purpose, and older girls thought this was: "I know what I want to be when I grow up." Subsequently, the item became: "I think there is a reason for my life."

Interviewers identified questions that respondents found difficult, and words were identified in the local Rajasthani language (Marwari) to be used if a respondent had difficulty understanding Hindi. Items needed to be short and simple in English, as the Hindi equivalents tended to be longer. The Hindi translation was sometimes too formal and thus difficult for respondents, so the vocabulary needed to be further simplified. The following is an example of how a shorter and more concrete statement was better translated and subsequently better understood by respondents. This self-confidence item was used after revising Option 1: *We'd like you to decide whether you are more like the first child or the second child: 1) Some children like the person they are; 2) Other children often wish they were someone else;* we changed "some children like the kind of person they are" to "some children like the person they are."

Example of Simplified Wording: Decision-Making Item 2 was: "If I can't figure something out, I try different solutions until one works." The first phrase was changed to "When I have a problem …" because the translation of the original was difficult and not understood by respondents during pre-testing.

Even after simplifying and rephrasing to make items easier to understand, we found that respondents still struggled with non-standard answer options, such as those found in the forced choice section. For each question, respondents must choose which of four items is most like her and least like her (see *Section 4.3.4: Forced Choice Questions* for details). They therefore needed to understand and memorize all four options, for example:

Enumerator read: "I am going to read four statements, please tell me which one is the most like you, and which one is the least like you"

Statement	Response
A. I am satisfied with myself	
B. I can control my temper	Most like her: |____|
C. I am hard working	Least like her: |____|
D. I am good at telling others my ideas and feelings	

We addressed this issue by printing each series of four items on separate pieces of paper so that the interviewer could point them out to the girl while reading each item and answer option, and let the girl read it once more before she made a choice. For girls who could not read the items, interviewers were instructed to repeat the sentences two or three times. Interviewers were given the following guidance:

- Read all sentences slowly one by one, pointing to them on the printed version.
- Pause, ask the girl whether she understands all sentences or if she wants you to repeat/explain one of them (if the girl is very young, you can ask whether she needs explanation after each question).
- Repeat the sentences that raise issue; if the girl has trouble with a particular word, find synonyms; if it is still not understood, explain using the explanations discussed in training.
- Once the girl understands, ask her which statement is the most like her. Note the answer.
- Ask which statement is the least like her. Note the answer.

Following the two pre-tests and corresponding revisions, we conducted the pilot with 357 girls participating in GEP in nine government schools in Phalodi and Lohawat Tehsils in India. GEP participants in school on the day of the assessment were randomly selected with the goal of interviewing 35 − 40 students per school with stratification across grade levels. The purpose of the pilot and survey procedures were explained to the sample of students in group settings where they could ask questions. Each interviewer individually sought the student's assent. When a student hesitated or seemed concerned with going forward, the social mobilizer spoke with the student in the local language. The interview proceeded only if the student assented without pressure.

Each interviewer was supervised by core study team members who observed the interviews and took notes. Daily team debriefs resulted in problem identification, word changes and reordering of life skills sections. The survey firm performed spot checks for missing data at the school while interviews were being conducted and entered data into a spreadsheet. Data validation checked for allowed values for each item.

2.3.3 Testing the Girl Questionnaire's Adaptation and Administration
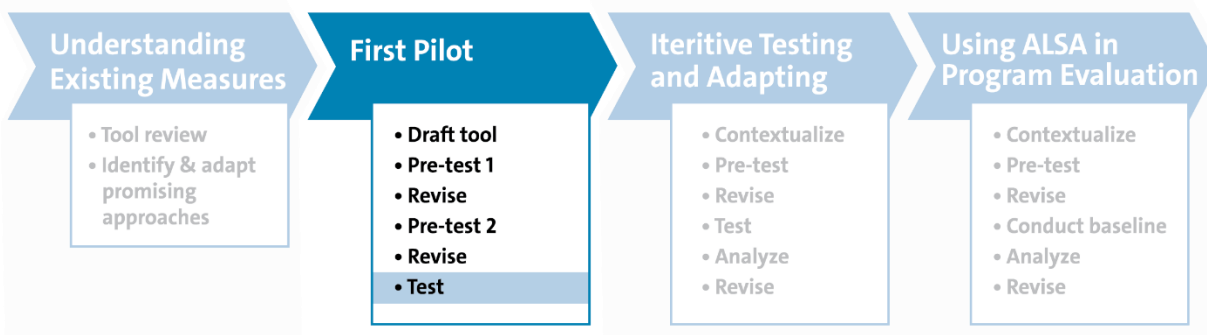
We planned to conduct numerous tests with data from the girl questionnaire in order to further refine it, including: 1) reliability test using Cronbach's alpha to estimate how consistent the items

are within each life skill scale; 2) confirmatory factor analysis of individual items to see how they load on the life skills construct (or scale); 3) correlation analysis among the constructs and with external measures for validation; and 4) distribution analysis primarily using skewness and kurtosis to ensure the tool adequately distinguishes between girls with high and low levels of the various life skills.

Reliability

"Reliability" here means that the questions for each life skill hold together and correlate well with one another, suggesting that the various questions are getting at the same underlying construct. In psychometric tests like the items used in the ALSA tool, reliability, or the internal consistency of the set of items for each life skill, is often estimated using Cronbach's alpha. Cronbach's alpha is a function of the number of items in a test, the average covariance between item-pairs and the variance of the total score. The resulting alpha coefficient ranges from 0 to 1. If all the life skill items are entirely independent from one another (i.e., are not correlated or share no covariance), then $\alpha = 0$; and, if all the items have high covariance, then α will approach 1. In other words, the higher the alpha coefficient, the more the items have shared covariance and measure the same underlying concept.

Although the standards for what makes a "good" alpha coefficient are somewhat arbitrary and depend on a theoretical knowledge of the scale in question, many methodologists recommend a minimum alpha coefficient between 0.65 and 0.8 (or higher). Alpha coefficients that are less than 0.5 are typically considered to have unacceptably low internal consistency. However, as noted above, low alphas can also be driven by the number of items and the distribution of scores and need to be considered in that context. An alpha above 0.85 may mean that there are too many items in the scale measuring the same construct, i.e. that some items are nearly duplicative and could be removed for greater time efficiency.

In addition to Cronbach's alpha, we examined data distributions. We looked at skewness, or symmetry of the distribution, and kurtosis, or the height and sharpness of the peak of the distribution relative to the rest of the data. We chose rule-of-thumb cutoffs as close to 0 for skewness and around 3 for kurtosis. These statistics helped us decide which estimators or tests should perform best with data distributed like the pilot data. High kurtosis alerted us to outliers in the distribution tail(s).

Validity

We used multiple measures over time to assess if the tool was measuring what we intended to measure (known as construct validity; see *Section 1.4.2: Addressing Validity*Addressing Validity for more). In the first pilot, the approaches to assess validity included: 1) a mirror drawing exercise; 2) a parent questionnaire; and 3) confirmatory factor analysis.

Bias

We included the following in the girl questionnaire to assess bias: 1) a mix of positive and negative worded items; 2) free list exercises; 3) anchoring vignettes; and 4) forced choice questions.

For a deeper discussion of the ways in which we assessed and improved ALSA tool reliability, validity and bias, see *Chapter 4: Continuous Improvement through Analysis*. For a snapshot of the modules and third-party surveys used during pilots, see Table 31. ALSA Development over Time on page 62.

2.3.4 Summary of Pilot Results

In the India pilot, Room to Read sought to test the whole process, including adaptation, interviewer training, data collection and analysis. We wanted to understand which of these elements were most challenging, including particularly to get an idea of how much adaptation would be required to use the measures across very different cultural and socioeconomic contexts from those in which they had previously been used and validated.

The India pilot provided valuable information on each stage of the process that informed the later pilots. We learned that thoughtful and careful translation, pre-testing and revision are insufficient to fulfilling our goal of having a reliable girl questionnaire. During the pilot, we continued to adjust the questionnaire, such as shifting a challenging module to later in the survey and providing respondents with show cards. We learned that adapting would take far more effort and resources than we put into the first pilot.

With regard to the life skills measures themselves, the most significant finding was that as implemented, they showed low reliability. For example, girls who rate themselves highly on one question related to Self-Control should be more likely to rate themselves highly on another question related to Self-Control. Where this is not the case, it suggests that the two questions measure different things or are not well understood by respondents. Due to such issues, we ran our basic measures including Cronbach's alpha, kurtosis and skewness for the life skills measures, but no other analysis. Data from India showed low alphas across all measures (see Table 1). Note that this analysis includes only those skills tested using a Likert scale.

Table 4. Cronbach's Alpha for Life Skills Measures (Pilot 1 - India)

Life Skills with Likert Scales	Item #s	Alpha
Self-Confidence – 1	A1-A6	0.39
Self-Confidence – 2	A7-A16	0.50
Expressing & Managing Emotions	B1-B6	0.41
Empathy	C1-C5	0.44
Self-Control	D1-D8	0.40
Decision-Making	E1-E5	0.26
Perseverance	F1-F10	0.51
Communication	G1-G10	0.50
Relationship-Building – 1	H9-H16	0.65
Relationship-Building – 2	H17-H23	0.48

The most straightforward interpretation of this finding is that respondents did not understand the questions as intended despite the revisions made during and after pre-testing. Interviewer feedback suggested that girls misinterpreted or misunderstood the meaning of items. The Hindi translations were sometimes too formal and not well understood by some respondents. Respondents did not always recall the response options by the time they understood the question.

Regarding Relationship-Building, the measures we tested included peer nomination against which we tested Measures 1 and 2. We found very low correlation, which given the low alphas, was difficult to interpret. Creative Problem-Solving was tested through a different approach based on the Challenging Situations Task (CST)[9] and the mapping exercise from the OECD PISA survey.[10] As noted in *Section 2.2*, these exercises involved presenting girls with challenges to solve. The results of the hypothetical situations (I1-I3) suggest the items are consistent in measuring the underlying construct.

Overall, we concluded that translation into a new language and/or adaptation to a new cultural context were more challenging than initially anticipated.

2.3.5   Girl Questionnaire Changes
The girl questionnaire was revised following the findings from the first pilot to: 1) improve respondent understanding by adding practice questions, instructions, altering response options and adding and cutting items; 2) add items to be tested to improve reliability; and 3) add validation measures. Most of the items added were sourced from the surveys identified in the tool review.

Improving Respondent Comprehension and Ease
One of the ways we improved respondent understanding was by adding a set of practice questions to familiarize respondents with the style of questions and the scale/response options, as shown on the following page.

---

[9] Denham, S, P. Ji, and B. Hamre (2010). "Compendium of Preschool Through Elementary School Social-Emotional Learning and Associated Assessment Measures." [Retrieved August 2021 from: https://casel.org/compendium-of-preschool-through-elementary-school-social%E2%80%90emotional-learning-and-associated-assessment-measures/]
[10] OECD (2004), "Problem Solving for Tomorrow's World: First Measures of Cross-Curricular Competencies from PISA 2003," OECD, Paris. [Retrieved December 2020 from:
http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/34009000.pdf]

Enumerator instructions: The following questions are for the girl to learn how to answer the questions. Please do not rush through these questions; make sure she understands this before proceeding.

ENUMERATOR READ "During the time we are going to be together, I am going to read some statements. Please indicate whether each statement is not true, slightly true, mostly true or completely true for you. Before this starts, we will first try some practice questions, so you know how to give your answers."

Enumerator instructions: introduce the scale and show the scale to the girl with the gradation in different options.

ENUMERATOR READ: "There are no right or wrong answers; this is only about what is true for you. If you don't understand a question, you can always tell me, and I can explain more. If you don't want to answer, just tell me and I'll move to the next question "

	Not True	Slightly True	Mostly True	Completely True
I like candy.	1	2	3	4
I spend time with my friends.	1	2	3	4
I know how to speak Arabic.	1	2	3	4
I like waking up early in the morning.	1	2	3	4

Another change that improved respondent understanding was replacing "very true" to "mostly true" as noted in the above scale.

To improve respondent ease with the questionnaire, we moved a difficult life skills module (Self-Confidence) to nearly last among the life skills.

Changes Made to Improve Life Skills Measures
We reversed the Likert scales to start with "not true" or "strongly disagree" rather than end with them in an effort to reduce large right tails in the data that suggest social desirability bias.

We also dropped the mapping question for Creative problem-solving, as it appeared to measure mathematical ability more than Creative problem-solving.

2.3.6   Girl Questionnaire Adaptation and Administration Changes

As noted in the summary of results from the India pilot, adaptation of the life skills measures to the girls' context was likely the primary driver of low reliability scores. So much so that some planned analyses were dropped for the first pilot (e.g., we could not use the vignettes for anchoring in the absence of reliable life skills construct data). Our adaptation of the girl questionnaire required more attention to translation and contextualization. To accomplish this, we increased resources in subsequent pilots for adaptation and administration, as described in Section 3 below.

> Observation is a valuable tool. During pre-testing in Pilot 2, for example, the social mobilizers noticed that "mostly true" was better understood than "very true" in the response scale and we saw an increase in response variation suggesting a decrease in acquiescence.

3. Continuous Improvement through Pilots
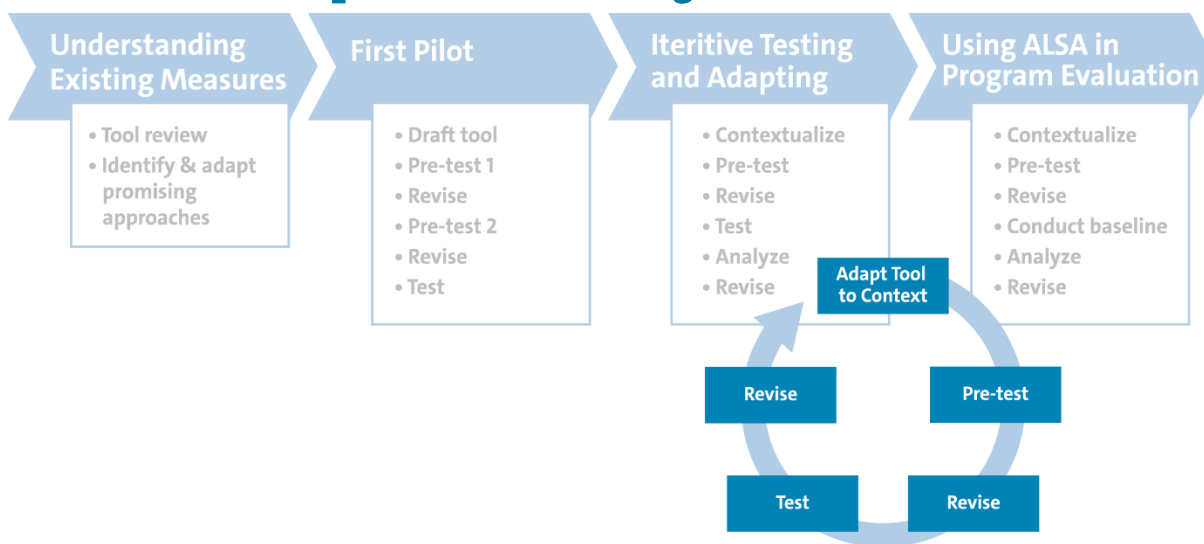
Following Pilot 1, we focused our attention and resources more on adapting the girl questionnaire for the next two pilot locations (Tanzania and Cambodia). As noted above, we looked to change the life skills constructs based on what appeared a poor fit for the study population and added items we considered during the design of the first ALSA tool but had not tested. While keeping in mind how interrelated validity and reliability are, we sought to further test the life skills construct validity by adding items from a behavioral questionnaire geared to self-reporting by children 11 to 17 years

called the Strength and Difficulties Questionnaire (SDQ).[11] The SDQ has been administered and validated in several countries. It addresses five scales: emotional symptoms, conduct problems, hyperactivity, peer problems and prosocial behavior, some of which we used to compare with our life skills scales.

This section explores the iterative learning process including what we learned about adapting the girl questionnaire, pre-testing it, training interviewers and implementing the girl questionnaire. The next step of learning through data analysis is detailed in *Chapter 4: Continuous Improvement through Analysis*. As noted in Figure 4 on page 1, the second pilot greatly enhanced pre-testing and the third pilot further refined the instructions in the questionnaire. Following these pilots, the questionnaire was much more reliable.

3.1 Adapting the Girl Questionnaire
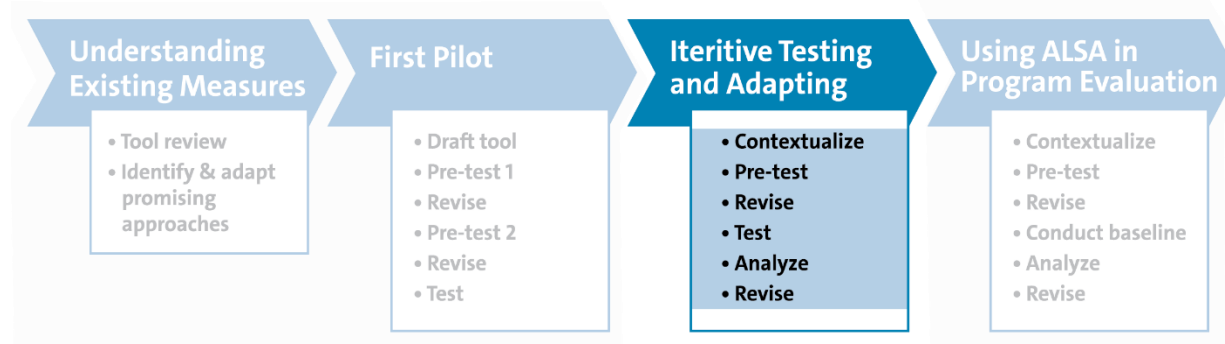
In addition to our ALSA development team, the team responsible for contextualizing the ALSA includes the Research, Monitoring and Evaluation Lead and GEP team in each of our country offices. The GEP team includes a Manager or Director who leads the country-level program, Program Officers and/or Associates, and social mobilizers (or mentors) who liaise with GEP girls, their families and schools. Since social mobilizers are well known to families and schools, they work one-on-one with girls during pre-testing to learn how the girls understand the questions. This is especially important as we learned in the India pilot that close attention to word choice and how girls understand items and key words is critical to a successful survey.

The ALSA team reviews the existing girl questionnaire in detail guided by the questions below, which serve as an adaptation guide for the girl questionnaire. Some of the most challenging efforts include adapting the girl questionnaire to how skills manifest in each context, and to survey language(s), including finding synonyms for interviewers to use when respondents do not understand.

---

[11] See the Youth in Mind website: https://www.sdqinfo.org/py/sdqinfo/c0.py for SDQ tools and information.

Questions to discuss for each life skill (or skill group[12]) include:

- What does this life skill mean?
- How does this skill represent itself in your context?
- How does the tool seek to measure this?
- Does this measurement strategy make sense? Are there critical elements of this skill that are missing?

Questions to discuss for each scale or activity include:

- Will girls understand the instructions the interviewer provides to them?
- After hearing these instructions, will the questions/activity make sense to them?
- What further explanation might be needed to make sure girls understand?

Questions to discuss for each item include:

- How does this item relate to the life skill it is intended to measure? Is it a good measure of this skill?
- Is it locally relevant? Does it reflect the way this skill represents itself in your context?
- Is there any cultural sensitivity that needs to be considered in this question?
- Can all terms in the question be easily translated into the language(s) without changing their meaning?
- Once translated, will girls easily understand these terms? If not, are there synonyms that can be used for those girls who need more explanation?

As the team discusses the girl questionnaire item by item, they take notes and document decisions. This documentation is important to understanding how and why the pilot girl questionnaire changes as we hope to standardize it as much as possible across contexts.

Once the team is satisfied with the questionnaire, they have an independent translator translate the tool from English to the appropriate language(s), and then have it back-translated by a different translator to English. The team reviews the resulting versions, discusses differences and reconciles them prior to pre-testing.

Pilot 1 taught us that getting the "right" words to fit the context and our study population is critical. While adapting the girls' self-report tool for Tanzania, we struggled with finding the right words for the Expressing and Managing Emotions life skill, because there are two words that convey "emotions" in Kiswahili as spoken by this population. One of the words is understood as very general, including not just emotions, but physical feelings such as hunger, pain, etc. The other word

---

[12] Note that in the tool there are two groups of skills that have been combined because of the degree to which they overlap conceptually: 1) Perseverance and Self-Control; and 2) Empathy and Expressing & Managing Emotions.

technically means emotions, but in common usage in the communities where the ALSA was piloted, it has a strong sexual connotation.

To resolve this, we introduced that life skill section with a highly abbreviated version of the "Feeling Words" module from the Life Skills curriculum. We also added some items and dropped "I pay attention when I do things" due to very little response variance and poor fit in the construct. The following is how the life skill was measured in Pilot 2.

In this section, we are going to talk about emotions. Emotions are also called mood and feelings here. Can you think about the feelings you had recently? Over the past seven days, you were at school; you were at home; you interacted with people. In all these different situations, your feelings/emotions may have changed. Can you remember some of them? Can you name them? (Probe: Okay, what other feelings/emotions can you remember?)

B10	1 Happiness 2 Sadness 3 Loneliness 4 Anger 5 Jealousy 6 Stress 7 Fear	8 Worry/Anxiety 9 Excitement 10 Pride 11 Sympathy 12 Regret 13 Guilt	14 Other: 14_1 _____ 14_2 _____ 14_3 _____							
B11	How many emotions or feelings did the girl list? (Enumerator: Do not ask, just circle the correct number)		0	1	2	3	4	5	6	7+
B12	Did it seem difficult for the girl to come up with an answer? (Enumerator: Don't ask this question, just circle the correct answer)		1. Very hard		2. Hard		3. Easy		4. Very easy	

Thank you. Emotions are what guide your reactions to situations. If you like the situation, you will feel happy, and you may smile or laugh. If you don't like the situation, you may feel angry or sad.

Here are some faces that show different people with a variety of emotions. Can you link the following emotions to the corresponding face? A) Angry; B) Surprised; C) Sad; D) Happy.

1.   2.   3.   4.
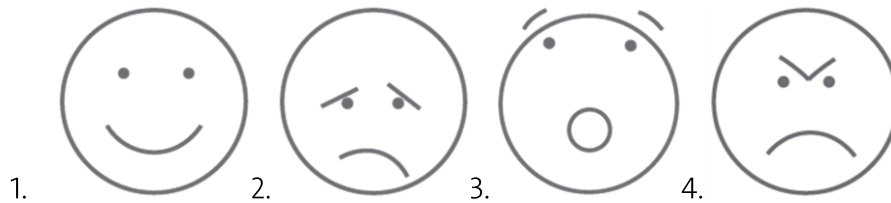
ENUMERATOR READ: Very good! There are a lots of emotions and feelings. Someone can feel lonely, excited, nervous, relieved. All of these are emotions. Now, I am going to read some statements about things that you may feel or do. Please tell me for each of these statements how much the following is true for you. Is it not true for you, slightly true for you, mostly true for you or completely true for you?"

		Not True	Slightly True	Mostly True	Completely True
B1	I understand my moods, feelings and emotions.	1	2	3	4
B2	I can name the feelings I have.	1	2	3	4
B3	I can explain why I am feeling a certain way.	1	2	3	4
B4	I understand why I react the way I do to different situations.	1	2	3	4
B5	When I'm upset, I understand how I am feeling before I take action.	1	2	3	4
B6	My moods and feeling can change my behavior.	1	2	3	4
B7	When I feel nervous or stressed, I know how to calm or relax myself.	1	2	3	4
B8	I know what makes me happy.	1	2	3	4
B9	I can describe my most and least favorite parts of each day.	1	2	3	4

In Pilot 3 (Cambodia), we pre-tested a new introduction to the scales, as they continued to challenge respondents. We took what we learned and revised the introduction to the questions which resulted in the following introduction.

Enumerator instructions: The following questions are for the girl to learn how to answer it. Please do not rush; make sure she understands the scale perfectly before proceeding.

Read: "During the time we are going to be together, I am going to read some statements. I'll ask you to indicate whether each statement is not true, slightly true, mostly true or completely true for you. First, we will first try some practice questions so you know how to give your answers."

Enumerator instructions: Introduce the scale and show the scale to the girl with the gradation in different options.

Read: "When I read a statement, you will have to think if this sentence applies to you. If you think: "this is not me, I am not this person at all, this isn't true for me" then you should select "not true." If you think that this is a little bit like you, but only to a small extent, you should select "slightly true." If you think that most of the time you're like this person/this sentence applies to you, you should select "mostly true." Finally if you think that you're like the person described by the sentence, then you should select "completely true."

Let me tell you an example on how to select the answer. For example, if I say "I like the color blue." You should think about yourself, and how does this sentence apply to you. If you think it has nothing to do with you (in our example, if you don't like blue), you should select "not true." If you think this sentence has some element of truth for you (you find blue okay, but you much prefer yellow and pink), you should select "slightly true." If you think that this sentence is true for you most of the time (you like almost all kinds of blue, but there are some kinds of blue you don't like), you should select "mostly true." Finally, if you think this sentence is always true for you (blue is your favorite color), you can select "completely true."

	Not True	Slightly True	Mostly True	Completely True
I spend time with my friends. (*Probe: "think about yourself, how does this sentence apply to you? Is it "not true," "Slightly true," "mostly true," or "completely true"?*)	1	2	3	4
I like ice cream. (*Probe: "why did you give that answer? how often do you eat ice cream?"*)	1	2	3	4
I know how to speak French.	1	2	3	4
I like waking up before five in the morning.	1	2	3	4

In addition to challenges around translations of concepts like emotions, we learned that despite pre-testing, sometimes girls in one context understand items as they are intended, and in another they do not, despite intense care to adaptation, translation and pre-testing. For example, in Vietnam, this pair of positive and negative items: "If someone I trust asks me to do something I know is wrong, it is easy for me to say no" and "If someone I trust asks me to do something I know is wrong, it is hard for me to say no" resulted in data opposite of what was expected. Girls appeared to find the positive behavior (easy to resist peer pressure) undesirable and the negative behavior (hard to say no to peers) desirable. Possible reasons for this include translation issues, vocabulary issues or cultural factors. See *Section 4.3.1: Positive and Negative Items* for more on using these items to address bias.

> "The ALSA tool was developed for global fit, so it's very important to country to contextualize. Even the questions, when asked in English may convey a different meaning than in a local language. In pre-tests, girls understood some questions differently than we expected. We had to do a lot of translation back and forth. It's very important to back-translate. Beyond language issues, conditions such as the education system may be different, so it's important to make sure the wording is locally relevant."
>
> -GEP Staff

3.2 Pre-Testing

Rigorous pre-testing is critical to the success of building a new tool. We conduct a pre-test after adaptation and translation and prior to each pilot. The pre-tests are conducted by our social mobilizers who administer the girl questionnaire while simultaneously conducting cognitive interviews with respondents to assess their understanding of and responses to the assessment.

Cognitive interviews using think-aloud and probing questions during pre-testing proved to be a productive approach for improving the girl questionnaire. They helped to identify misalignment between items as intended versus the respondents' understanding of those items. This process helped us to identify when concepts and language needed to be clarified or simplified and how to explain and train girls on the scales so that they understand them.

We use probes during the pre-test cognitive interviews, such as:

- Can you tell me a little more about your answer?
- What did the question mean to you?
- What does [key word/phrase] mean to you?
- What were you thinking about as you answered that question?

Asking questions like these is especially important when a girl hesitates to answer, looks confused or absent, or answers in an unexpected or illogical manner. We improved the girl questionnaire by pre-testing not only questions but also words/phrases and ways to introduce the scales. The introduction to the scales described above was adapted from the previous pilot and then tested

with groups of girls by reading the introduction and asking who understood and who did not. We asked one girl who understood to explain it to another who did not, while the rest of the group listened.

Following the pre-test, the team discusses girls' understanding of the items and words, the translation(s) used and any confusion that arose during the focus groups and/or interviews. They revise the girl questionnaire accordingly. Typically, the team makes any needed changes to the translations at this stage, but if the changes are lengthy and significant, we recommend another round of independent translation and back-translation.

3.3 Training Interviewers

Based on our experience, we learned that at least around four days of interviewer training is necessary to fully prepare interviewers to conduct an efficient and high quality ALSA, especially the girl questionnaire. Although the girl questionnaire is the core of the ALSA and the longest tool, the ALSA also includes other data collection tools (see *Annex 3: Validation Instruments*). Training must include all the tools included in one's ALSA.

Interviewer training includes:

- Introduction to the study topic and objectives
- Detailed review of the instruments in the ALSA pilot with special focus on understanding each item in the girl questionnaire
- Synonyms for key words and creation of an approved list of synonyms for the girl questionnaire
- Ethical conduct and assent processes
- Building rapport and establishing trust
- Roles and responsibilities, logistics and timeline
- Interview role-play with the girl questionnaire

Interviewers need to understand the study topic, objectives, the girl questionnaire and the assent process. They must be prepared to establish trust and a good rapport with respondents and effectively use the questionnaire by asking all questions and describing activities as they are written. The only deviations permitted are to use approved synonyms. Having found the latter to be a necessary part of successful interviews, we established a process during training to identify

Protecting children
Protecting children is a primary responsibility that all Room to Read contractors are required to follow per our Child Protection Policy. Each country office's policy is slightly different in order to align with relevant local laws. We train interviewers on human subjects research ethics. All of our survey protocols, including tools and consent and assent language and processes, are annually reviewed by the WCG IRB (Institutional Review Board).

synonyms for key words. Otherwise, if a respondent is having difficulty, the interviewer can restate the question or use a different tone of voice. These requirements are critical to respondent comprehension and maintaining survey fidelity.

While training interviewers—observing and encouraging good administration habits—we help them to use the right voice and attitude throughout survey administration. This includes seeing trust building as a process rather than a one-time exercise at the beginning of the survey. Building trust can reduce social desirability bias. We believe that continually reinforcing confidentiality in training has helped to normalize score distributions.

Role-playing serves several purposes: 1) interviewers develop fluency, becoming at ease with the questions, activities and the instructions they will give respondents; 2) they may discover reasons to discuss a word/phrase translation or meaning again and revise the synonym list and/or ALSA translation; and 3) opportunity for the study team to observe interviewers and make decisions about who to hire for data collection. We train enough interviewers so that any dropping out does not affect the data collection timeline, and we engage the best candidates following a review of their performance during training.

> "When training enumerators, one needs to provide a lot of detail: what they should and shouldn't do, and how to make girls feel comfortable during the interview and ensure proper understanding. Take time to let the enumerators practice using the questionnaire. Practice with each other, but also with girls in non-program schools."
>
> -GEP Staff

We recommend that each trainee conduct one to two interviews with GEP members as part of their training since good performance in role play does not assure good performance during data collection. This is especially important when working with children and youth.

With the above efforts, we lock the tool once pilot data collection begins so that the data are based on the same instrument across all schools.

3.4 Implementing the Girl Questionnaire

Good implementation means more, of course, than survey fidelity. It includes a smooth operation that fulfills the research protocol as reviewed and approved by an IRB.

Research Ethics

As noted above, our survey protocols are reviewed by an institutional review board. We obtain parental consent and student assent prior to interviews. Interviewer training is intensive and includes role-playing practice interviews. The research team and interviewers maintain confidentiality of ALSA participants including girls and other respondents.

Confidentiality is an agreement to protect the privacy of research participants. Protecting privacy means that interviewers keep information about respondents secure, allowing only the research team access to information about respondents and their responses. Maintaining confidentiality increases the level of trust that respondents have with interviewers, protects the dignity of respondents, and allows interviewers to get more honest information. Violating the privacy of respondents could result in physical and/or social harm to respondents or their families.

Ethical review and approvals are obtained by Room to Read's Global Office for all our life skills assessments through a global IRB. The country office teams are responsible for obtaining informed assent from the girls and consent from their parents.

Examples of how interviewers can protect respondent privacy are:

- Not talking with anyone outside the study team (including respondent's relatives) about the study
- Not taking photos that would allow the respondent to be identified
- Ensuring that documents such as consent, assent, sampling lists and research notes are stored in a secure place
- For paper-administered surveys, following safety protocols and submitting paper versions daily to a supervisor

The one exception to confidentiality is to report abuse of someone under age 18 by following the study's communication protocol.

Communication Protocol

Following the study's communication protocol includes everything from reporting on a prescribed schedule to a team leader to daily data back-up to following verbal and non-verbal communication training guidelines with respondents. Some of the most critical for ALSA participants are behaving and speaking in ways that convey respect, active listening and non-judgment.

The communication protocol changes from country to country, but always includes at a minimum the basic requirements above.

Roles and Responsibilities

Smooth implementation means good communication and clear roles and responsibilities. The Country Office Research, Monitoring and Evaluation Lead has ultimate responsibility for implementing the ALSA. They choose, for example, who to hire as data collectors (e.g., interviewers or a survey firm). As noted above, social mobilizers have existing relationships with students' families and schools, so they conduct the parental consent process and obtain permission from school authorities (see Table 5).

Table 5. Roles and Responsibilities for ALSA Adaptation and Implementation

Staff	Roles & Responsibilities
ALSA Team Leader[13]	• Has ultimate responsibility for ALSA adaptation and implementation • Guides process and ensures its success • Supports team in their efforts • Engages interviewers or survey firm • Assures that adequate resources available • Consults with support staff as needed
Program Staff	• Discuss ALSA adaptation
Program Mentors (Social Mobilizers)	• Discuss ALSA adaptation • Obtain parental consent • Conduct pre-test and cognitive interviews with girls • Obtain school permission for pilot/baseline
Interviewers or Survey Firm	• Implement ALSA • *If a firm,* conduct training and quality control in conjunction with Room to Read, and prepare data for analysis.
Translators	• Translate ALSA to local language(s) • Back-translate to English *(separate translator from above)* • Repeat the above process if there are any major revisions of the ALSA
Other Potential Staff Needed	• Conduct data entry if the ALSA is done by hand with paper and pencil rather than with electronic data gathering tools (e.g., tablets) • Program tablet/device • Back check interviews and other data quality checks • Conduct data analysis

4. Continuous Improvement through Analysis

Reliability, validity, minimizing bias and good administration are essential for a high-quality assessment tool. We continue to improve the ALSA through ongoing testing, data analysis, and use in new and previously tested environments. This chapter describes our efforts to create a high-quality standardized girl questionnaire by addressing reliability, validity, bias and administration mode. Some of the methods we tested have worked better than others.

4.1 Addressing Reliability

To test the girl questionnaire's reliability, we use Cronbach's alpha and confirmatory factor analysis to explore internal consistency. We conducted test-retest on sampled girls to compare their answers between the two survey events. Throughout analysis, we look at results: 1) at the life skills level; 2)

---

[13] In Room to Read, this is the country RM&E Lead.

of individual life skills measures; and 3) of items within each life skills measure. We illustrate findings with a variety of measures at different levels of analysis.

4.1.1 Internal Consistency

The internal consistency of the life skills measures—how well the life skills constructs measure what we want them to measure—has improved since the first pilot as illustrated in the alpha scores for each measure (shown in Figure 3 on page iv and further detailed in Table 28 on page 58). We use confirmatory factor analysis for each life skills scale to determine how well each item in the scale loads on the factor, which is the underlying construct or life skill. High loading (defined here as 0.40 or better) means the item is relevant in defining the life skills construct.
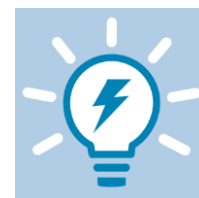
Check factor loadings and move items when appropriate

In Pilot 2, we sought to better understand which items within each life skills construct loaded high. In Table 6 below, we look at the Expressing and Managing Emotions construct to illustrate how we combine our conceptual understanding of the life skill and analytical results to build the measure. The results of Pilot 2 for this life skill showed significantly improved internal consistency indicating that girls replied in a coherent way to the items and suggesting that the items were better understood than in Pilot 1. We ran a confirmatory factor analysis with the intention of keeping items with factor loadings above 0.40. Although B6 and B7 load lower than the other items (highlighted), they were retained, as these items were the only items about *regulating* emotions, an important element of the construct that we did not want to drop entirely.[14]

Table 6. Confirmatory Factor Analysis for Expressing and Managing Emotions (Pilot 2)

Variable	B1	B2	B3	B4	B5	B6	B7	B8	B9
Factor 1	0.62	0.69	0.65	0.68	0.55	-0.34	0.39	0.57	0.48

In another pilot, we tested different measures for Decision-Making and Expressing and Managing Emotions (see Table 7). We rejected the alternative measures.

---

[14] We later moved B7 to the self-control construct, where it has been a better fit.

Table 7. Results of Multiple Measures Tested (Pilot 5)

Construct	Content	Reliability (alpha)	Skewness	Kurtosis	p-value of Normality Test
Self-Confidence	A1-A12	0.54	-0.29	2.67	0.09
Self-Confidence (no negative items)	A1-A8	0.59	-0.45	2.38	0.01
Self-Confidence (only negative items)	A9-A12	0.59	-0.66	3.43	0.00
Emotions: main	B4-B12	0.69	-0.28	2.37	0.24
Emotions: alternate	B13-B22	0.59	-0.05	2.29	0.66
Empathy	C1-C9	0.64	-0.51	2.84	0.00
Self-Control	D1-D12	0.54	-0.43	2.67	0.04
Self-Control (no negative items)	D1-D8	0.68	-0.81	3.12	0.00
Self-Control (only negative items)	D9-D12	0.55	-0.77	3.89	0.00
Decision-Making: main	E1-E7	0.45	-0.77	3.37	0.00
Decision-Making: alternative	E8-E11	0.41	0.32	2.79	0.33
Perseverance	F1-F6	0.66	-1.03	3.92	0.00
Communication	G1-G9	0.64	-0.52	2.58	0.00
Relationship-Building with friends	H1-H8	0.70	-1.18	4.24	0.00
Relationship-Building with adults	H9-H19	0.72	-0.87	3.51	0.00
Relationship-Building: measure 3	H20-H29	0.57	-0.49	2.80	0.00
Relationship-Building (no negative items)	H20-H29	0.71	-1.04	3.71	0.00
# of solutions in problem-solving	I1-I3	0.54	0.70	3.86	0.00

Results sometimes leave us with new questions. For example, life skills constructs that are stable and perform well in some ALSA administrations perform less well in others. When comparing the scores between pilot and baseline in Bangladesh, for example, of the nine life skills constructs, one-third performed similarly, another third better, and a final third worse (see Table 8).

A great deal has been written about Cronbach's alpha, its use (and overuse), and reliance on it, and what we can reasonably conclude based on it.[15] The fluctuation in alpha scores over the pilots and baselines suggests challenges adapting the girl questionnaire for each context within a country. We reflect more on why some life skills' internal consistency fluctuates within a country between the pilot and baseline in *Section 6.1 Critical Reflection*.

Table 8. Bangladesh Pilot and Baseline Alpha Scores

Construct	Pilot Alpha	Baseline Alpha
Self-Confidence	0.71	0.63
Expressing & Managing Emotions	0.67	0.66
Empathy	0.72	0.70
Self-Control	0.62	0.62
Decision-Making	0.61	0.58
Perseverance	0.63	0.67
Communication	0.73	0.65
Relationship with Friends	0.83	0.87
Relationship with Adults	0.80	0.84

Along with Cronbach's alpha, we looked at Guttman's lambda to cross verify reliability scores (see Table 9 on the following page). Data distributions continue to skew to the right. To understand what is happening within each scale, we conducted confirmatory factor analysis. It showed negative items loading low and negative: the Self-Control scale, for example, includes four negative items (see "Factor Analysis: Example from Bangladesh Baseline" in *Annex 4: Additional Analyses*). The Perseverance construct, consisting of five items, with adequate loading (between 0.43 – 0.66) but a somewhat low alpha (0.67) suggests that the number of items may be insufficient.

---

[15] See for example Taber, K.S. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Res Sci Educ 48, 1273 – 1296 (2018). https://doi.org/10.1007/s11165-016-9602-2.

Table 9. Reliability Scores in Bangladesh Baseline

Construct	α	λ-1	λ-2	λ-3	λ-4	λ-5	λ-6
Self-Confidence	0.63	0.57	0.65	0.63	0.77	0.63	0.67
Expressing/Managing Emotions	0.66	0.57	0.66	0.66	0.74	0.64	0.65
Empathy	0.70	0.62	0.71	0.71	0.79	0.69	0.71
Self-Control	0.62	0.57	0.64	0.63	0.74	0.62	0.65
Decision-Making	0.58	0.50	0.58	0.56	0.69	0.56	0.57
Perseverance	0.67	0.54	0.68	0.68	0.77	0.67	0.65
Communication	0.65	0.60	0.69	0.67	0.78	0.68	0.68
Relationship with Friends	0.87	0.77	0.88	0.88	0.90	0.85	0.87
Relationship with Adults	0.84	0.76	0.85	0.84	0.89	0.82	0.85

4.1.2 Test-Retest

We conducted a small sample test-retest in Pilot 5 (Nepal). Test-retest is a common practice to validate a tool by piloting twice with the same respondents, in this case with 40 randomly sampled girls (26% of the total sample) with two and three weeks between interviews. We compared the test and retest data by survey section and counting the number of discrepancies. We expected differences since the questionnaire measures psychological aspects which may vary to some extent from day to day.

Table 10. Test-Retest Discrepancies (Pilot 5)

Dimension	Variables	# Comparisons	# Discrepancies	% Discrepancy
Background Questions	Z1-Z8	320	0	0%
Self-Confidence	A1-A12	480	209	44%
Expressing/Managing Emotions 1	B2-B12	440	136	31%
Expressing/Managing Emotions 2	B13-B22	400	144	36%
Empathy	C1-C9	360	180	50%
Self-Control	D1-D12	480	246	51%
Decision Making 1	E1-E7	280	68	24%
Decision Making 2	E8-E12	200	64	32%
Perseverance	F1-F6	240	99	41%
Communication	G1-G9	360	156	43%
Relationship-Building	H1-H29	1160	466	40%
Peer Nomination	HB1_1-HB7	760	490	64%
Problem-Solving	I1_1_1-I3_2	1640	184	11%
Strength and Difficulties	SDQ1-SDQ13	520	248	48%
Enumerator Rating	ENUM1-ENUM4	160	86	54%
Vignettes Empathy	Canch_3-Canch_1	120	50	42%
Vignettes Self-Control	Danch_2-Danch_1	120	62	52%
Forced Choice	FC1_MOST-FC6_LEAST	480	276	58%

We then tested the correlations for the life skills measures (Pearson's correlation). We found that at the aggregate level, most constructs are highly and significantly correlated with the exception of Decision making (see aggregate results in Table 11 on the following page and item-level results in Table 41 on page 127). The relationships illustrated in these results caused us to ponder if we are measuring life skills as a skill (malleable), a mood (changes daily) or a trait (to some degree

immutable). We recognize that life skills manifest themselves on a continuum that may fluctuate in a short time if they are malleable, in a medium timeframe if they change slowly, or in a long timeframe or never if they rarely change.

Table 11. Test-Retest Correlations (Pilot 5)

Dimension	Correlation Coefficient
Self-Confidence	0.69
Expressing and Managing Emotions 1	0.59
Expressing and Managing Emotions 2	0.73
Empathy	0.48
Self-Control	0.52
Decision Making 1	0.30
Decision Making 2	0.21
Perseverance	0.63
Communication	0.50
Relationship-Building 1	0.53
Relationship-Building 2	0.29
Relationship-Building 3	0.27
Problem-Solving Average Number of Answers	0.57
Strength and Difficulties Questionnaire 1	0.44
Strength and Difficulties Questionnaire 2	0.48
Strength and Difficulties Questionnaire 3	0.48

4.1.3 Next Steps to Improve ALSA Reliability

We continue to improve scales based on an extraordinary amount of data collected since 2016. As a result of our desire to have pilot data from multiple contexts, we did not always have enough time to analyze, reflect and adjust between one pilot and the next. Instead, we have deeply reflected on results from several pilots together, including in workshops in 2018 and 2020. Our plans for 2020 were affected by the global COVID-19 pandemic, resulting in delayed or truncated data collection, data analysis and reporting. We expect to eventually identify a smaller number of items for each life skill measure as we validate the measures and find them reliable. See, for example, Annex 5: Domain Analysis. We look forward to catching up with our timeline as the pandemic abates and periodically updating this report with new findings.

4.2 Addressing Validity

We have sought to assess and improve several types of validity during our development of the ALSA. One of these is *content validity*—the degree to which each of our measures fully encompasses its intended domain—which in our case starts with having based most of the life skills constructs on validated measures and carefully adapting them based on expertise in survey and item design, girls' education and local contexts.

Face validity examines how well GEP girls (intended respondents) believe the survey questions measure the life skills they were intended to measure. The ALSA team used cognitive interviews—using probes during pre-tests and focus groups or one-on-one interviews with GEP girls—to investigate how they understood the questions. Following the first pilot, these methods were used in every location as: 1) part of adapting the tool to the context; and 2) to standardize the tool across contexts. We believe the increase in pre-testing greatly contributes to ALSA reliability and validity.

At what point are we satisfied with a life skill construct?

This is a question we need to assess with each ALSA administration. For example, after Pilot 5, the Self-Confidence scale correlated well with social mobilizer ranking (0.25***), interviewer feedback (0.25***), and forced choice rank (0.23**), but not with teacher results (0.20). The measure continues to perform well, sometimes significantly correlated with other measures, but not with all other measures that we might expect it to in any given administration. We have not yet reached a definitive conclusion regarding how to interpret such cases.

We have also examined ALSA tool validity by comparing girls' self-reports with external data and others' perspectives, task-based measures and statistics such as the results of confirmatory factor analysis to determine if the items for each life skill construct are measuring the construct (underlying factor).

In each country, we learn from pre-testing and implementing the ALSA. We typically pre-test in one location. We have learned, however, from district-level analysis of pilot data that there can be differences in understanding between girls in different locations. For example, when we dug into the data by district in Pilot 6 (Sri Lanka), we learned that for most of the life skills, the correlations (and thus, validation results) are brought down by one northern district, while data collected in a different district had better psychometric properties. This could be due to translation/adaptation issues and/or cultural differences in respondents leading to bias in the northern district data. These results may indicate tension between our desire to standardize the ALSA and respondent comprehension across contexts within a country, as well as between countries.

In this section, we explore our efforts to validate the ALSA using third-party assessment, task-based measures with girls and external data such as girls' grades.

4.2.1 Third-party Assessment

Over the past six years, we have tested girls' self-reports with reports from parents, teachers, GEP social mobilizers, interviewers and peers (see *Annex 3: Validation Instruments*). Third-party assessments for triangulation with girls' self-reports can be useful but also present challenges, as third parties should only be asked about that which they know. Parents may or may not, for example, be aware of their daughters' emotions and opinions, or some of their behaviors outside the home environment. In addition, third parties also have their own biases.

We found that teachers, social mobilizers and interviewers show the greatest promise for third-party validation (see Table 13 later in this section). We dropped the peer nomination exercise in the girls' self-report tool, which demonstrated some promise but was difficult to administer; we also had concerns about its validity, as popularity among girls may influence whom they nominate.

Below we describe our experience with some of the tools we have used.

Example: Teacher's Survey Results (Pilot 5)

Six teachers filled out a reduced version of the Devereux Student Strengths Assessment (DESSA), each for 10 girls. We created several indicators including Self-Confidence, Perseverance (in DESSA language "goal oriented"), Communication/cooperation, Relationship-Building and Creative Problem-Solving. Based on previous experience and pre-testing, we excluded the DESSA items around Empathy, Decision-Making and Emotional Awareness, as we found that teachers do not know the girls well enough to answer in those areas. We also added items describing anxious behavior to create an anxiety indicator.

As shown in Table 12, indicators generally have good internal reliability based on Cronbach's alpha. Only the anxiety index is weak and has an asymmetrical distribution. Yet when we compared these measures with relevant life skills, we found either little correlation or, in one case, a puzzling negative correlation between teachers' perspective and girls' self-report. For the negative correlation (related to relationship-building with friends), we hypothesize that what girls perceive they need to do to make friends may be perceived as negative behavior by teachers.

Table 12. Results of Teachers' Assessment

Teachers' Measures	Alpha	Skewness	Kurtosis	Correlations with Life Skills Measures	Significance Level
Self-Confidence	0.83	-0.09	2.48	• Self-Confidence	None
Anxiety Index: Girl is afraid, worry, nervous or shy	0.38	0.83	3.24	• Expressing & Managing Emotions • Self-Control	None -0.23**
Perseverance/Goal-oriented	0.86	-0.23	2.29	• Decision-Making 1 • Decision-Making 2 • Perseverance	0.28* None None
Communication, Cooperation	0.64	0.05	2.23	• Communication	0.24*
Relationship-Building, Social awareness	0.65	-0.15	2.05	• Relationship-Building with Friends	-0.30**
Creative Problem-Solving	0.65	-0.4	2.30	• # of Solutions Given by Respondents	0.29**

Example: Addressing Bias in Social Mobilizer Tool

One of the most useful tools asks social mobilizers—who know the girls well—about their perceptions of their students. We address potential bias with social mobilizers who may want to influence the ALSA results by asking social mobilizers to rank girls along four dimensions: Empathy, Emotional Awareness, Self-Confidence and Communication. We ask each social mobilizer to rank, for example, two groups of 7 to 10 girls who are randomly selected among their students. By ranking girls, social mobilizers are forced to give an opinion on them and provide implicit information about girls' weaknesses.

In contrast to the teacher and social mobilizer tools, we dropped the parent tool because results showed that parents were biased toward their own daughters. In some contexts, parents' ability to understand and respond were hampered by low literacy and education, or by their awareness of only some of their daughters' behaviors in areas of interest. For example, parents had difficulty answering questions about their daughters' self-confidence using the DESSA parent-reported scale on items such as, "how often over the past six month did your daughter say good things about the future"/"show awareness of her personal strengths"/"keep trying when facing difficulties"? In addition, talking with parents required significant resources outside school settings given that interviewers talked with girls and other third parties at schools.

Table 13. Results of Using Third-Party Assessments

	Parents	Teachers	Social Mobilizers	Enumerators	Peers
Skill Areas Assessed	• Self-Confidence, Perseverance, Communication, Relationship-Building, Creative Problem-Solving	• Self-Confidence, Perseverance, Communication, Relationship-Building, Creative Problem-Solving	• Empathy, Emotional Awareness, Self-Confidence, Communication	• Self-Confidence, Communication	• Relationship-Building, Empathy
Advantages	• Know girls well (home context)	• May know girls well (school context) • May be less biased than parents • Literate and educated • Can be interviewed at school	• Know girls well (multiple contexts) • Understand relevant outcomes very well	• Understand relevant outcomes • Likely to be least biased • Easy to assess 100% of sample	• Know other girls well • Already being surveyed • Peer nomination generates multiple data points
Disadvantages	• Literacy & education levels may be low • Positively biased • Logistically challenging	• Knowledge of girls uneven and context-bound • May have limited time	• May be very biased (assume evaluation of own work)	• Very limited knowledge of girls	• Time-consuming and difficult to administer • Validity concerns around popularity
Results	• Reliability low • Correlations low	• Correlations in some areas (e.g., ~0.30 for creative problem-solving), weak or negative in others (self-confidence)	• Correlations in some areas (e.g., ~0.25 for self-confidence), weak or negative in others (empathy)	• Correlations highly significant even where low	• Correlations around 0.20 in relevant areas, highly statistically significant

4.2.2 Task-based Measures

Mirror Drawing Exercise

The mirror drawing exercise was a task-based measure we tried in the early pilots. As an exercise in the girls' self-report tool, it was meant to assess the respondent's perseverance by asking her to replicate images with pencil and paper without being able to look directly at her hand; instead, she could see it reflected in a mirror (see *Annex 2: Pilot 1 India*). We wanted to use the mirror drawing exercise results to compare with the Perseverance construct. The girl's effort and results in the exercise were observed and recorded by the interviewer. We also thought the exercise was a good break during a survey with mostly closed-answer questions with Likert scales.

The disadvantages of this exercise were: 1) additional training was needed for interviewers to understand how to set up the exercise and properly observe and record it; 2) it was time-consuming in an already lengthy survey; 3) the results were challenging to interpret, as we were not looking to test how well the girls drew; and 4) it was resource intense as the exercise required each interviewer to have a mirror, a stopwatch to time how long girls spent on each attempt and a way to block the respondent seeing her hand. It also increased interviewer training time. The results of the exercise showed little to no correlation (-0.06 to 0.09) between the Perseverance construct measure and the number of attempted drawings, total time spent and time spent on the last drawing. We decided after including the exercise in multiple contexts to drop it from the ALSA.

Roleplay Exercise

In the Vietnam pilot, we added a roleplay exercise to evaluate girls' life skills in real life situations. The Vietnam team designed six scenarios, each describing a situation that could occur in the girl's real life. In each situation, the girl was given an objective and then had to interact with a roleplay partner personifying a relevant interlocutor in the situation. This interaction was observed by a "rater," i.e., an external observer.

The roleplay exercise comprised four steps:

a. The girl was given a choice between two randomly selected scenarios and selected the one she preferred to play.
b. The girl had five minutes to read the scenario and prepare.
c. The girl and her roleplay partner played the scenario. The roleplay partner had instructions on how to behave, what to say, what his/her own objective was and what kind of emotions/attitude s/he should show.
d. The girl and roleplay partner debriefed on the roleplay with questions about the girl's emotional management and decision-making, which included a decision-making exercise in which the girl is asked to choose between two schools. Both raters took note of the questions she asked to understand her decision-making process.

An additional decision-making exercise was completed by the girl alone.

Both the observer and the roleplay partner independently completed a questionnaire about the girl's performance, such as how she spoke, stood, behaved, etc.

We first asked girls to tell us what emotions they felt during the roleplay and what techniques they used to control those emotions. We wanted to measure whether girls could express their emotions and demonstrate an understanding of emotional self-regulation techniques. On average, girls mentioned 1.3 emotions (1 in the non-GEP sample and 1.4 in the GEP sample) and 0.8 methods to manage them (0.5 in non-GEP sample and 0.9 in GEP sample).

The roleplay exercise was stressful for girls, as more than half said they were nervous (52%), nearly half (46%) said they were afraid and many said they were sad (30%). Only 16% said they were happy. In terms of how they felt physically during the roleplay, more than one in five said she quivered or shivered. We are reviewing how to implement the roleplay exercise without stressing respondents.

Girls in the roleplay mentioned that they mostly managed their emotions through movement, breathing and drinking water, as well as keeping a positive mindset, smiling and talking. Yet more than one in five girls did not mention any self-management techniques (34% in the non-GEP sample and 13% in the GEP sample). Regarding what girls felt their roleplay partner must have felt, on average, girls mentioned 0.7 emotions (0.6 in the non-GEP group and 0.7 in the GEP group). They named angry, sad, doubtful, sympathetic/understanding or ill-at-ease the most. We hope to use these data to measure empathy. Finally, we asked whether the girls knew any other solutions than the ones they came up with during the roleplay.

The raters scored girls in the dimensions of Communication, Expressing and Managing Emotions and Self-Confidence. The data show that raters mainly disagree on girl's stance, posture and gestures (see Table 14 where shaded cells denote a minimum of 25% disagreement between raters). This may be because one of the raters is the roleplay partner, who may find it difficult to adequately note the girl's physical movements while playing. Another hypothesis is that the roleplay partner is closer to the girl during the exercise and is a better observer of how she breathes, moves, etc. Another hypothesis is that the difference does not result from observation discrepancies but from subjective interpretation. We will continue to test roleplay-based approaches and include a debrief session between the raters to understand the source(s) of discrepancy.

Table 14. Discrepancy Between Roleplay Raters' Observations

Life Skill	Observation category	Yes	Raters Disagree	No
Com	Girl speaks in a steady tone of voice.	40%	25%	35%
Com	Girl articulates clearly so her words can be understood.	44%	22%	33%
Com	Girl asks one or more clarifying questions.	13%	14%	73%
Com	Girl mainly looks at roleplay partner in the face during the exercise	49%	30%	21%
Com	Girl gives relevant answers to questions	59%	22%	19%
Emotion	During or after exercise, girl names at least one feeling she experienced (apart from nervous)	86%	10%	5%
Emotion	Girl is able to list one strategy she used to manage her feelings.	59%	13%	29%
Emotion	Girl is able to list at least one feeling her roleplay partner might experience	86%	8%	6%
Emotion	Girl does not fidget with her hands	8%	29%	63%
Emotion	Girl breathes steadily	22%	33%	44%
Emotion	Girl's stance is steady, natural, relaxed	24%	19%	57%
Emotion	Girl does not repeatedly touch her hair/ear or nose	35%	29%	37%
Self-conf	Girl has relaxed posture (no shoulders bent) during the exercise	16%	29%	56%
Self-conf	Girl faces her body toward her roleplay partner.	54%	27%	19%
Self-conf	Girl is persistent in pursuing an acceptable solution to the problem	40%	13%	48%
Self-conf	Girl's resolution to the conflict reflects her self-respect	49%	14%	37%
Self-conf	When given a reason why she cannot have what she wants, she is able to offer new	46%	21%	33%
	Girl decided when the situation was resolved.	24%	27%	49%

Roleplay Decision-Making Exercise 1

For this decision-making exercise, girls were asked: "Imagine you need to make a decision about which school you will attend for high school. One school is very close by in your village. Another is in

the district capital, ~10km away." They were then asked: "Who are some people you could ask for advice?" Their final did not preference any particular decision; rather, the exercise sought to assess their decision-making process itself.

Girls were able to answer this question, typically with family members and teachers. They were given the opportunity to ask questions as part of their decision-making. They asked on average 4.8 questions to help them make the decision before actually deciding.

After deciding, they were asked to list:

a. The advantages of the school they selected
b. The disadvantages of the school they selected
c. The advantages of the other school

Based on the data, girls did not understand how to describe the advantage and disadvantages, e.g., some girls listed disadvantages when asked about advantages. The objective of these three questions was to see how far the girl went in weighing pros and cons before making a decision. Many girls did not or were not able to describe disadvantages of their selected school or advantages of the other school.

Roleplay Decision-Making Exercise 2
The final decision-making exercise was completed by the girl alone. The instructions were:

Marina wants to choose her future career. Help her organize her thoughts by putting the following in a reasonable chronological order:

A. Being a teacher allows me to stay close to here and to my family, and to work with children
B. My teacher thinks I would be good at being a doctor and could even become a specialist
C. I decide to become a teacher
D. Being a doctor has a better salary that being a teacher
E. I could be a teacher or a doctor
F. My parents told me they prefer me to be a teacher in the school because it will be safer
G. Being a doctor means working at night and in the hospital that is far away from my family's place

DM [___|___|___|___|___|___|___|]

In the GEP sessions, girls learn that they should first look at their options, list the pros and cons, ask for advice and then decide. We would therefore expect girls to put them in order as follows:

i.   start with E
ii.   select [D, A, G] with no ideal order
iii. select [F, B]
iv.   select C at the end (final decision)

This exercise proved extremely difficult for respondents. Yet when we created an indicator describing the order in which the girl organized A-G in the exercise with an increment of 1 each time the girl ordered them correctly (according to the steps i-iv described above), the resulting score was larger for the GEP than the non-GEP group, and it correlates significantly with the indicator from the first decision-making exercise.

Overall, the role play exercise may be a useful validation tool if used with the full ALSA sample. We found some low but significant correlations between results from the role play and life skills constructs, especially between the Relationship-Building ALSA scale and the Communication and Self-Confidence scores from the roleplay, though less so between Self-Confidence ALSA scale and the Self-Confidence score from the roleplay. We plan to revise the exercise and interviewer and rater training to improve the girls' experience and to further test the exercise as an ALSA validation measure.

There are also significant disadvantages to this exercise. It is very demanding on data collectors and challenging to create relevant scenarios and an objective scoring rubric. In addition, it takes a lot of time and the girls who participated in our first attempt were nervous during the exercise.

4.2.3   External Data

Comparing existing external data with primary data is a standard means of validating a new item or tool. In Nepal, we collected academic results for several subjects (Nepali, English, Math, Social Studies, Science, Health, Moral Education, Computer Science) and girls' mean total grades for comparison. We expected Self-Control and Perseverance to be correlated with school performance as measured by grades. We thought we might see a correlation between Communication and some subjects, and that Self-Confidence might be related to girls' grades.

We found data collection and reconciliation of the external data with the survey data using the girls' identification numbers to prepare the data for analysis to be time-consuming. In this initial pilot, we found a statistically significant correlation with our Self-Confidence measure but were not able to detect a relationship to our other skill measures (Table 15). Given the wide range of factors known to influence grades, to explore these relationships in greater depth we will likely need larger samples than the 142 girls included in this Nepal pilot, as well as additional data collection to support a multivariate regression analysis. We will investigate this further in future ALSA administrations.

Table 15. *Life Skills Measures' Correlations with Girls' Grades*

Life Skill	Grade (Mean)
Perseverance	0.06
Self-Control	0.01
Communication	0.05
Self-Confidence	0.14*

4.2.4 Next Steps to Improve Girl Questionnaire's Validity

To improve the validity of the ALSA, we are testing alternative task-based measures such as the roleplay exercise. We are improving the alignment of our life skills constructs with the skills definitions using resources like the Harvard Explore SEL Taxonomy Project. Sometimes, scales are built on other scales, which are in turn built on other scales, and the concept definition on which the first scale was based gets lost. We are in the process of revisiting our scales.

4.3 Addressing Bias

Variance and distribution are important because low variability among responses suggests response bias. A ceiling effect—a high percentage of the respondents selecting the highest option for an item or group of items—can be caused by what is defined in the scientific literature as "social desirability bias."[16] Overly positive responses may be due to girls' desire to please or impress the interviewer and to present themselves in the best possible light, as well as their fear of being judged. To minimize social desirability bias and address the low response variance observed in Pilot 2, in Pilot 3 (Cambodia) we added or strengthened introductions to each scale, reminding girls that there are no good or bad answers, and that we want them to be honest. We also rephrased or deleted items that seemed to be biased or judgmental.

Build on others' work, adapt & test

Below is an example of how we explored variance and data distributions between different measures of Decision-Making and the resulting decision to keep the first measure. The Decision-Making constructs in Pilot 5 contained 12 items. Items E1 to E7 belonged to the first measure, and E8 to E12 to the alternative measure. The alternative measure asked respondents "how difficult is it for you?" rather than "how true is it for you?." Girls selected an answer from "Very difficult, somewhat difficult, Neutral, Somewhat easy, Very easy." We hypothesized that we would get a higher variance with the alternative measure, as girls would better understand the "difficulty" perspective than the how "true" one.

We built the main measure by averaging E1 - E7, after removing E5 as it loaded very low on the main factor. Likewise, we removed E12 from the alternate indicator and computed it by averaging E8 − E11.

---

[16] Fisher, R. J. (1993). "Social desirability bias and the validity of indirect questioning." Journal of Consumer Research, 20, 303-315; Nederhof, A. J. (1985). "Methods of coping with social desirability bias: a review." European Journal of Social Psychology, 15: 263-280

Both indicators displayed a low alpha in this pilot, but the alternate measure displays a higher variance and a less skewed distribution (see Table 16 and Figure 8 below).

Table 16. Decision-Making Analysis Results (Pilot 5)

Indicator	Reliability (alpha)	Skewness	Kurtosis	Variance
Decision-Making, Main Measure (N=87)	0.45	-0.77	3.37	0.23
Decision-Making, Alternate Measure (N=69)	0.41	0.32	2.79	0.70

Figure 8. Distributions of Main and Alternate Decision-Making Measures (Pilot 5)
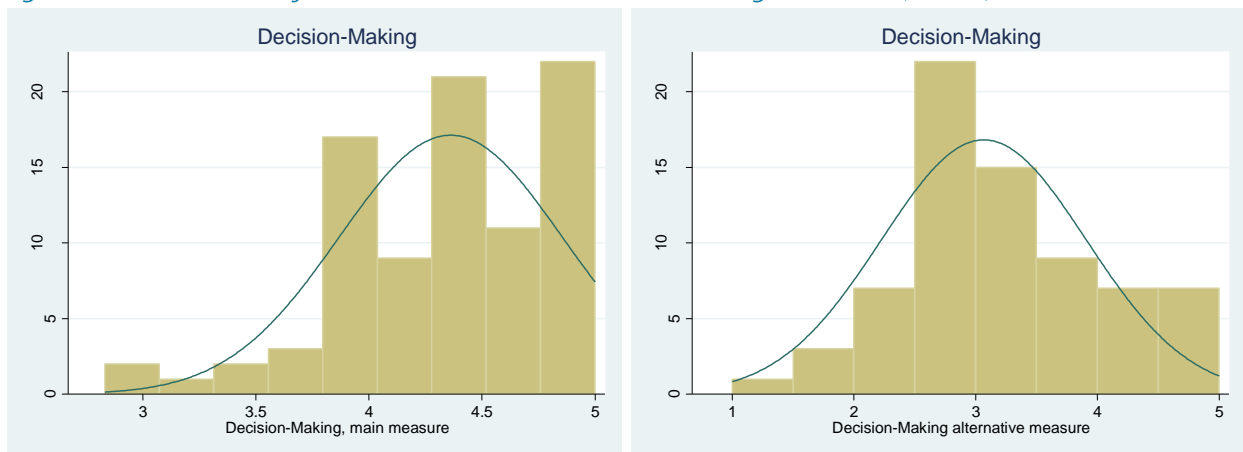
In terms of validity, the main measure is correlated positively and significantly with the teacher's indicator of girl's perseverance/motivation, and both correlate with the SDQ hyperactivity indicator (see Table 17).[17]

Table 17. Decision-Making Validity Test Results (Pilot 5)

| Decision-Making indicators | Decision-Making, Main | | Decision-Making, Alternative | |
	Correlation	N	Correlation	N
SDQ Hyperactivity	0.27***	87	0.31***	70
Girl Perseverance/Goal-oriented (teacher)	0.28*	36	0.26	13

* Correlation significant at 10%; *** Correlation significant at 1%

---

[17] Following several tests, we dropped the SDQ measures as they were insufficiently aligned with GEP life skills.

We also looked at the correlation between ranks from the forced choice questions pertaining to Decision-Making and the ranks obtained with the Decision-Making Likert scales (correlation coefficient = -0.02). The lack of correlation could be from bias, but we concluded it was more likely that the measure was not yet good enough, especially given the low alpha. Following this analysis, we chose to retain the response options related to how "true" a statement was for the respondents.

Decision-Making has proven the most difficult life skill to measure with consistency as seen by Cronbach's alpha across pilots and evaluations to date. As with our other life skills scales with negative items, those items continue to load low (or negative) on the main factor.

Test new items in multiple ways

We address potential bias in many ways including using positive and negative items, anchoring vignettes, free listing and forced choice questions. We look at the results of each in this section. Our findings show that transforming data to remove acquiescence bias improves the properties of some of the scales. We have not drawn a conclusion at this point, as we are still working to validate the scales.

4.3.1. Positive and Negative Items

Including positive and negative items may reduce acquiescence bias (the tendency for respondents to avoid disagreement with an interviewer's statement).[18] The following life skills contained negative items: Decision-Making, Self-Control, Self-Confidence and Relationship-Building. See Table 18 for the pairs of positive and negative items used in Pilot 6 (Sri Lanka).

---

[18] Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity and predictive power. *Journal of Personality and Social Psychology*, 113(1), 117 – 143. http://dx.doi.org/10.1037/pspp0000096.

Table 18. Pairs of Positive/Negative Items (Pilot 6)

Item	Life Skill	Item Text
E6		When I am around my friends, I change the way I act so that I will fit in.
E9		I do not change the way I act around my friends just to fit in.
E7	Decision-Making	If someone I trust asks me to do something I know is wrong, it is hard for me to say no.
E2		If someone I trust asks me to do something I know is wrong, it is easy for me to say no.
D3		It is hard to wait for something I want.
D7		It is not hard to wait for something I want.
D6	Self-Control	Many times, I lose focus before completing a task.
D12		I am able to complete tasks without losing focus.
D13		My parents tell me I do things without thinking.
D2		My parents tell me I am thoughtful when completing a task.
H8		It is hard for me to always be kind and caring with my friends.
H2	Relationship-Building	It is easy for me to always be kind and caring with my friends.
H11		I find it hard to make friends.
H5		I find it easy to make friends.
A11		I am afraid I am going to fail.
A3		I am not afraid that I will fail.
A12	Self-Confidence	I don't think I have many good qualities/skills.
A2		I feel that I have a lot of good qualities.
A13		Sometimes, I don't feel confident about myself.
A8		I feel confident about myself.

Addressing Bias by Transforming Data with Negative Items

We enhanced our use of the pairs by using them to balance the life skills scales: "Content-balanced scales would allow researchers to easily control for acquiescence at the item level by centering each individual's set of item responses on their within-person mean (i.e., their mean response to the full item set, without reversing the false-keyed items; see Soto et al., 2008)."[19]

We identified ten pairs of items consisting of one positive and one negative item with opposite implications for personality. For each respondent we created an **acquiescence score (AS)**: the mean

---

[19] Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity and predictive power. Journal of Personality and Social Psychology, 113, 117 – 143.

of their paired items (10 pairs of 2), and a **response extremeness score (RES)**: the standard deviation of their 20 responses.

We then applied the following transformation to our data where Xt is the transformed data and X is the raw data:

$$Xt = \frac{X - AS}{RES}$$

The results appear as "centered" and include the following (see Table 19):

- *Mean Score of All Transformed Items*: Average of all items transformed for a scale (row 3).
- *Mean Score of Selected Items Transformed*: Average of the highest loading items and then transformed (row 4).
- *Mean Score of Items Selected after Transformation*: Average of Items selected that best hold together after transformation.

The data without transformation are:

- *Mean Score of All Items*: Average of all items in each scale (row 1)
- *Mean Score of Highest Loading Items*: After checking for item covariance and running a confirmatory factor analysis, we decided to keep some items and average them within each life skills construct (row 2).
- *Factor Scores*: Factor analysis allows us to predict and save factor scores that can be understood as the underlying variable and used for validation.
- *Balanced Scores*: For scales with negatively keyed items, we built "balanced scores," averaging a similar number of positively and negatively worded items (row 5).

The results for Self-Control in the last pilot illustrate that items hold well together. The negative items continue to lower the reliability score. For most of the mean scores, the data are normally distributed (see normality test results in the last column).

Table 19. Self-Control Analysis (Pilot 7)

Variable	Reliability (alpha)	Skewness	Kurtosis	p-value of Normality Test
Mean of all items	0.66	-0.36	3.21	0.11
Mean of best items	0.68	-0.43	3.37	0.08
Mean of all items centered	0.73	-0.11	2.78	0.60
Mean of best items centered	0.71	-0.21	3.10	0.51
Balance scale mean of items	0.47	-0.34	3.18	0.23

In the most recent pilots—Pilot 6 in Sri Lanka and Pilot 7 in Vietnam—negative items tend to load on a separate factor than the rest of the items in the life skills scales and/or load low. The literature on using positive and negative items offers no solution, as it too illustrates similar challenges. Colosi (2005) found that using negative items introduced new error where respondents are inconsistent between negative and positive items and more respondents say "don't know" to negative items than to positive ones.[20] Salazar (2015) found less internal consistency when including positive and negative items. He also notes that some authors have found that respondents with little education have more difficulty with negatively worded questions. As we also found, the presence of both positive and negative items decreases acquiescence bias, which is the purpose of using them, but decreases the factorial validity to the factor structure.[21]

This creates a conundrum for researchers and ALSA users: the positive/negative items decrease the presence of bias and the fit of the items on each construct, as negative items tend to correlate even when they are not on the same scale. In other words, a life skills scale with only positive items could look valid and reliable but inflates correlations, while one with both may be more valid and (by standard measures like Cronbach's alpha) less reliable.

Transforming data may not resolve issues

We can see the conundrum clearly in our results. The alpha is lowered by the negative items (see Table 20, row 5), which represents the average of similar numbers of positive and negative items in the Self-Control scale. When we applied our acquiescence bias measure to transform data on the Self-Control scale, for example, we created three variables: the means of: 1) all transformed items (called centered, row 3); 2) selected items transformed (row 4); and 3) items selected after transformation (row 7). The items in the scale hold well together, but less well when we choose the best items based on factor analysis before transforming the data (row 4, alpha=0.57) than when we choose the best items after transforming the data (row 7, alpha=0.69). Overall, transforming the data to remove the acquiescence bias seems to improve the properties of the scales in some dimensions, and not in others.

---

[20] Colosi, Robert, 2005. "Negatively Worded Questions Cause Respondent Confusion." Proceedings of the Survey Research Methods Section, American Statistical Association. 2896 – 2903.

[21] Salazar, Martín Solís, 2015. "The dilemma of combining positive and negative items in scales." Psicothema, Vol. 27, No. 2, 192 – 199. doi: 10.7334/psicothema2014.266

Table 20. Self-Control Analysis (Pilot 6)

Variable	Reliability (alpha)	Skewness	Kurtosis	p-value of Normality Test
Mean of all items	0.67	-0.12	2.71	0.80
Mean of best items	0.71	-0.34	2.29	0.01
Mean of all items centered	0.69	-0.44	3.97	0.00
Mean of best items centered	0.57	-0.61	3.85	0.00
Balanced scale mean of items	0.46	0.16	2.73	0.92
Only positively keyed items	0.74	-0.37	2.31	0.00
Mean of centered best items (after data are transformed, items are selected)	0.69	-0.79	5.48	0.00

Possible solutions[22] for this conundrum are:

- Determine if respondents understand the negative items during pre-test
- Separate positive and negative items on the same topic in the survey
- Remind respondents of the importance of their answers
- If the interviewer thinks the respondent is answering without thinking and may not understand the item, then the interviewer can re-state the item with a moderated voice to bring the respondent's attention back to the survey
- Interviewer should note if a respondent appears to be answering without thinking, as the information can be used to assess answering patterns

We also look at potential correlation between the acquiescence and response extremeness scores and respondent age and average school grades. As seen in the example below from Pilot 6 (Sri Lanka), the response extremeness score is correlated with general grade average.

---

[22] Salazar, Martín Solís. 2015. "The dilemma of combining positive and negative items in scales." *Psicothema*, Vol. 27, No. 2, 192 – 199. doi: 10.7334/psicothema2014.266

Table 21. Correlation: Acquiescence and Response Extremeness Scores
with Respondent Age and Average Grade

	Age	Avg Grade
Acquiescence Score	0.0132	-0.119*
Response Extremeness Score	-0.147*	0.250***
N	292	

* p<0.05   **<0.01   ***<0.001

4.3.2. Anchoring Vignettes

Anchoring vignettes are scenario-type questions in which girls are asked to respond to short stories about the behavior of other hypothetical girls their own age. We included three behaviors, one "good," one "average," and one "bad" in each vignette. Respondents were asked to give us their opinion about the girls described in the vignettes. These questions are meant to adjust for reference bias (the tendency of different respondents to gauge their own responses against different implicit mental benchmarks) so that girls' answers about themselves can be more accurately interpreted.

Table 22. Scores on Re-Scaled Items Based on Vignettes

Position of Girl's Answer Relative to Vignettes	Re-scaled Rating
Lower than the "bad" behavior	1
Equal to the "bad" behavior	2
Between the "bad" behavior and the "average" behavior	3
Equal to the "average" behavior	4
Higher than the average behavior, lower than the "good" one	5
Equal to the "good" behavior	6
Higher than the "good" behavior	7

Using girls' responses to the vignettes, we re-scale the data based on whether the girl rated herself lower than what she rated the "bad behavior," higher than the "good behavior," between the lowest and the average behavior, etc. (see Table 22). Therefore, for the life skills of Empathy, Emotional Awareness, Self-Control, Decision Making and Perseverance, each girl has two scores. One that is based only on her self-rating and the other based on her answers to vignettes.

One hypothesis we made while implementing vignettes was that girls would rank them in order ("bad" behavior getting a lower rating than "good" behavior). When a girl gives the same ranking to two distinct vignettes, we call it a "tie" and when she gives two different but reversed rankings to two vignettes, we call it a "violation." We judge the quality of our vignettes by checking the

percentage of ties and violations. Vignettes that are doing what they are designed to do have fewer ties and violations.

We tested the vignettes in the Bangladesh and Nepal pilots to a limited extent, and more so in baselines in India, Cambodia, Tanzani,a and Bangladesh. Except for the vignettes related to perseverance in India, none of the vignettes have performed adequately in any context. Below are results from the Cambodia baseline (Table 23). The vignettes display a high level of both ties (more than 70% for five dimensions and 60% for the sixth) and violations (25% to 50% for five dimensions and 15% for the sixth).

Table 23. Vignette Analysis Percentage of Ties and Violations

Variable	Sample	N	Mean	Standard Deviation
Percentage of ties "XYZ can manage her emotion"	Total	1199	0.77	0.42
	Project	600	0.77	0.42
	Comparison	599	0.77	0.42
Percentage of ties "XYZ understands others' feeling"	Total	1199	0.76	0.43
	Project	600	0.76	0.43
	Comparison	599	0.76	0.42
Percentage of ties "XYZ is well behaved in class "	Total	1199	0.80	0.40
	Project	600	0.80	0.40
	Comparison	599	0.79	0.41
Percentage of ties "when XYZ has a problem, she tries a lot of solutions to solve it"	Total	1199	0.82	0.38
	Project	600	0.82	0.38
	Comparison	599	0.83	0.38
Percentage of ties "XYZ thinks about the consequences of her actions before making a decision"	Total	1199	0.72	0.45
	Project	600	0.71	0.45
	Comparison	599	0.72	0.45
Percentage of ties "XYZ is committed to her goal"	Total	1199	0.61	0.49
	Project	600	0.62	0.49
	Comparison	599	0.60	0.49
Percentage of violation "XYZ can manage her emotion"	Total	1199	0.32	0.47
	Project	600	0.32	0.46
	Comparison	599	0.32	0.47
Percentage of violation "XYZ understands others' feeling"	Total	1199	0.31	0.46
	Project	600	0.35	0.48
	Comparison	599	0.27	0.44
Percentage of violation" XYZ is well behaved in class"	Total	1199	0.48	0.50
	Project	600	0.50	0.50
	Comparison	599	0.47	0.50

None of our pilot analyses allowed us to conclude that rescaled items based on vignettes were more reliable than non-rescaled items. They were time consuming for the girls and interviewers in a survey that takes on average 40 minutes for a baseline and up to one hour for a pilot. We included the vignettes in the Tanzania baseline, but dropped them from the Sri Lanka pilot in 2018, and have not used them since.

4.3.3. Free List Exercises

Free list exercises assess girls' familiarity and comfort levels within a given topic. Since the first pilot, we have asked girls to think about and name the emotions they experienced in the past seven days. We ask interviewers to assess how easy it was for the girl to do so. We use the results to check for girls' conceptual understanding. We assume that girls who list a larger number of distinct emotions, on average, have a more diverse emotional vocabulary and perhaps also a greater self-awareness regarding distinctions between similar emotional states. In contrast, girls with a limited emotional vocabulary and/or limited emotional self-awareness may list a small number of more general emotions such as "happy" or "sad." We learned, for example, that older girls tend to name more emotions than younger girls and that older girls also name emotions that younger girls do not.

Other free listing exercises we have tried include asking girls to name:

1) Solutions to problems
2) Ways to calm down when angry
3) Qualities she has that make her a good communicator
4) Her good qualities
5) Her qualities that help her to be a leader

We expected GEP girls to mention more qualities than girls in comparison groups. The exercises are time-limited so that respondents have the same amount of time, as some girls lean toward introversion while others lean toward extraversion. Our hypothesis was not borne out by girls interviewed in the Vietnam pilot, although GEP girls more than non-GEP girls named the emotions they experienced during the roleplay exercise.

We expect the number of: 1) emotions a girl says she experienced in the past week; and 2) ways to calm down when angry to correlate with Empathy and Expressing and Managing Emotions. We find weak but statistically significant correlations with some of the measures in Pilots 6 and 7 (see Table 24).

Table 24. Correlations Between Expressing & Managing Emotions and Girls' Free Listing

Likert-based Expressing & Managing Emotions Construct	Pilot 6	Pilot 7	Pilot 6	Pilot 7
	Number of Emotions Named		Number of Solutions Cited	
Mean of all items	0.18**	0.16*	0.15**	0.12
Mean of best items	0.16**	0.13	0.13*	0.17*
Mean of all items centered	0.10	0.21*	0.09	0.12
Mean of best items centered	0.08	-	0.09	-

* p<0.05, ** p<0.01, *** p<0.001

4.3.4. Forced Choice Questions

The forced choice questions consist of six sets of four items. Each of the four items comes from either the Self-Confidence, Decision-Making, Empathy or Perseverance life skills measures. Girls were asked to select the item that is "most like her" and the one that is "least like her." Below is one of the forced choice statements.

I am going to read four statements, please tell me which one is the most like you, and which one is the least like you.

FC1	
A. In general, I am satisfied with myself (happy, pleased). B. I do not visit with friends to play when I need to prepare for studies. C. I can work alone with concentration. D. If I need to buy something, I know how to save my money.	Most like her: |___| Least like her: |___|

We looked at how the results from three previous pilots load on the main factor in confirmatory factor analysis. By doing this, we created new forced choice questions where the first set of items contains the highest loading items for the four dimensions, and the sixth has the lowest loading ones, e.g., the last forced choice question contains the items that loaded lowest among the forced choice options, on average, across the previous three pilots combined. We used the new forced choice questions in Pilot 5 (Nepal).

For each girl, we obtained a set of six "most like me" variables and six "least like me" variables. We created four "forced choice scores" based on these variables where each forced choice score corresponds to one of the life skills of Self-Confidence, Decision Making, Empathy or Perseverance. For each life skill, the score increases by one if the girl selects this dimension as "most like her" and decreases by one if she selects it as "least like her." The results include four scores ranging from -6 to +6 for each respondent. They allow us to compare dimensions and say that a girl has, for example, more perseverance or more self-confidence.

Ordering the dimensions by scores, we generated ranks, ranging from 1 to 4, for each dimension and each girl. If a girl selected mainly items from the Empathy life skill as "most like her" and mainly items from Self-Confidence life skill as "least like her," she will have a high score in Empathy and a low score in Self-Confidence, e.g., Empathy will be ranked 1 and Self-Confidence ranked 4, while the other two dimensions will be ranked 2 and 3.

To check whether these ranks have meaning, we created similar rankings based on the Likert scale scores for all four life skills. If the rankings were random, the distance between the two ranks would

be 1.25. The average distance between our dimensions is smaller (in absolute terms) than 1.25 on three of the four life skills where only the distance on Decision-Making is close to 1.25 (see Table 25).

Table 25. Distance Between Forced Choice and Life Skills Rankings

Life Skill	Average Forced Choice Rank	Average Likert Item Rank	Average Distance (Absolute Value)
Self-Confidence**	2.1	2.2	0.82
Empathy**	2.5	2.5	0.89
Decision-Making	2.5	2.1	1.14
Perseverance	2.8	3.2	0.98

We found positive and significant (at $p<0.05$) correlations between the forced choice and Self-Confidence rankings (0.23) and forced choice and Empathy (0.25). No significant correlations were found for Decision-Making (-0.02) or Perseverance (-0.01). We observed a lack of symmetry in the data for the Perseverance scale, and we know it is a dimension that may be particularly prone to desirability bias. The forced choice responses suggest that the Perseverance data are strongly (positive) biased.

We plan to further explore the use of forced choice questions.

4.4 Addressing the Girl Questionnaire Administration Mode

Prior to Pilot 7 in 2019, interviewers exclusively conducted the questionnaire face-to-face with respondents. In Pilot 7 (Vietnam), we tested a self-administered paper questionnaire with girls to compare with the face-to-face interview. Self-administration included the interviewer introducing the questionnaire to a girl and walking her through the practice questions, then leaving her to complete the questionnaire by herself.

Potential for decreasing bias with self-administration

Our main objective with self-administration was to reduce social desirability bias, which in our case results in overly positive answers and scores skewed to the right. Self-administration reduces costs. We also wanted to see its effect (if any) on non-response. Within each school, girls were randomly sampled for the self-administered version or interviewed.

We expected more unanswered questions from self-administration due to cognitive burden and respondent fatigue.[23] In contrast, an interviewer may help respondents with fatigue and motivate them to answer all questions. An interviewer asks each question systematically from the first to the last, while respondents with a paper questionnaire may skip around and then miss questions. Finally, an interviewer can clarify questions that a respondent does not understand.

Indeed, we did find more missing data from self-administration than interviews (see Table 26). The only variables left empty in the latter group are in the household roster, where age, last grade completed and work for some family members were unknown. All other questions are systematically filled in.

An interviewer conveys many social norms through her appearance and way of talking, without sharing her opinion with respondents. If the interviewer is from a city and girls are from a rural area, she can influence girls' behavior and responses. On the other hand, if the interviewer is from the community, girls may feel they must respect community norms. Researchers also questioned whether girls might be afraid of being excluded from the program or lose financial support if they do not respond well.

Table 26. Percentage of Missing Data by Survey Mode

Number of Empty Cells	Self-Administered N=73	Interviews N=76
0	23%	61%
1 − 2	26%	28%
3 − 4	21%	9%
5 − 8	25%	3%
10+	5%	0%

The number of unanswered questions (item non-response) in the self-administered version are far greater than we expected. We believe respondent comprehension and lack of interviewer support or influence are the reasons. We address the possibility of interviewers pressuring respondents into giving an answer during training. Interviewers can and do help respondents with fatigue through the long interview process by encouraging them to continue.

---

[23] See De Leeuw E.D., van der Zouwen J. (1988). "Data quality in telephone and face-to-face surveys: a comparative meta-analysis," in R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, & J. Waksberg (Eds.), *Telephone survey methodology*, (pp. 283-299). New York: Wiley.

We expect that girls completing the questionnaire by themselves may feel more comfortable with the relative privacy of self-administration and give fewer positive-leaning responses. We therefore expect the self-administered average scores to be lower than those from the interviews. Except for Self-Control, the average score on constructs is higher among the interviewed girls, though differences are small, and most are not statistically significant (Table 27). Additional testing with self-administration is needed, especially with technology such as tablets where we can build in support for girls, e.g., girls can ask for the meaning of a word or for the questions to be read aloud.

Table 27. Differences in Construct Score between Self-Administered and Interviewed Girls (Pilot 7)

Scores: Mean of All Items	Self-administered (Mean)	Interviewed (Mean)	Difference	p-value of Equality Test
Decision-Making	2.78	2.80	0.02	0.76
Perseverance	3.21	3.25	0.03	0.66
Self-Control	2.88	2.83	-0.05	0.41
Expressing & Managing Emotions	2.95	3.10	0.15	0.04**
Self-Confidence	2.87	2.98	0.11	0.09*
Empathy	2.78	2.88	0.10	0.20
Relationship-Building	3.14	3.21	0.07	0.28
Communication	3.10	3.16	0.06	0.40

4.5 Using ALSA in Program Evaluation

Understanding Existing Measures
- Tool review
- Identify & adapt promising approaches

First Pilot
- Draft tool
- Pre-test 1
- Revise
- Pre-test 2
- Revise
- Test

Iteritive Testing and Adapting
- Contextualize
- Pre-test
- Revise
- Test
- Analyze
- Revise

Using ALSA in Program Evaluation
- **Contextualize**
- **Pre-test**
- **Revise**
- **Conduct baseline**
- **Analyze**
- **Revise**
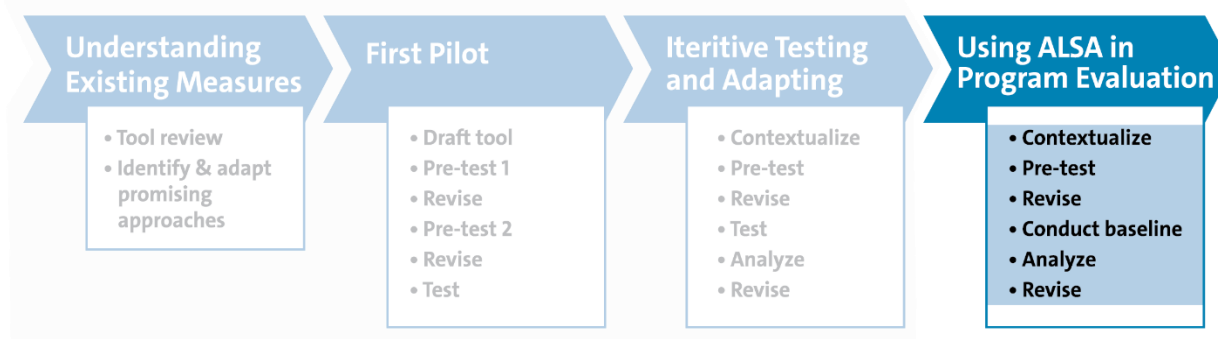
Starting in 2017, we used the tool for program baseline in countries where it was piloted. Respondents have been girls between Grades 6 and 9 and mostly between ages 11 and 17 years. We

use the best fit evaluation method for each context. In countries where we are implementing GEP in a sufficiently large number of schools at a time such that our power calculations suggest that we are likely to identify a measurable difference from comparison schools,[24] we conduct a quasi-experimental evaluation. Where we are not, we use pre-post evaluation. Over time, we have adjusted our cohort-based approach to a repeated cross-section, because differences in pre-post data of a cohort of girls may be a result of natural maturation.

For example, the impact evaluations in India and Cambodia include cohorts interviewed when the girls were in Grades 6 and 7, respectively. They were then interviewed again at the end of lower secondary school (typically grades 8 or 9). For evaluations that started in 2019 or later, we interview early secondary school students using repeated cross-sections in the same grade: first when we introduce GEP (so the 8th/9th graders who we interview are non-participants), and again two to four years later with 8th/9th graders who have participated in the program. The advantages to this approach include being able to conduct a more rigorous pre-post evaluation when we do not have enough schools for an impact evaluation (which is usually the case) and removing concerns about whether it is appropriate and valid to use the same instrument for beginning 6th graders and ending 8th/9th graders. We are currently working on evaluations in six countries: quasi-experiments in India, Cambodia and Nepal, and pre-post evaluations in Tanzania, Bangladesh and Vietnam.

The results enable us to understand the effectiveness of our program and learn how changes in life skills are associated with other outcomes such as school completion, successful transitions from secondary school and agency over life decisions. Collecting comparable data across diverse contexts has enabled us to reflect on our program design and make continuous improvements to the questionnaire and ALSA process.

Despite extensive adaptation of the ALSA to context, we see large variation in the tool's reliability even within a context where we have piloted and conducted a baseline as seen in Table 28. We tested multiple life skills measures and items within measures, some from the sources identified in Table 3 and some of our own design; here you can see where life skills measures were tested and, in some cases, dropped. We reflect on these results in Section 6.1.

COVID 19's Effects on Program Evaluation and Expectations
"We are implementing the baseline and endline in two different situations. We are not using staff in schools or delivering the standard number of sessions a month; we have lowered our expectations. Our concern is that the results may not be what we expected."

Tanzania Team Member

---

[24] Currently defined as launching the program in 20 or more new schools in a country in a single school year, though the assumptions behind this calculation will continue to be refined based on new data.

Table 28. Reliability Results: Cronbach's Alpha

Skill Area	India	Tanzania	Cambodia	Bangladesh	India	Cambodia	Bangladesh	Tanzania	Nepal	Sri Lanka	Vietnam
Self-Confidence (2 measures)	0.39	0.57									
	0.50	0.69	0.72	0.71	0.50	0.66	0.60	0.58	0.59	0.70	0.78
Expressing & Managing Emotions	0.41	0.70	0.72	0.67	0.63	0.62	0.63	0.48	0.69	0.68	0.63
Empathy	0.44	0.65	0.69	0.72	0.60	0.66	0.66	0.45	0.64	0.74	0.73
Self-Control	0.40	0.73	0.69	0.62	0.54	0.72	0.59	0.53	0.68	0.67	0.68
Decision-Making	0.26	0.62	0.57	0.61	0.51	0.65	0.50	0.58	0.45	0.61	0.53
Perseverance	0.51	0.69	0.68	0.63	0.61	0.64	0.61	0.56	0.66	0.64	0.77
Communication	0.49	0.79	0.72	0.73	0.56	0.65	0.64	0.69	0.64	0.74	0.72
Relationship-Building	0.68	0.87	0.78	0.63	0.84	0.85	0.84	0.85	0.71	0.72	0.67

Pre-testing and piloting are critical - reliability scores for subscales increased dramatically after more pre-testing days were added (from ~0.4 to ~0.7).

Attribution

Our commitment to girls leads us to use all the tools at our disposal to understand their needs and how GEP facilitates their growth. We are conducting impact evaluations in India and Cambodia where baselines were conducted in 2017. We collected data from Grade 6 girls in India and Grade 7 girls in Cambodia in 24 GEP and 24 non-GEP schools. We will compare data between the GEP (treatment) and non-GEP (comparison) at endline with the same cohort of girls.

GEP schools are selected based on a set of global criteria[25] that include:

---

[25] Note that schools are not selected with the aim of being representative of any particular geography; rather they are selected based on Room to Read's assessment of the need and the likelihood of program success. As such, we have not attempted to make any comparison between program schools and any regional or national averages; comparisons are instead made only between baseline and endline results and between program and comparison schools.

- The program is only introduced in government schools
- Schools demonstrate gender disparity in enrollments, attendance, completion and/or grades
- The school principal/headmaster, parents and community representatives are supportive of the program and its goals

For the comparison group, we sought schools in the same or neighboring districts that also met the above criteria and matched the project group schools as closely as possible in student-teacher ratios, school type (girls' only or coeducational) and geographic location (urban or rural).

Partial endline data collection was conducted in India during the COVID-19 global pandemic in 2020. The analysis is underway. The Cambodia endline data collection was completed in late 2020.

Solutions Can Present Challenges

Some of the solutions to challenges that we experienced in turn presented new challenges. Here, we present the Self-Control measure as an example of a scale that we modified for Pilot 2 with great success vis-à-vis reliability but with largely skewed data (Table 29). In Pilot 2, we rephrased four of the original items and added items to get at peer pressure (D10 − D12), so that the Self-Control measure consisted of the following:

Table 29. Self-Control Items (Pilot 2)

D1	I refrain from visiting with friends so I can prepare for studies.
D2	I pay attention and resist distractions in class.
D3	I follow instructions given by adults, **even when some friends do not**.
D4	I do my work on time.
D5	I allow others to speak without interruption.
D6	I am polite to adults and classmates, **even when others are not**.
D7	I can control my temper.
D8	I can remain calm even when someone insults or provokes/bullies me.
D9	I am able to stop myself from panicking when I am nervous before exams.
D10	I am able to resist having a boyfriend, even if many of my friends have one
D11	I can have a boyfriend without engaging in sexual activity.
D12	I do not engage in bad behaviors, even if my friends want me to do so.

Table 30. Factor Loading (Pilot 2)

Variable	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Factor 1	0.29	0.73	0.67	0.61	0.54	0.66	0.58	0.61	0.60	0.68	0.68	0.78

The internal consistency for the Self-Control measure improved significantly (Cronbach's alpha = 0.73, up from 0.40) between Pilots 1 and 2, indicating that girls replied in a coherent way to the items and suggesting that comprehension improved. We found, however, a lack of variability in how girls responded to the items (variance = 0.12), suggesting social desirability bias, e.g., they say they are always in control of their temper and never engage in bad behavior or give in to peer pressure (see Figure 9). Lack of discrimination between scale options 3 and 4 may also have contributed to the highly skewed data. In looking at the data distribution, we found them highly skewed (-1.50) and with a heavy right tail (kurtosis=6.32). We needed to improve upon the measure in the next pilot.

In Pilot 3 (Cambodia), we added three negative items to increase response variance, dropped the items measuring peer pressure $(D10 - D12)$ and modified the introduction to remind girls that there are no right or wrong answers and that they should think about their friends and family when responding. Cronbach's alpha results were similar to Tanzania (0.67), but the negative items now loaded negatively. Without them, variance, skewness and kurtosis were within the range of acceptability.

Figure 9. Data Distribution in Pilots 2 and 3
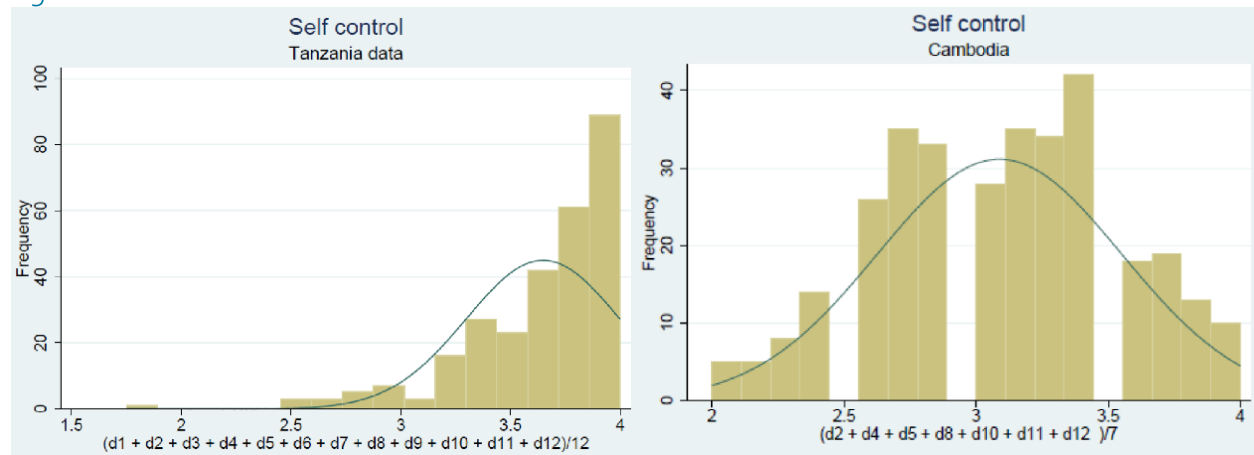
We found that negative items — added in the hope of increasing response variance and normal distribution of the data — loaded negatively and/or were low on a factor. Some negatively worded items, however, are loading high and negatively on the main factor once reverse-coded, such as D3 "I have a hard time finishing what I started" and D6 "It is hard to wait for something I want." While negatively worded items have often been found to load on a separate factor, we were not expecting them to load high and negative. The questions we chose to add to later pre-tests sought to address the following questions: Do respondents develop an "answering routine?" Do they really understand the negatively worded questions? Do interviewers understand negative items?

As reliability and validity are interrelated, some life skills measures, such as Self-Control, were not sufficiently validated during the first three pilots. While continuously exploring validation measures, we made small revisions to contextualize the tool, continuing to face challenges with negative items. In Pilot 4, they loaded negatively again, but low this time, and without them the data were very skewed.

As seen in the examples above, the "big three" concerns about reliability, validity and bias are interrelated. In efforts to improve one, we affect another.

Overall, the girl questionnaire's performance is somewhat stable. Some of the life skills measures perform better than others. We are satisfied with the Relationship-Building measure, and we continue to work on improving all life skills measures. See Figure 10 for the range of Cronbach's alpha on life skills measures through the Vietnam baseline in 2019.

Figure 10. Range of Alpha Scores by Skill Area across All Pilots (excluding India)
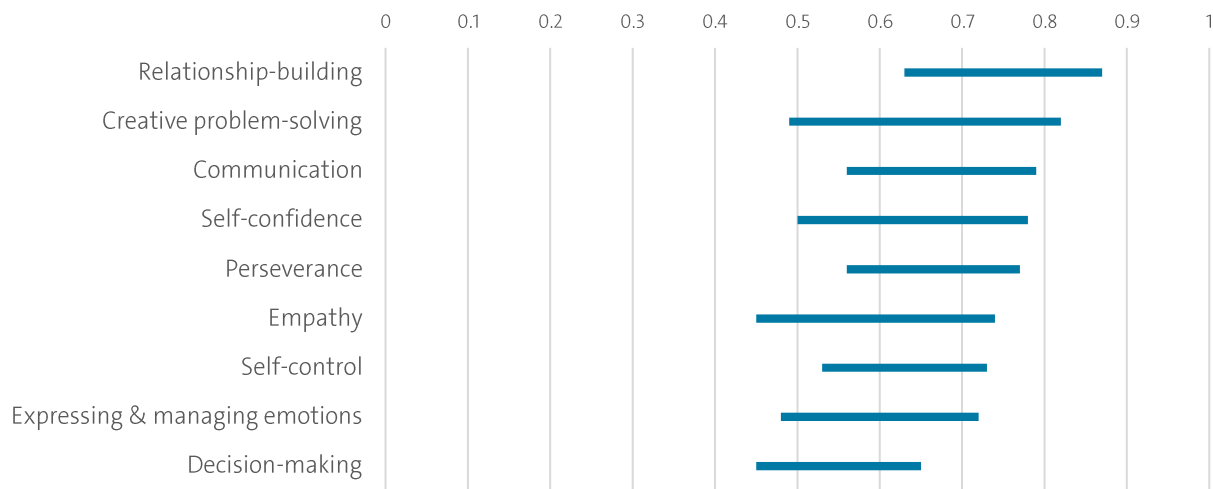
Improving the questionnaire involves a wide range of methods. As noted above, some efforts have proved more useful toward achieving a reliable and valid girl questionnaire, the core of the ALSA. Table 31 provides a snapshot view of our piloting efforts. The blue-shaded rows across the girl questionnaire were piloted as part of our efforts to identify and address bias, while the gray-shaded rows are used to address validity.

Table 31. ALSA Development over Time

	Country	India	Tanzania	Cambodia	Bangladesh	Nepal	Sri Lanka	Vietnam
	Year	2016	2016	2016	2017	2018	2018	2018
	Type	Pilot	Pilot	Pilot	Pilot	Pilot	Pilot	Pilot
Girl Questionnaire	Demographic / Student Profile	●	●	●	●	●	●	●
	Life Skills Modules	●	●	●	●	●	●	●
	Listing*: Emotions	●	●	●	●	●	●	●
	Listing: Peer Nomination	●	●	●	●	●		
	Listing: Solutions to Problems	●	●	●	●	●	●	●
	Listing: Ways to Calm Down						●	●
	Listing: Qualities of a Good Communicator						●	●
	Listing: Respondents Good Qualities						●	●
	Listing: Qualities of a Leader						●	●
	Forced Choice	●		●	●	●		
	Anchoring Vignettes	●	●	●	●	●		
	Map Exercise	●						
	School & Career Aspirations						●	●
	Voice						●	●
	Mirror Drawing	●	●	●	●			
	Strength & Difficulties		●	●	●	●		
	Roleplay Exercise							●
Third Parties	Social Mobilizer Ranking of Girls				●	●	●	●
	Parent Survey	●	●	●	●			
	Teacher Survey				●	●	●	
	Interviewer Questions					●	●	●

*Free listing exercises ask girls to brainstorm by naming, for example, all the emotions they felt in the past week.

5. Room To Read's Pathway Forward

As part of our efforts to use life skills measurement to improve programming for girls, we will continue to improve the ALSA by piloting in new contexts (Laos is next) and using ALSA in GEP evaluations. Alongside this report, we are also releasing a companion how-to guide, which is available at https://www.roomtoread.org/alsa-how-to-guide. We hope the guide will improve the validity of the life skills measures and help standardize the girl questionnaire across contexts. We are also starting to pilot ALSA with boys in some of the countries in which we work as part of engaging boys to further our goal of greater gender equality.

> **Piloting the ALSA with boys**
> In May and June of 2019, Room to Read completed a formative research pilot study in Cambodia to inform our decision-making regarding engaging with boys. This was the first pilot in the process; others will build on it. We seek to answer the following questions:
>
> - Where do we find the biggest gaps in boys' life skills and in their knowledge and attitudes around gender?
> - To what extent are these gaps like those observed in girls in the same context, and where are the key differences that our program should consider?
> - How do boys and others in their communities think about gender roles and challenges faced by boys?
>
> The study included:
>
> 1. A quantitative portion with 90 boys using a version of the ALSA supplemented with questions on boys' gender attitudes, most of which came from widely-used international surveys on boys and masculinity such as the Gender Equitable Men (GEM) scale, the International Men and Gender Equality Survey (IMAGES) and the "Man Box" scale.
> 2. A qualitative portion consisting of focus group discussions with boys, girls, male caregivers (mostly the boys' fathers) and female caregivers (mostly the boys' mothers).
>
> The quantitative survey shows results for life skills, knowledge and attitudes that are broadly similar to Cambodian girls of the same age in a previous assessment with some notable exceptions. In the skills assessment, girls showed higher scores in Empathy while boys scored higher in Creative Problem-Solving. In the area of knowledge and attitudes, boys were more likely to feel that boys' education was more important than girls' education, and to show higher levels of knowledge about early marriage but lower levels of knowledge about menstruation. They were also less likely to agree that one should seek a trusted adult if one is being abused. Boys showed some respect for girls' bodily autonomy with boys agreeing (70%) that a girl has a right not to be touched by boys if she says "no," but high levels of agreement with regard to men's control over household roles and resources around decision-making and caregiving.
>
> In the qualitative assessment, respondents regardless of age and gender tended to express egalitarian attitudes in the abstract, but often more gender-normative views around specific details of men and women's behavior and roles.

6. Conclusion

We continue to refine and test the girl questionnaire in pilots and GEP evaluations. The girl survey became sufficiently stable after the first three pilots to use as a baseline instrument (with modules for knowledge and attitudes, and sometimes voice and aspirations), while continuing to improve it. In this section, we share our reflections on challenges in developing the ALSA, especially the girl questionnaire, and share our successes, issues we continue to grapple with and our next steps.

6.1 Critical Reflection

In the past six years, we learned many lessons about building an ALSA tool. Developing a standardized ALSA for adolescent girls is a great challenge. Engaging in this work has taught everyone who worked/works on it to be creative, open and humble. We have not resolved every challenge, but we have made progress in many areas and are convinced of the value of continued development. We see the benefit of:

- Testing internal consistency using multiple methods
- Testing validity using within-tool measures and exercises, third-party assessment and external data
- Looking for and mitigating bias in the data
- Using an iterative approach to ALSA development

Below we reflect on critical issues that challenge ALSA development and use.

6.1.1   Internal Consistency

Our objective is to look at each construct and how well items hang together to represent a life skill. We know that relying on Cronbach's alpha exclusively will not tell us what we need to know about internal consistency. As noted in *Section 4.1.1: Internal Consistency*, alpha is an often-abused measure and a reductionist approach to looking at internal consistency. As an organization working with psychometricians and our own data to better understand what influences alpha, Room to Read's path to understanding the internal consistency of our constructs has become more complex and challenging. In the process, we learned much about what we do not know, which is essential in assessment development.

We may be seeing a broad range of alphas across contexts and between pilots and baselines because:

- We put considerable resources into contextualizing and pre-testing in the pilots and possibly less than what was needed in the baselines, out of a sense that our instrument was sufficiently tested.
- The baselines being generally larger in scale may also have made quality control more challenging.
- The baselines typically covered more geography than the pilots, and so might have been better adapted for some respondents than others as we saw in Sri Lanka (see *Section 1.4.2: Addressing Validity*). We continue to find contextualization challenging.

Differences in alphas between life skills constructs may result from:

- Some life skills constructs being more culturally and/or linguistically bound and thus contextualized more readily to certain environments. It could also be that we did a better job with the contextualization process in some environments than others due to variations in available resources or in the skill levels of team members.
- Internal consistency variation across life skills constructs, which could partly be related to the number of items within each construct and could suggest that some constructs are more complex than others or less unidimensional. In some cases, we may not have found a valid measure for the outcome of interest, and we need to go back to the literature.

The small-sample test-retest (Section 4.1.2) item-level results suggest that we still have challenges with internal consistency. They also highlight the role that time plays in life skills measurement. Overall, we find ourselves in an uncomfortable but good position in that we know more about the challenges and seek to address big questions such as: What is the threshold for "good enough" measurement? How much precision is functional? How much is needed to make good decisions?

6.1.2   Time Horizon

In developing the ALSA, we have considered the time horizon of measurement, as changes in life skills may take time to manifest. We have made judgments about the reasonable intervention dosage before expecting observable, measurable changes in life skills within the timing of the program. We recognize that a small change in some life skills could lead to big changes in later outcomes, while a big change in some life skills could lead to small changes in outcomes. We want the implementation of the ALSA to inform our understanding of these time horizons so that we can strengthen our program.

Another time horizon of import is how well our life skills measures are capturing a given construct as a *skill* (something malleable to programmatic intervention) versus as a *mood* (changes daily) versus as a *trait* (may be largely immutable) (see 4.1.2). To help us program more effectively, the ALSA should focus on changes in skills over time. Therein lies another time-related challenge, as we know that during adolescence there can be a high degree of volatility and change in these types of skills and attitudes. For example, self-confidence tends to drop in mid-adolescence (around age 13) such that younger adolescents (around age 10) may score higher in self-confidence than 13-year-olds.

We see this in our life skills measures where scores are correlated to such a degree that may indicate that constructs are insufficiently differentiated. In looking at these issues, we find that younger girls have less vocabulary and knowledge compared with older girls. They do not perceive themselves as individuals as much as older girls do. We find that among five life skills, older girls' scores have higher alphas than the scores of younger girls. The life skills constructs with higher alphas are different in the raw data than in the transformed data as seen in Table 32's boldfaced columns.

Table 32. Alpha Results by Respondent Age

	RAW DATA			TRANSFORMED DATA		
	All Girls	Girls Younger than 13	Girls 13 or Older	All Girls	Girls Younger than 13	Girls 13 or Older
Self Confidence	0.70	0.68	**0.75**	0.72	0.71	**0.76**
Expressing & Managing Emotions	0.68	0.67	**0.70**	0.53	0.52	**0.58**
Empathy	0.74	0.74	**0.72**	0.54	0.51	**0.60**
Self-Control	0.67	0.69	**0.71**	0.69	0.69	**0.69**
Decision-Making	0.61	0.58	**0.70**	0.67	0.66	**0.70**
Perseverance	0.64	0.62	**0.68**	0.61	0.61	**0.60**
Communication	0.74	0.75	**0.71**	0.55	0.53	**0.61**
Relationship-Building	0.72	0.72	**0.70**	0.70	0.71	**0.69**

Transforming the data improves the alpha on life skills constructs with negatively worded items more so than on scales with no negative items. Upon addressing acquiescence bias by transforming data, the life skills constructs are better differentiated. The correlation coefficient between life skills decreases with transforming data, meaning that part of the common variance in our life skills constructs was initially due to acquiescence bias.

We may also see desirability bias contributing to high life skills scores in a baseline followed by lower life skills scores at endline after youth have a better understanding of the life skills concepts and more fairly rate themselves. Thus, life skills measures need to consider how respondents' skills evolve over time and with maturation.

Lastly, we hypothesize that in some of our validation techniques, such as vignettes and forced choice, reference points may evolve differently in our program versus control/comparison groups over time. For example, GEP and non-GEP girls in an impact evaluation may rate themselves as being in the middle of their cohort, which would result in a no difference between GEP and non-GEP girls, even if the GEP girls' life skills grow considerably more than do non-GEP girls. Therefore, the "cohort" reference point may not be appropriate. The ALSA tools might better address these concerns.

Teasing apart these issues and addressing them creates enormous challenges that we have not fully resolved. One potential approach is using Item Response Theory (IRT) to address the challenges of measuring life skills over time with a cohort where maturation, and not just the intervention, affect life skills measures. IRT describes the relationship between a respondent's response to an item on the questionnaire and their score on an unobservable (latent) construct, such as a life skill. As we

enter the next phase of development of the ALSA, we will include a detailed review of the potential for IRT to address some of the range of challenges we have encountered to date.

6.1.3   Relying on Self-Report

The field of psychometrics, on which we rely to develop assessment tools like the ALSA, relies greatly on self-report. As researchers, however, we want to triangulate results from multiple sources. We therefore have an overall discomfort relying mostly on self-report, so we need to find more ways to use task-based measures and external data sources.

We are happy with the efforts we made to gather external data; however we did find this extremely time-consuming as the data are not readily available in a useable form. As with our efforts with internal consistency, we now know more about what we do not know and are looking into dissonance between self-report results and external data (e.g., grades) and third-party results.

We are encouraged by our findings to date with respect to third-party assessments, having learned that teachers and social mobilizers are among the most promising as they are readily accessible at schools and may be more reliable third-party assessors. Asking parents to answer questions can be challenging as they may not have opportunities to observe some relevant behaviors in the home environment, and also in some contexts due to literacy levels. Questionnaire implementation with parents is also challenging due to the assessment being conducted at schools, so interviewing parents requires more time in coordinating implementation location and logistics.

We can also balance self-report with direct knowledge and attitude questions in the girl questionnaire. The questions help us understand intervention effects, and a benefit of this approach is that the results can be easily validated.

6.1.4   Contextualizing the Girl Questionnaire

There is a need to adapt the ALSA with great care not only for each context, but also possibly for sub-contexts, as girls from different districts may understand the questions differently. This finding further challenges our goal to have a standardized ALSA tool that requires adapting to:

- How skills manifest may vary across geographies, culture, age groups and between relatively privileged and non-privileged groups within a context, including within countries, across languages and dialects, and gender;
- Structural (political, social and economic) factors that affect how life skills translate into ultimate life outcomes; and
- Local stakeholders' views on which skills are important due to variances across contexts.

Any of the above contextualizing efforts may align better with the local GEP intervention, but negatively affect standardization of the ALSA. We find that we cannot fully standardize the life skills measures for all demographic realities across students and contexts. Despite the challenges, striving for comparability is important for many reasons, including justice and equity concerns.

6.2 Summing Up and Moving Forward

We intend for this synthesis report to be a living document, as the story of ALSA development for GEP is ongoing. Below we summarize what we are happy about as of this writing, what we wish we had done differently and what is next for ALSA development at Room to Read.

6.2.1 What We are Happy About

We appreciate the work done by our collaborative team of experts within and outside of Room to Read who followed good research science principles to understand existing measures, select those that met GEP needs and adapted them. We recognize the ALSA development as an iterative process with trial and error. We are happy about the following:

- Building on existing scales for our life skills measures
- Increasing interviewer training time to four days plus a practice day
- Including cognitive interviews in pre-tests and tool adaptation
- Dedicating resources for adapting the ALSA for each country
- Testing negative/positive item pairs and negative items
- Using a combination of Cronbach's alpha and confirmatory factor analysis as primary analytical tools and exploring others, e.g., Guttman's lambda
- Testing anchoring vignettes, free listing, task-based exercises and roleplay, which add complexity to ALSA administration and/or analysis
- Testing ways to transform data
- Achieving relatively stable life skills constructs

6.2.2 What We Wish We Had Done Differently

We wish we had conducted detailed analysis of each data set immediately after each ALSA administration. By using the results to fuel a team discussion of what worked well, what did not, how to improve life skills scales and internal consistency and how to address bias and validity, we might have identified challenges and solutions more quickly. We also wish we had tested some additional scales for life skills constructs that have not performed as well as hoped.

6.2.3 What We Intend to do Going Forward

Going forward, we are adjusting assessment timelines to allow for more reflection after each administration, testing new scales and testing self-administration with a larger sample. In addition, we are working on better aligning measurement constructs to the life skills definitions using resources like the Harvard Taxonomy Project[26] to explore social and emotional learning. Results may help us with the overlap of life skills found in the domain analysis (see *Annex 5: Domain Analysis*), although we are questioning if our assessment is attempting to disentangle constructs that conceptually cannot be separated. We expect to eventually identify a smaller number of dimensions for each life skill as we validate the measures and find them reliable.

---

[26] See more at https://easel.gse.harvard.edu/taxonomy-project.

To validate the girl questionnaire, we want to test alternative task-based measures and have an explicit plan for testing validations measures.

We are discussing how to address the time horizon issues, since one of the reasons we conduct longitudinal studies is to examine our program's working development hypothesis. Evaluation of girls' life skills is meant to examine the life skills that GEP is teaching, including life skills that are linked with longer-term life outcomes. We want evaluations to demonstrate if we have breakdowns in our chain of hypotheses.

We are considering the unit of measurement. While we have been measuring at the individual level, we are looking at the enabling environment level and the interaction between the two. Assessing the conditions in which young people learn life skills may in some cases be more appropriate, meaningful and actionable than assessing learning outcomes at the level of individual youth.

Lastly, we want to be careful about how we make meaning out of the data, and how we analyze and interpret findings. We do this by either comparing treatment and comparison/control groups or pre-post evaluations. We look at time horizon issues and distributions while being aware of a higher desirability bias at baseline than two to three years later at endline when girls better understand the concepts they are being asked about and their position vis-à-vis those concepts/skills and are more comfortable in sharing their growing self-awareness. We are considering how to interpret findings when we see a decrease or no change in life skills scores, and which domains are most affected by these phenomena.

Annex 1: Sample ALSA Tool

The following is a version of the Life Skills Assessment Pilot Tool used in Vietnam in 2019.

NOTE: Words in **bold** are important for the sentence's meaning. They require specific attention while translating and training interviewers.

General Information

S. No.	Questions	Responses
001	School name	
002	Girl ID	|__|__|__|__|__|
003	Name of the child (copy from the roster)	
004	Date of interview	|__|__| / |__|__| / |__|__|
005	Start time of interview	|__|__| : |__|__|
006	End time of interview	|__|__| : |__|__|
007	Enumerator name	
008	Enumerator ID	
009	District Name	

ENUMERATOR READ: Hello, my name is _____, and I am working with an organization called Room to Read. I am here today to understand a little bit more about you and other girls your age so that Room to Read can design programs that help girls succeed. You heard this before from [name of person who obtain assent], but I'd like to go over what the questionnaire is about.

If you agree to help, I am going to ask you questions about you, and your opinion about some stories. You can always say "no" and choose not to answer a question if you don't want to: just ask me, and I'll move to the next question. But if you answer, please take your time and answer honestly. This is not a test about what you've learned in any lessons, it's only about you, how you think, how you feel, and what happens in your life. There are no right or wrong answers, we just want to understand how girls your age act and think. Your answers will not be shared with anyone and will remain between you and me. Your name will not be linked to your answers after this survey. If you are willing to participate, we will be talking for about an hour.

Do you have any questions? If yes, listen carefully and respond to them. If a question is asked that you cannot answer, follow the instructions from training or in your field manual.

Are you willing to participate?   Circle student's response: Yes   No

Section 1: Demographic Information

Enumerator Instructions: Ask the girl the following background questions

	Questions	Responses	No Response
Z1	How old are you?	|__|__|	
Z2	Have you ever seen your mother or female guardian read? (ENUMERATOR: If girl does not understand, explain: "'Female guardian' means a woman who lives with you and takes care of you, if you do not live with your mother.")	1. Yes 2. No 3. No mother or female guardian	
Z3	Have you ever seen your mother or female guardian write?	1. Yes 2. No 3. No mother or female guardian	
Z4	Have you ever seen your father or male guardian read? (ENUMERATOR: If girl does not understand, explain: "'Male guardian' means a man who lives with you and takes care of you, if you do not live with your father.")	1. Yes 2. No 3. No father or male guardian	
Z5	Have you ever seen your father or male guardian write?	1. Yes 2. No 3. No father or male guardian	
Z6	Are you currently receiving academic tutoring outside of school?	1. Yes 2. No	
Z7	What is your religion?	1. Buddhist 2. Catholic or Christian 3. Other/No religion	

ENUMERATOR Instructions: For each respondent, we want to capture information about all the members of their household, including parents, siblings and extended family. Information should be collected for anyone living in the respondent's household at the time of this interview. It is important to ask specifically about the respondent's household/people living in their house. For "Relationship to Respondent," use the following codes:

1. Father
2. Mother
3. Brother
4. Sister
5. Grandfather
6. Grandmother
7. Uncle
8. Aunt
9. Cousin
10. In-law (all)
11. Domestic servant
12. Other

You should first take all names (fill first column) and then go line by line for details.

ENUMERATOR READ We are now going to discuss about the people you live with. This information will not be used to contact any of your family or household members. This information will only be

used to help us understand more about our program participants and their background. This information will not be shared with the school or other students.

Can you tell me how many people you're living with? The people that share the same food, and sleep under the same roof as you most of the week, can you give me their name?

ENUMERATOR INSTRUCTIONS: Once you have all names, then ask questions to complete the table.

(ex you said "John" was living with you? how is "John" related to you? how old is "John," what is his age? What is the last grade "John" completed? Is "John" married? Does he work for an income?)

Questions for All Household Members				Household members age 4+	Household members age 15+	
No.	Relationship to Respondent	Sex	Age	Last grade completed	Marital Status	Work outside the home
1	(use codes from above)	1. M 2. F	|_|_|	|_|_|	1.Single 2.Married 3.Divorced 4.Widow/er	1. Yes 2. No
2			|_|_|	|_|_|	1.Single 2.Married 3.Divorced 4.Widow/er	1. Yes 2. No
3			|_|_|	|_|_|	1.Single 2.Married 3.Divorced 4.Widow/er	1. Yes 2. No
4			|_|_|	|_|_|	1.Single 2.Married 3.Divorced 4.Widow/er	1. Yes 2. No
5			|_|_|	|_|_|	1.Single 2.Married 3.Divorced 4.Widow/er	1. Yes 2. No
6			|_|_|	|_|_|	1.Single 2.Married 3.Divorced 4.Widow/er	1. Yes 2. No
	Add more rows as needed					

ENUMERATOR READ: Thank you for telling me about the people living in your household. Next, we want to understand a little more about the place where you live. Again, we are only asking this information to help us understand more about our program participants. This information will not be shared with the school or other students. Which of the following is present in your household (owned by anyone living in your house)?

	Item	Response	No Response
1	Radio/radio cassette	1. Yes 2. No	
2	Smart phone	1. Yes 2. No	
3	Television	1. Yes 2. No	
4	Motorcycle/Tuktuk	1. Yes 2. No	
5	Automobile/Car	1. Yes 2. No	
6	Refrigerator	1. Yes 2. No	
7	Computer	1. Yes 2. No	
8	Washing machine?	1. Yes 2. No	

ENUMERATOR READ: We also want to understand more about the place where you live. This information will not be shared with your school or other students.

	Question	Response	No Response
1	What type of roof does your house have?	1. Natural (thatch, palm leaf, sod, etc.) 2. Rudimentary (mat, bamboo, wood planks, etc.) 3. Finished (metal, wood, tile, etc.)	
2	What type of floor does your house have?	1. Natural (earth, sand, etc.) 2. Rudimentary (mat, palm/bamboo, etc.) 3. Finished (metal, wood, tile, cement, shingles, etc.)	

3	Does your house have electricity?	1. Yes 2. No	

**** Introduction to the questions in the tool

Enumerator instructions: The following questions are for the girl to learn how to answer the questions. Please do not rush through these questions; make sure she understands this before proceeding.

ENUMERATOR READ "During the time we are going to be together, I am going to read some statements. Please indicate whether each statement is *not true, slightly true, mostly true* or *completely true* for you. If you don't know, or you don't understand the question, that's okay too. You can just tell me.

Enumerator instructions: Show the printed-out scale to the girl with the gradation in different options.

ENUMERATOR READ: This paper shows the different ways you can answer each question, arranged in a line. After I read each question, you can either say your answer out loud, or you can just point to it on the paper. There are four options:

- **Not True** means that you feel the statement I read is not true at all for you.
- **Slightly True** means that you feel the statement I read might be true sometimes, or a little bit true, but is mostly not true for you.
- **Mostly True** means that the statement I read is true most of the time, or in most situations, but sometimes it might not be true for you.
- **Completely True** means that the statement is totally true for you, in all situations all the time.

Adapt the practice questions to your context.

Before this starts, we will first try some practice questions so you know how to give your answers. Remember, there are no right or wrong answers; we just want to know how you feel and think about some things. Please answer based on your own thoughts and feelings, and remember that this is private, so your answers will stay just between you and me. We want to know how true these statements are *for you*, so you don't have to answer based on what you think is right or wrong. If you don't understand a question you should always tell me, and I can explain more. If you don't want to answer, just tell me and I'll go to the next question.

Practice Questions

		Not True	**Slightly True**	**Mostly True**	**Completely True**	**No Response**
	I like ice cream.	1	2	3	4	
	I like spending time with my friends.	1	2	3	4	
	I know how to speak **Chinese**.	1	2	3	4	
	It is easy for me to wake up before five in the morning.	1	2	3	4	

Choose something difficult for youth

Section 2. Life Skills

2.1 Decision-Making

ENUMERATOR READ: "Thanks a lot. I am going to read some more sentences now. Please tell me for each of these statements how true that statement is for you. Is it *not true* for you, *slightly true* for you, *mostly true* for you or *completely true* for you?"

		Not True	**Slightly True**	**Mostly True**	**Completely True**	**No Response**
E1	When I have problems at school, I am good at **finding ways to solve them**.	1	2	3	4	
E2	If someone I trust asks me to do something I know is wrong, it is easy for me to say "no."	1	2	3	4	
E3	When I make a decision, I think about what might happen as a result.	1	2	3	4	
E4	If I have a **big problem**, I try different solutions to solve it.	1	2	3	4	
E5	I can admit when I have made a mistake.	1	2	3	4	
E6	When I am around my friends, I change the way I act so that I will **fit in**.	1	2	3	4	

		Not True	Slightly True	Mostly True	Completely True	No Response
E7	If someone I trust asks me to do something I know is wrong, it is hard for me to say "no."	1	2	3	4	
E8	If I need to buy something, I know how to **save my money.**	1	2	3	4	
E9	I do not change the way I act around my friends just to **fit in**.	1	2	3	4	
E10	I know what I want to be in the future.	1	2	3	4	
E11	I have a plan to **achieve my goals**	1	2	3	4	

2.2 Perseverance & Self-Control

ENUMERATOR READ: "Now I'm going to read some more statements that can describe how different people think and act. Some of these statements will apply to you a lot and some will not apply to you as much. Please tell me for each of these statements how true that statement is for you. Is it *not true* for you, *slightly true* for you, *mostly true* for you or *completely true* for you?"

		Not True	Slightly True	Mostly True	Completely True	No Response
F1	I make sure that I finish whatever I start to do before starting something new	1	2	3	4	
F2	I **try very hard** even when I don't see the results of my efforts.	1	2	3	4	
F3	Even *if I* **feel like quitting**, I keep working hard	1	2	3	4	
F4	**Big difficulties** do not discourage me. I try again.	1	2	3	4	
F5	When I do something, I **work hard at it**	1	2	3	4	
D1	I try to **focus** on the lesson in class.	1	2	3	4	
D2	My parents tell me I am **thoughtful** when completing a task.	1	2	3	4	

		Not True	Slightly True	Mostly True	Completely True	No Response
D3	It is hard to wait for something I want.	1	2	3	4	
D4	I make time to study.	1	2	3	4	
D4	I am polite to adults and classmates, even when I do not agree with what they are saying.	1	2	3	4	
D5	Many times I **lose focus** before completing a task.	1	2	3	4	
D6	It is not **hard** to wait for something I want.	1	2	3	4	
D7	I can control my anger even if someone is not nice to me.	1	2	3	4	
D9	I can remain calm even when someone hurts my feelings or **teases** me.	1	2	3	4	
D10	I am able to stop myself from panicking when I am **nervous** before exams.	1	2	3	4	
D11	When I feel nervous or **scared**, I know how to calm or relax myself.	1	2	3	4	
D13	I am able to complete tasks without losing **focus**.	1	2	3	4	
D14	My parents **tell me** I do things without thinking.	1	2	3	4	

2.3 Emotions and Empathy

In this section, we are going to talk about feelings. Feelings can also be called emotions.

Can you think about the feelings you had recently? Over the past seven days (one week), you were at school; you were at home; you interacted with people. In all these different situations, your feelings may have changed. Can you remember some of them? Can you name them? (*Probe*: Thanks for your answer, what other feelings can you remember?)

Enumerator: circle the corresponding emotion below. If the girl uses only very general words like "good" or "bad," do not circle anything for these responses.

B1	1 Happiness 2 Sadness 3 Loneliness 4 Anger 5 Jealousy 6 Stress 7 Fear	8 Worry/Anxiety 9 Excitement 10 Pride 11 Empathy/Concern 12 Regret 13 Guilt 14 Disappointment	15 Love 16 Hope 17 Other: 17_1 _____ 17_2 _____ 17_3 _____								
B2	(Enumerator: how many answers did the girl give? Count the number of responses the girl gave in B1 and circle the number)	0	1	2	3	4	5	6	7+
B3	(Enumerator: How easy was it for the girl to answer the question? Circle your response)	1. Very hard		2. Hard		3. Easy		4. Very easy	

ENUMERATOR READ: Thank you. Emotions are what guide your reactions to situations. If something happens that you like, you will feel happy, and you may smile, or laugh. If something happens that you don't like, you may feel angry or sad.

Enumerator instructions: Show the page with the four printed out faces to the girl. You don't need to write anything down for this section.

ENUMERATOR READ: Here are some faces that show different people with a variety of emotions. Can you pick the face that matches each emotion word?

Which of these faces looks **surprised**? (Wait for the girl to indicate one of the faces)

Ok, good. Which one of the faces looks **sad**? (Wait for the girl to indicate one of the faces)

Which one of the faces looks **angry?** (Wait for the girl to indicate one of the faces)

Which one of the faces looks **happy**? (Wait for the girl to indicate one of the faces)
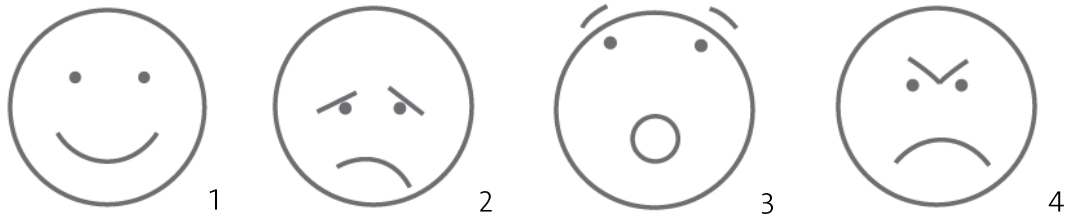
Great. Please keep thinking about you and your emotions over the past few days or week as I am going to read some statements about things that you may feel or do. Remember that this is about you and how you feel, and not about what you have learned in class. There are no right or wrong answers to these questions.

Please tell me for each of these statements how much the following is true for you. Is it *not true* for you, *slightly true* for you, *mostly true* for you or *completely true* for you?"

		Not True	Slightly True	Mostly True	Completely True	No Response
B4	I understand my feelings.	1	2	3	4	
B5	I can understand the **reasons** behind my emotions.	1	2	3	4	
B6	I understand why I feel differently in different situations.	1	2	3	4	
B7	I am aware of how my feelings change my behavior.	1	2	3	4	
B8	I know what makes me happy.	1	2	3	4	
B9	I can describe the best and worst moments of yesterday.	1	2	3	4	
B10	I share my feelings with adults in my family.	1	2	3	4	
B11	When I am **upset**, I can manage my feelings before I take action.	1	2	3	4	
C1	I feel bad when one of my friends is **insulted** or **teased**.	1	2	3	4	
C2	I understand how the people I am close to feel.	1	2	3	4	
C3	I can understand other people's **situations**.	1	2	3	4	
C4	Even if I **feel down**, when I see other people feel excited, I start to feel excited too.	1	2	3	4	
C5	I know how to make other people feel happier.	1	2	3	4	
C6	When I see someone being picked on, I feel sorry for him or her.	1	2	3	4	

		Not True	Slightly True	Mostly True	Completely True	No Response
C7	I get a strong urge to help when I see someone who is upset.	1	2	3	4	
C8	It makes me sad when I see someone who is alone	1	2	3	4	

2.4 Relationship-Building

		Not True	Slightly True	Mostly True	Completely True	No Response
H1	I can choose my friends well.	1	2	3	4	
H2	It is easy for me to always be **kind and caring** with my friends.	1	2	3	4	
H3	It is important to me to help other students in my school.	1	2	3	4	
H4	I **support** my friends when they do the right thing.	1	2	3	4	
H5	I find it easy to make friends.	1	2	3	4	
H6	I **encourage** my friends to be the best they can be.	1	2	3	4	
H7	I help my close friends **feel good about themselves**.	1	2	3	4	
H8	It is **hard** for me to always be kind and caring with my friends.	1	2	3	4	
H9	I am always ready when a close friend needs me.	1	2	3	4	
H10	I would stand up for my friends if another person was causing them trouble.	1	2	3	4	
H11	I find it hard to make friends.	1	2	3	4	
H12	I have **strong women** in my life whom I admire.	1	2	3	4	

2.5 Communication

ENUMERATOR READ: "Thanks a lot; we have made a lot of progress. I am now going to read some sentences, describing some behaviors. Please tell me for each of these statements how true that statement is for you. Is it *not true* for you, *slightly true* for you, *mostly true* for you or *completely true* for you?"

		Not True	Slightly True	Mostly True	Completely True	No Response
G1	I ask my teachers for help when I need it.	1	2	3	4	
G2	If I get angry with a friend, I can talk to them and **make things better.**	1	2	3	4	
G3	I can work well with others.	1	2	3	4	
G5	I get along well with my classmates.	1	2	3	4	
G6	I feel confident enough to talk in front of **a group**.	1	2	3	4	
G7	When a friend is talking to me, I make eye contact (look at his/her eyes).	1	2	3	4	
G8	I ask questions when I do not understand what someone is telling me.	1	2	3	4	
G9	I know how to get support/help if I have a problem.	1	2	3	4	

2.6 Self-Confidence

ENUMERATOR READ: "Great, you are doing well. Now I am going to read some more statements. Think about yourself at school and with your family. Please tell me for each of these statements how true that statement is for you. Is it *not true* for you, *slightly true* for you, *mostly true* for you or *completely true* for you?"

		Not True	Slightly True	Mostly True	Completely True	No Response
A1	In general, I am happy with **the way I am**	1	2	3	4	
A2	I feel that I have a lot of good qualities.	1	2	3	4	

		Not True	Slightly True	Mostly True	Completely True	No Response
A3	I am not afraid that I will **fail**.	1	2	3	4	
A4	I can do things as well as most other girls my age.	1	2	3	4	
A5	I feel that I am just as important as anybody else.	1	2	3	4	
A6	I can make decisions that will help me in my life.	1	2	3	4	
A7	I know what I need to do to be a better person.	1	2	3	4	
A8	I feel **confident** in myself.	1	2	3	4	
A9	I think a lot about **how I want to live my life** in the future.	1	2	3	4	
A10	I feel **positively** about myself.	1	2	3	4	
A11	I am afraid I am going to fail.	1	2	3	4	
A12	I don't think I have many good qualities/skills.	1	2	3	4	
A13	Sometimes, I don't feel confident in myself.	1	2	3	4	

2.7 Creative Problem-Solving

ENUMERATOR INSTRUCTIONS: For the three following questions, you should not give any suggestion to the girls. Read the statement, then let them talk and note their answers down, probe to see whether they can think of anything else.

| I1 | ENUMERATOR READ "Linh and Thu are sisters. They have one mango, and they both want it, so they are starting to fight about the mango. What can the mother of the girls do to help solve this problem? Think about every possible thing she can do. (Probe: good, what else can she do?)" | |

I1_1	Circle all solutions the girl mentions:	
	1. Cut the mango in two parts/tell them to share it.	1
	2. Give them another mango, or another fruit type.	2
	3. Let the girls solve their problem on their own.	3
	4. Go ask the father to solve it.	4
	5. Take the mango away/ Eat the mango herself/give it to someone else.	5
	6. Give the mango to the one that deserves it most.	6
	7. Give them money to go buy another mango.	7
	8. Yell at the girls/ Beat the girls.	8
	9. Reason with the girls, tell them they shouldn't fight.	9
	10. Give the mango to the girls later, after they have stopped fighting.	10
	11. Other 1 2 3 4+	11
	…………………………………	12
		13
	…………………………………	14
		15
	…………………………………	16
		17
I1_2	What would you do if you were the mother? Pick only one answer. Note for enumerator: if the girl comes up with a new item, enumerator should circle this answer in I1_1	|__|
I2	Thao and Lam are talking in class. This prevents Ngoc from listening to the teacher. What can Ngoc do? Think about every possible thing she can do. (Probe: good, what else can she do?)	
I2_1	Circle all solutions the girl mentions:	
	1. Scolding Thao and Lam.	1
	2. Politely ask Thao and Lam to stop talking.	2
	3. Move closer to the teacher/away from Thao and Lam.	3
	4. Start talking with them/join their conversation.	4
	5. Talk with another person in class.	5
	6. Tell the teacher that she cannot hear the lesson because Thao and Lam are talking.	6
	7. Not do anything/just sit there.	7
	8. Get upset/cry.	8
	9. Tell Thao and Lam what she thinks after the class is finished.	9
	10. Other 1 2 3 4+	10
I2_2	What would you do if you were Ngoc? Pick only one answer. Note for enumerator: if the girl comes up with a new item, enumerator should circle this answer in I2_1	|__|
I3	Phuong comes back home after receiving a bad grade/low mark on an exam. She is afraid her father will be angry about it. What can she do? Think about every possible thing she can do. (Probe: good, what else can she do?)	

I3_1	Circle all solutions the girl mentions:	
	1. Talk to her mother/sibling/teacher, ask them to tell the father.	1
	2. Talk to the father that day regardless.	2
	3. Wait to tell the father when he is in a good mood.	3
	4. Do not say anything to the father unless he asks.	4
	5. Make excuses to the father (e.g. the teacher is very tough/the exam was hard).	5
	6. Cry	6
	7. Promise the father she will work hard.	7
	8. Delay going home.	8
	9. Destroy/change/hide the report card.	9
	10. Ask for forgiveness.	10
	11. Other 1 2 3 4+	11
I3_2	What would you do if you were Phuong? Pick only one answer. Note for enumerator: if the girl comes up with a new item, enumerator should circle this answer in I3_1	|__|

Section 3. Knowledge Questions

We're almost done. I just have a few more questions to ask you. These questions will work a little differently. I'd just like you to tell me whether you *agree* or whether you *disagree*. If you don't know, or you don't understand the question, that's okay too. You can just tell me and I'll move to the next question.

		Agree	Disagree	No Response
K1	It is more important for boys to complete 12th grade than girls.	1	2	
K2	Boys are smarter in school than girls.	1	2	
K3	It is not important to me to finish secondary school before getting married.	1	2	
K4	Girls cannot become engineers or scientists because these are and should be only men's jobs.	1	2	
K5	Girls have the right to not **be touched** by a boy even if it hurts the boy's feelings.	1	2	
K6	Men should make all the decisions in the household.	1	2	
K7	If you are experiencing **abuse**, you should ask a trusted adult for help.	1	2	
K8	Boys do not experience **abuse**, only girls do.	1	2	

		Agree	Disagree	No Response
K9	Girls should not express their opinions in front of elders	1	2	
K10	A girl cannot be a leader and a good daughter at the same time	1	2	
K11	Marrying under the age of 18 is **illegal**.	1	2	
K12	It is okay for parents to decide who and when a girl marries, even if she doesn't agree.	1	2	
K13	Early marriage can make it harder to complete secondary school.	1	2	
K14	Girls **have a right** to go to school.	1	2	
K15	Menstruation is a sign that a girl can get pregnant.	1	2	
K16	You should stop washing yourself when you are menstruating.	1	2	
K17	Menstruation is normal and healthy.	1	2	

Ok, we're done! Thank you for talking with me today. I really enjoyed getting the chance to speak with you and learn a little bit about you. I hope it was interesting for you too. Remember that what we talked about is just between you and me. Thanks again, and enjoy the rest of your day!

Section 5: Additional Questions for the Enumerator Only

How true are each of the following statements? Circle your answers.

	Statement	Not True	Slightly True	Mostly True	Completely True
ENUM1	This girl spoke loudly and clearly. It was easy to hear what she said.	1	2	4	5
ENUM2	This girl made eye contact (looked in my eyes) during the interview.	1	2	4	5
ENUM3	This girl asked questions when she did not understand.	1	2	4	5
ENUM4	This girl sat facing me during the interview; she did not turn to the side or away.	1	2	4	5

Annex 2: Pilot 1 India

The following is the Life Skills Assessment Pilot Tool used in India in 2016 following two pre-tests.

Section 0: General Information

S.No	Questions	Responses
001	Associated School (Pre-Sample) Identifier (DISE Code)	|__|__|__|__|__|__|__|__|__|__|__|
002	School name	
003	Tehsil/Block	|__|__|
004	Village name	|_____|
005	Unique Child Identifier	|__|__|__|__|__|__|
006	Name of the Child (Copy from the roster)	
007	Date of First Interview	|__|__|/|__|__|/|__|__|
008	Time of Start of First Interview	|__|__| : |__|__| HH:MM
009	Enumerator ID	

		|___|___|
010	Child's religion	Hindu
		Muslim
011	Caste	

Verbal Informed Consent to be Interviewed   CHILD GEP ID:

Instructions to Interviewer: (1) This form is to be used to obtain assent from all children over age 12 and younger than age 18. (2) Read the following statements to the selected respondent and answer any questions the respondent may have. Do not begin the interview until all questions have been addressed, and the respondent has agreed to participate in the study. Do not interview the respondent if he/she does not give assent.

Hello, my name is _____

I am talking with school age adolescent girls in the area. The information I collect will be used in a study to benefit adolescent girls like you.

I would like to ask you some questions about the work you do, your education and perceptions.

You may choose not to answer any question, if we do so, we will move to the next one, and you may stop the interview at any time if you don't want to answer any more questions.

Your answers to the questions will be kept private and no one else will know what you said.

Your name will not be used in any reports.

It will take about 60 minutes to talk with you today.

Do you have any questions about the study?

May we begin?

Interviewer Certification of Consent:

My signature affirms that I have read the verbal informed consent statement to the child, and I have answered any questions asked about the study. The respondent agreed to be interviewed.

_____Respondent agreed to be interviewed

_____Respondent did not agree to be interviewed→End

Print Interviewer's Name: _____

Interviewer's Signature: _____

Date: _____

Thanks a lot for participating in the survey. All along the questionnaire, we are going to ask you questions about you, your opinion about various stories. You can always refuse to answer if you don't want to, just ask me, and I'll move to the next question. But if you answer, please take your time and answer truly, there are no right or wrong answers, we just want to understand how girls your age act and think. Your answers will remain between you and me, and your name will not be linked to your answers after this survey.

Section 1: Demographic Information

Z Demographic info: Enumerator ask the child:

Z1	How old are you?	|__|__|__|__|__|
Z2	What is your grade?	1. 6th grade 2. 7th grade 3. 8th grade 4. 9th grade 5. 10th grade 6. 11th grade 7. 12th grade
Z3	How many siblings do you have?	
Z4	Have you ever seen your mother read?	1. Yes 0. No

Z5	Have you ever seen your mother write?	1. Yes 0. No
Z6	Have you ever seen your father read?	1. Yes 0. No
Z7	Have you ever seen your father write?	1. Yes 0. No

Section 2: Life skills

Self confidence

READ OUT LOUD: The following pairs of sentences are talking about two kinds of kids. We would like you to decide whether you are more like the first kid or more like the second kid. Then we would like you to decide whether that is only sort of true for you or really true for you.

ENUMERATOR: read the two sentences; ask the girl whether she is more like the first or the second kid. Circle the answers she gives you. Then ask again "you told me you were more like this kid, is that "sort of true for you" or "really true for you"?

A1_1	We'd like you to decide whether you are more like the first child or the second child	
	Some children often get mad at themselves	1
	Other children are pretty pleased with themselves.	2
A1_2	Thanks, now we would like you to decide whether that is only sort of true for you or really true for you	a.sort of true b.very true
A2_1	We'd like you to decide whether you are more like the first child or the second child	
	Some children don't like the way they are leading their life	1
	Other children do like the way they are leading their life.	2
A2_2	Thanks, now we would like you to decide whether that is only sort of true for you or really true for you	a.sort of true b.very true
A3_1	We'd like you to decide whether you are more like the first child or the second child	
	Some children are happy with themselves most of the time.	1
	Other children are often not happy with themselves.	2

A3_2	Thanks, now we would like you to decide whether that is only sort of true for you or really true for you	a.sort of true b.very true
4_1	We'd like you to decide whether you are more like the first child or the second child	
	Some children like the person they are.	1
	Other children often wish they were someone else.	2
A4_2	Thanks, now we would like you to decide whether that is only sort of true for you or really true for you	a.sort of true b.very true
A5_1	We'd like you to decide whether you are more like the first child or the second child	
	Some children are very happy being the way they are.	1
	Other children wish they were different.	2
A5_2	Thanks, now we would like you to decide whether that is only sort of true for you or really true for you	a.sort of true b.very true
A6_1	We'd like you to decide whether you are more like the first child or the second child	
	Some children aren't very happy with the way they do a lot of things.	1
	Other children think the way they do things is fine.	2
A6_2	Thanks, now we would like you to decide whether that is only sort of true for you or really true for you	a.sort of true b.very true

Now, I am going to read some statements about you. Think about yourself in relation to school and family. Please tell me if you really strongly agree, agree, disagree or strongly disagree with the statement		Strongly Agree	Agree	Disagree	Strongly Disagree
A7	In general, I am satisfied with myself (happy, pleased).	1	2	3	4
A8	I feel that I have a number of good qualities (good things about me).	1	2	3	4
A9	I am able to do things as well as most other people (I can do things just as good as other people).	1	2	3	4
A10	I feel that I am just as important as anybody else	1	2	3	4

A11	I wish I could think better of myself.	1	2	3	4
A12	I can make a difference in my own life.	1	2	3	4
A13	I know what I need to do to be a better person.	1	2	3	4
A14	I think that my life has "purpose" or meaning (I think that there is a reason for my life).	1	2	3	4
A15	I think a lot about the future and how I want to live my life.	1	2	3	4
A16	I have a positive attitude about myself	1	2	3	4

Expressing and Managing Emotions

ENUMERATOR READ: Now, I am going to read some statements about things that you may feel or do. Please tell me for each of these statements much the following is true for you is it "completely true for you" "very true for you" "slightly true for you" "not true for you."

		Completely true	Very true	Slightly true	Not true
B1	I understand my moods and feelings	1	2	3	4
B2	I understand why I do the things I do	1	2	3	4
B3	When I'm upset I notice how I am feeling before I take action	1	2	3	4
B4	I am aware of how my moods affect the way I treat other people	1	2	3	4
B5	I can identify and name the emotions I feel	1	2	3	4
B6	I pay attention when I do things	1	2	3	4

ENUMERATOR READ: Now we are going to ask some questions about the last week, i.e. the last 7 days:

B9	Think about a situation in the past 7 days where you felt a strong emotion. I don't need you to tell me about the situation, but can you name this emotion?	
B10	What other emotions did you feel in the last month?	1 Happiness 2 Sadness

	(enumerator: don't read them, circle if emotion mentioned)	3 Loneliness
		4 Anger
		5 Jealousy
		6 Stress
		7 Fear
		8 Other:
		8_1 ....
		8_2 ... .
		8_3 ... .

B11	How many emotions did the girl list?	0 1 2 3 4 5 6 7 or more

B12	Enumerator: don't ask this question: Did that seem difficult for the girl to come up with an answer?	1. Very hard	2	3	4	5. Very easy

Empathy

ENUMERATOR READ: Thanks a lot; you are doing a very good job. We are now going to read some sentences. Please tell me for each of these statements, how much the following is true for you. Is it "completely true for you" "very true for you" "slightly true for you" "not true for you"?

		Completely true	Very true	Slightly true	Not true
C1	I feel bad when someone gets their feelings hurt PROBE: Can you describe how you came up with your answer?	1	2	3	4
C2	I'm happy when others succeed	1	2	3	4
C3	I understand how those close to me feel	1	2	3	4
C4	I can put myself in someone else's shoes (place)	1	2	3	4

C5	I can imagine what other people are feeling	1	2	3	4

Self-Control

Directions: I am going to read a number of different statements that can describe how different people think and act. Please tell me for each of these statements much the following is true for you. Is it "completely true for you" "very true for you" "slightly true for you" "not true for you?"

	Completely true	Very true	Slightly true	Not true
D1. I come to class prepared	1	2	3	4
D2. I pay attention and resist distractions in class.	1	2	3	4
D3. I remember and follow directions.	1	2	3	4
D4. I get to work right away, instead of waiting until the last minute	1	2	3	4
D5. I allow others to speak without interruption	1	2	3	4
D6. I am polite to adults and classmates	1	2	3	4
D7. I can control my temper	1	2	3	4
D8. I can remain calm even when criticized or otherwise provoked	1	2	3	4

Thanks a lot, now I am going to read sentences about girls your age. I would like you to give me your opinion about these girls. Remember, there are no right and wrong answers.

Question	Completely true	Very true	Slightly true	Not true
D9_2 Y is attentive to what the teacher says She do her homework, but she often need to work a lot at the last minute How much do you think the following sentence is true: **"Y is well-behaved in class"**	1	2	3	4
B8_1 X wants to go out to play but her mother would not let her go. She gets angry and breaks her pencil; say she will not eat anything until the mother let her out. How much do you	1	2	3	4

think the following statement is true **"X can manage her emotions?"**				
C6_2 Y likes teasing others, but she stops when she realizes that she is hurting the person. How much do you think the following statement is true **"Y understands others' feelings?"**	1	2	3	4
D9_1 X often has to ask her friend what the teacher said because she thinks about something else. She is often late on her school work. How much do you think the following sentence is true: **"X is well-behaved in class"**	1	2	3	4
B8_2 Y wants to go out and play but her mother would not let her go. She tries to convince her mother, and as it does not work, she go to her room and cries. How much do you think the following statement is true **"Y can manage her emotions?"**	1	2	3	4
C6_1 X is never touched when someone around her is sad. How much do you think the following sentence is true: **"X understands others' feelings?"**	1	2	3	4
D9_3 Z always sits in the first raw so that she avoids being distracted by others, she tries to organize herself to do her homework on a day-to-day basis. How much do you think the following sentence is true: **"Z is well-behaved in class"**	1	2	3	4
B8_3 Z wants to go out and play but the mother would not let her go. She is unhappy about it, she tries to discuss with her mother to understand why, and then stays at home. How much do you think the following statement is true **"Z can manage her emotions?"**	1	2	3	4
C6_3 Z's friends tell her they feel comfortable sharing their feelings with her. She is easily upset when she sees someone crying. How much do you think the following statement is true **"Z understands others' feelings?"**	1	2	3	4

Decision-Making

ENUMERATOR READ: Thanks a lot; we are making good progress in this questionnaire.

I am going to read some sentences now. Please tell me for each of these statements how much it is true for you. Is it "completely true for you" "very true for you" "slightly true for you" "not true for you?"

	Please indicate how much the following is true for you	Completely true	Very true	Slightly true	Not true
E1	When I have problems at school, I am good at finding ways to solve them	1	2	3	4
E2	When I have a problem, I try different solutions until one works.	1	2	3	4
E3	When I make a decision, I think about what might happen afterwards.	1	2	3	4
E4	I take responsibility for my mistakes.	1	2	3	4
E5	I can say "no" when someone wants me to do things that are wrong or dangerous.	1	2	3	4

Perseverance

ENUMERATOR READ: Below a number of different statements that can describe how different people think and act. Some of these statements will apply to you a lot and some will not apply to you as much. Please rate how true each statement is for you. Please tell me for each of these statements how is it true for you. Is it "Completely true for you" "very true for you" "Slightly true for you" "Not true for you"?

	Completely true	Very true	Slightly true	Not true
F1. I finish whatever I begin	1	2	3	4
F2. I work independently with focus	1	2	3	4
F3. I tried very hard even after experiencing failure	1	2	3	4
F4. I stay committed to my goals	1	2	3	4
F5. I keep working hard even when I feel like quitting	1	2	3	4
F6. New ideas and projects sometimes distract me from previous ones	1	2	3	4

F7. Delays and obstacles do not discourage me. I bounce back from disappointments faster than most people do.	1	2	3	4
F8. I have been very focused with a certain idea or project for a short time but later lost interest	1	2	3	4
F9 I am a hard worker	1	2	3	4
F10 I often set a goal but later choose to pursue (follow) a different one	1	2	3	4

Communication

ENUMERATOR READ: thanks a lot; we have done a lot of progress. I am now going to read some sentences, describing some behaviors. Tell me if these sentences apply to you rarely, occasionally, frequently or usually.

	Never/ very rarely	Occasionally	Frequently	Almost always
G1. I know how to disagree without starting a fight or an argument.	1	2	3	4
G2. I ask my teachers for help when I need it.	1	2	3	4
G3. If I get angry with a friend, I can talk about it and make things better.	1	2	3	4
G4. I am able to work well with others.	1	2	3	4
G5. It is easy for me to join a conversation that other students have already started.	1	2	3	4
G6. I enjoy being around people who are different from me.	1	2	3	4
G7 I get along well with other people my age.	1	2	3	4
G8 I am good at making friends.	1	2	3	4
G9 I am good at telling others about my ideas and feelings.	1	2	3	4
G10 I am good at listening to other people.	1	2	3	4

ENUMERATOR READ: I am again going to read some sentences about girls your age, and then I will ask you your opinion about these girls. There is no right or wrong answers; I am just asking you to tell me what you think.

	Completely true	Very true	Slightly true	Not true
D11_1 X often gets into fights and gets aggressive at times. How much do you think the following sentence is true: **"X can control her temper"**	1	2	3	4
E9_2 Y did not work enough and gets a bad grade. She is ashamed, and when she comes back home she tells her parent that the test was very hard, and if she had known she would have work a lot more. How much do you think the following statement is true "**Y takes responsibility for her mistake?"**	1	2	3	4
E7_1 X cannot open the door of her house. She sits down and waits for her parents to come back. How much do you think the following sentence is true **"When X has a problem, she tries different solutions until one works?"**	1	2	3	4
E8_2 Y is on her way to school, and she stops to talk with her auntie. Suddenly she hears the bell and realizes she is getting late, so she runs to school. How much do you think the following sentence is true **"Y is thinking about the consequences of his action before making a decision?"**	1	2	3	4
E9_1 "X didn't work enough, and gets a bad grade. She gets angry at the teacher, and when she comes back to her home, she explains to her parent that the teacher is very tough" How much do you think how much do you think the following sentence is true **"X takes responsibility for her mistake?"**	1	2	3	4
F12_3 Z starts the school year her goal is to succeed She tries to organize her work and leisure time so that she won't be late on her lesson. Sometimes she would like to go play, but she tries to finish her work first. How much do you think the following sentence is true: **"X is committed to her goal?"**	1	2	3	4

I am going to read 4 statements, please tell me which one is the most like you, and which one is the least like you.
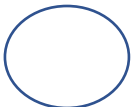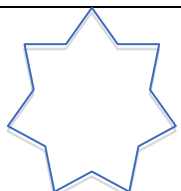
01 A. I am satisfied with myself B. I can control my temper C. I am hard working D. I am good at telling others my ideas and feelings	01_1 most like her: |_____| 01_2 least like her: |_____|
02 A. I have positive feelings about myself B. I am polite to adults and classmates C. I am a hard worker D. I am good at making friends	02_1 most like her: |_____| 02_2 least like her: |_____|
03 A. I feel my life has a purpose (meaning) B. I come to class prepared C. I finish whatever I begin D. I am able to work well with others	03_1 most like her: |_____| 03_2 least like her: |_____|
04 A. I feel that I am just as important as anybody else B. I can remain calm even when insulted or provoked C. Delays or obstacles do not discourage me; I bounce back from disappointments faster than most people D. I know how to disagree without starting a fight or an argument	04_1 most like her: |_____| 04_2 least like her: |_____|
05 A. I think a lot about the future and how I want to live my life B. I allow the other to speak without interruption C. I stay committed to my goals D. It is easy for me to join a conversation that other students have already started	05_1 most like her: |_____| 05_2 least like her: |_____|
06 A. I know what I need to do to be a better person B. I remember and follow direction C. New ideas and project sometimes distract me from previous ones D. I ask my teacher for help if I need it	06_1 most like her: |_____| 06_2 least like her: |_____|
07 A. I wish I could think better of myself B. I sometimes get distracted in class C. I often set goals but later choose to follow a different one D. I don't always get along well with other people my age	07_1 most like her: |_____| 07_2 least like her: |_____|
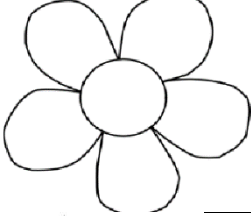
P. Mirror Drawing:

For this part, I'd like you to look at your hand in the mirror. While looking at your hand the whole time, try to trace the same picture than the one you see on the page next to you. Take as much time as you like. (interviewer record time spent on each image, what image was attempted and what image was done successfully; Errors recorded later).

Setup as follows: Image placed so that respondent only sees it in the mirror because the actual image is blocked by a piece of cardboard or something similar. Respondent attempts to trace each image. Start with simplest image, proceeding to one that is more difficult until respondent decides to quit. Record number of images attempted and time spent on each one. (Images below are suggestions)

	Image	Attempted by the girl	Successfully done by The girl	Time taken
P1		1. Yes 0. No	1. Yes 0. No	|__|__| M |__|__| S
P2		1. Yes 0. No	1. Yes 0. No	|__|__| M |__|__| S
P3		1. Yes 0. No	1. Yes 0. No	|__|__| M |__|__| S

| P4 | Flower picture | 1. Yes
0. No | 1. Yes
0. No | \|__\|__\| M \|__\|__\| S |
| P5 | | 1. Yes
0. No | 1. Yes
0. No | \|__\|__\| M \|__\|__\| S |

ENUMERATOR READ: I am again going to read some sentences about girls your age, and then I'll ask you your opinion about these girls. There is no right or wrong answers; I am just asking you to tell me what you think.

	Completely true	Very true	Slightly true	Not true
D11_3 Z can be criticized without getting upset. When she feels she is getting angry, she tries to calm herself down. How much do you think the following sentence is true: **"Z can control her temper?"**	1	2	3	4
F12_1 X goal is to pass to the next level. Starts the school year She does not have good grades. She switches to a new school there also she does not her homework and gets bad grades. How much do you think the following sentence is true **"X is committed to her goal?"**	1	2	3	4
E7_2 Y cannot open the door of her house. She tries to find another key. She finally goes and wait for her parent at the neighbor's house. How much do you think the following sentence is true **"when Y has a problem, she tries different solutions until one works?"**	1	2	3	4

E8_1 X's mother ask her to look after the food, she decides to go out to play, and the food gets burnt.				
How do you think the following statement is true **"X is thinking about the consequences of her action before making a decision?"**	1	2	3	4

Relationship-Building

ENUMERATOR READ I am going to read some statements. Please tell me for each of these statements how much it is true for you. Is it "completely true for you" "very true for you" "Slightly true for you" "not true for you"?

Please think about the current school year when answering these questions

	Completely true	Very true	Slightly true	Not true
I have a friend who...				
H9 Helps me when I am down.	1	2	3	4
H10 Will give me good advice	1	2	3	4
H11 Helps me to do the things I want to do.	1	2	3	4
H12 Makes me feel good about myself	1	2	3	4
H13 Cares about me	1	2	3	4
H14 I can do fun things with	1	2	3	4
H15 I can count on to be there for me if I have a problem or if I am sad	1	2	3	4
H16 I can talk to about a problem at school or at home	1	2	3	4
H17 I support my friends when they do the right thing.	1	2	3	4
H18 I encourage my friends to be the best they can be.	1	2	3	4
H19 I help my close friends feel good about themselves	1	2	3	4
H20 I find it hard to be kind and caring with my friends	1	2	3	4
H21 I am there when a close friend needs me	1	2	3	4

	Completely true	Very true	Slightly true	Not true
H22 I would stand up for my friends if another kid was causing them trouble	1	2	3	4
H23 I find it hard to make friends	1	2	3	4

ENUMERATOR READ: Thanks a lot. Now think about your class, and all of the children that study with you. I am going to ask you some questions about them.

H1 The teacher is asking you to work on a project with a team of students, whom would you pick?	Id1 |___|___|___|___|___|___| Id2 |___|___|___|___|___|___| Id3 |___|___|___|___|___|___| Id4 |___|___|___|___|___|___| Id5 |___|___|___|___|___|___|
H2 The teacher is asking you to work on a project with one student, whom would you pick?	Id |___|___|___|___|___|___|
H3 You are feeling sad, whom would you go to in the class? (cite as many as you want)	Id1 |___|___|___|___|___|___| Id2 |___|___|___|___|___|___| Id3 |___|___|___|___|___|___| Id4 |___|___|___|___|___|___| Id5 |___|___|___|___|___|___|
H4 You are feeling sad, whom would you go to in the class? (cite only one)	Id |___|___|___|___|___|___|
H5 You need help on schoolwork, who would you ask for help in the classroom (cite as many as you want)	Id1 |___|___|___|___|___|___| Id2 |___|___|___|___|___|___| Id3 |___|___|___|___|___|___| Id4 |___|___|___|___|___|___| Id5 |___|___|___|___|___|___|
H6 You need help on schoolwork, who would you ask for help in the classroom (cite only one)	Id |___|___|___|___|___|___|

H7 If one of your classmate needed help on schoolwork, who would you be happy to help. (cite as many as you want)	Id1 |___|___|___|___|___|___| Id2 |___|___|___|___|___|___| Id3 |___|___|___|___|___|___| Id4 |___|___|___|___|___|___| Id5 |___|___|___|___|___|___|
H8 If one of your classmate needed help on schoolwork, who would you be happy to help. (cite only one)	Id |___|___|___|___|___|___|

Creative problem solving

ENUMERATOR INSTRUCTIONS: please show the map to the girl and explain her what the distance between each city are. Then explain her the problem she has to solve.

I0. If you can't proceed with this section, please fill in the reason: 0. Girl cannot read 1. Girl doesn't want to answer 2. Girl can't answer	1 2 3	Enumerator: circle the reason and go to next section

I1:	
Lili wants to go from Nuben to Kado, what is the shortest route she can take?	
- Journey road: answer A B D F G E ….	|__||__||__||__||__||__|
- total number of km on this road (enumerator: ask the girl to compute, report her answer)	|__|__|__|__|__|__|__|

Fig 1 shows up a map of the area and F2 the distances between towns:

Figure 1. Map of roads between towns
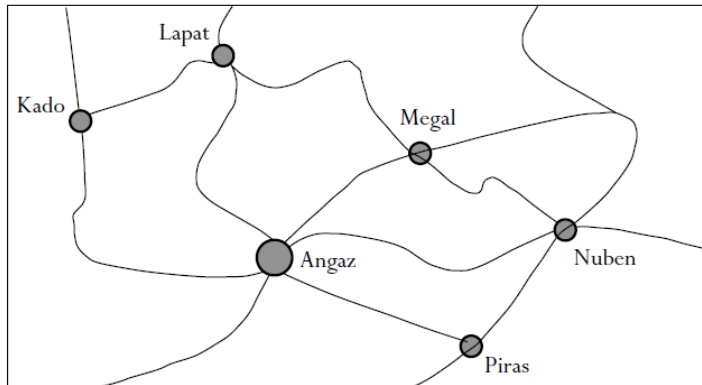
Figure 2. Shortest road distance of towns from each other in kilometres.

	Angaz	Kado	Lapat	Megal	Nuben	Piras
Angaz						
Kado	550					
Lapat	500	300				
Megal	300	850	550			
Nuben	500		1000	450		
Piras	300	850	800	600	250	

ENUMERATOR INSTRUCTIONS: for the three following questions, you should not give any suggestion to the girls, let them talk and note their answers down

I.2 X and Y are sisters, they have one fruit, and they both want it, they are starting to fight. What can the mother of the girls do? Think about every possible thing she can do. (Probe: good, what else can she do?)	
I2_1 circle all solutions the girl mentioned 1 Cut the apple in two parts 2. Go buy another apple 3. Give another fruit type 4. Let the girl solve their problem 5. Go ask the father to solve it 6. Give nothing. 7. Give the apple to the one that deserves it 8. Give the apple to the youngest one 9. Give the apple to one of them; tomorrow she will give another apple to the second. 10 Give them money so that they go buy another apple	1 2 3 4 5 6 7 8 9 10

11 Eat the apple herself 12 Yell at the girls 13 Other	11 12
I2_2 Which of the solutions would you pick if you were the mother.	|__|

I3 Z and F are in the classroom, they are talking, and this prevents Y from listening to the teacher. What can Y do? Think about every possible thing she can do. (Probe: good, what else can she do?)

I3_1 circle all solutions the girl mentioned
 	
      1. Yell at Z and F
      2. Kindly ask Z and F to stop talking
      3. Move closer to the teacher
      4. Start talking with them
      5. Talk with another person
      6. Report to the teacher that they disturb her
      7. Not do anything
      8. Get upset/ cries
      9. Tell the girls what she thinks after the class
      10. Other ....
	1
	2
	3
	4
	5
	6
	7
	8
	9

I3_2 Which of the solutions would you pick if you were this girl.	|__|

I4 Z comes back home after receiving a bad grade. She is afraid her father would be angry about it. What can she do? Think about every possible thing she can do. (Probe: good, what else can she do?)	
I4_1 circle all solutions the girl mentioned 1. Talk to the mother/sibling 2. Talk to the father anyway 3. Wait for the father to be in a better mood 4. Do not say anything to the father 5. Ask the teacher to intervene 6. Tell the father stories (the teacher is very tough /exam was hard) 7. Crying 8. Promise the father she will work hard 9. Other	1 2 3 4 5 6 7 8
I4_2 Which of the solutions would you pick if you were Y.	|__|

ENUMERATOR READ: Thanks a lot, we are almost done, I am again going to read some sentences about girls your age, and then I will ask you your opinion about these girls. There is no right or wrong answers; I am just asking you to tell me what you think.

	Completely true	Very true	Slightly true	Not true
F12_2 Y starts the school year; her goal is to pass to the next level. It is hard for her to study at home; she is very distracted by herself and prefers playing. After some weeks, she asks someone to help her with her homework. How much do you think the following sentence is true **"Y is committed to her goal?"**	1	2	3	4
E7_3 Z cannot open the door of her house. She tries to find another key, but does not find it. She asks the neighbor for a chair and climbs up the chair to enter from the window. From inside she can open the door How much do you think the following sentence is true **"When Z has a problem, she tries different solutions until one works?"**	1	2	3	4
D11_2 Y gets annoyed when others criticize her, and she sometimes criticizes them back, even if, most of the time, it is a calm person. How much do you think the following sentence is true: **"Y can control her temper?"**	1	2	3	4
E9_3 Z did not work enough and gets a bad grade. She is sorry, and when she gets home, she tells her parents she should have worked more. How much you think the following sentence is true **"Z takes responsibility for her mistake?"**	1	2	3	4
E8_3 Z wants to go and play, but she also knows that her mother is going to come back late, so she decides to stay at home to look after the younger brothers and sisters How much do you think the following sentence is true **"Z is thinking about the consequences of his action before making a decision"**	1	2	3	4

ENUMERATOR READ: Thanks a lot, we're done, you did a very good job, thanks for your patience in answering these questions.

Annex 3: Validation Instruments

Social Mobilizer Ranking Tool

The following is an instrument that GEP uses to get data from program staff (social mobilizers/mentors) about the girls they support regarding Empathy, Expressing and Managing Emotions, Communication and Self-Confidence. Asking social mobilizers to rank girls helps us get less biased data, as program staff may give answers that are skewed toward the positive, thinking that will make the program look good.

We are going to ask you to rank girls that you support as a social mobilizer. For each skill, you should think about the girls and how they display this skill. Then you can rank them, starting with the name and GEP code of the girl who is the strongest in this life skill, and order so that the weakest is at the end. If two girls seem similar to you, you should find something that makes a little difference and rank them accordingly.

We are going to use this data anonymously, to understand how the answers girls provide about themselves compares to the way they are seen by others. Remember that we are assessing how well the tool works, and not how well the girls are doing. So please answer truthfully.

Thanks a lot for your participation.

1. Empathy

Empathy is the ability to understand and value another person's feelings. A girl in 9th grade with strong skills in empathy would be likely to use empathetic words toward her peers, friends, teachers and family, and to demonstrate empathetic behavior to others when they are in need of support and care.

Rank	Girl's name	ID number
1st		|__|__|__|__|__|__|
2nd		|__|__|__|__|__|__|
3rd		|__|__|__|__|__|__|
4th		|__|__|__|__|__|__|
5th		

		|_|_|_|_|_|_|_|
6th		|_|_|_|_|_|_|_|
7th		|_|_|_|_|_|_|_|
8th		|_|_|_|_|_|_|_|
9th		|_|_|_|_|_|_|_|
10th		|_|_|_|_|_|_|_|

2. Expressing & Managing Emotions

The skill of expressing and managing emotions includes being aware of one's own emotions and recognizing them as important and valid. It also includes the ability to respond productively to situations where emotions may be strong. A girl in 9th grade with strong skills in Expressing & Managing Emotions would be able to name her feelings and express them assertively to peers, family and teachers without acting out in negative ways.

Rank	Girl's name	ID number
1st		|_|_|_|_|_|_|_|
2nd		|_|_|_|_|_|_|_|
3rd		|_|_|_|_|_|_|_|
4th		|_|_|_|_|_|_|_|

5th		|__|__|__|__|__|__|
6th		|__|__|__|__|__|__|
7th		|__|__|__|__|__|__|
8th		|__|__|__|__|__|__|
9th		|__|__|__|__|__|__|
10th		|__|__|__|__|__|__|

3. Communication

Communication includes self-expression and listening skills. A girl in 9th grade with strong skills in Communication would be able to demonstrate reflective listening, making I-statements, listening with empathy and refusal skills, as well as listening without interrupting her peers, friends, teachers and family.

Rank	Girl's name	ID number
1st		|__|__|__|__|__|__|
2nd		|__|__|__|__|__|__|
3rd		|__|__|__|__|__|__|
4th		|__|__|__|__|__|__|

5th		|__|__|__|__|__|__|
6th		|__|__|__|__|__|__|
7th		|__|__|__|__|__|__|
8th		|__|__|__|__|__|__|
9th		|__|__|__|__|__|__|
10th		|__|__|__|__|__|__|

4. Self-Confidence

Self-Confidence includes a sense of self, positive identity and a sense of competence. A girl in 9th grade with strong Self-Confidence would believe in her own self-worth and that she matters. She would have a strong sense of identity, and be able to describe her own positive attributes, strengths and talents.

Rank	Girl's name	ID number
1st		|__|__|__|__|__|__|
2nd		|__|__|__|__|__|__|
3rd		|__|__|__|__|__|__|
4th		

		|_|_|_|_|_|_|_|
5th		|_|_|_|_|_|_|_|	
6th		|_|_|_|_|_|_|_|	
7th		|_|_|_|_|_|_|_|	
8th		|_|_|_|_|_|_|_|	
9th		|_|_|_|_|_|_|_|	
10th		|_|_|_|_|_|_|_|	

S. No.	Questions	Responses
001	School name	
002	Girl ID	|__|__|__|__|__|__|
003	Name of girl	
004	Grade of girl	
005	Section of girl	
006	Your name	
	Year of teaching	
	Year/ Month teaching the girl	__ mm __ yy

Instructions: Think about everything you know about the girl above. The table below lists a set of behaviors that reflect girls' skills in a variety of different areas. Based on your experience with this girl, how likely do you think she would be to display each behavior in a relevant situation? For each behavior, circle the number corresponding to your answer: *Very Unlikely, Somewhat Unlikely, as likely/unlikely (neutral) Somewhat Likely* or *Very Likely.*

For example, if the behavior listed was "Perform well on a mathematics test," you would need to think about the girl's skill in mathematics. If she were given a mathematics test, how likely would she be to perform well? If your answer is "Likely," you would circle number 3. Note, if you don't know how skilled she is in mathematics, you could tick the box under "Don't Know/Not Applicable"

Always remember that if you don't know the answer you should simply tick the last column and move to next question. Do not select an answer if you don't know, just tick the last column.

How likely do you think the girl is to display the following behavior in a relevant situation?

		Very Unlikely	Somewhat Unlikely	Somewhat Likely	Very Likely	Don't Know
J1	Speak bravely in front of a group of people	1	2	3	4	
J2	Say good things about herself	1	2	3	4	
J3	Prepare herself when there is something she wants to do	1	2	3	4	
J4	Try to do her best	1	2	3	4	

		Very Unlikely	Somewhat Unlikely	Somewhat Likely	Very Likely	Don't Know
J5	Try to learn new things	1	2	3	4	
J6	Act respectfully (show respect) when participating in a game or competition	1	2	3	4	
J7	Respect another person's opinion	1	2	3	4	
J8	Contribute to group activities	1	2	3	4	
J9	Act as a leader to other girls her age	1	2	3	4	
J10	Show creativity in completing a task	1	2	3	4	
J11	Try challenging tasks	1	2	3	4	
J12	Say good things about the future	1	2	3	4	
J13	Cooperate with peers (girls her own age) or siblings	1	2	3	4	
J14	Work hard on school work	1	2	3	4	
J15	Express high expectations for herself	1	2	3	4	
J16	Pay attention when completing a task	1	2	3	4	
J17	Wait for others to finish speaking before she speaks	1	2	3	4	
J18	Show appreciation of or thank others	1	2	3	4	
J19	Focus on a task	1	2	3	4	
J20	Greet a person in a polite or respectful way	1	2	3	4	
J21	Show maturity in her behavior	1	2	3	4	
J22	Accept another choice when her first choice was unavailable	1	2	3	4	
J23	Ask questions to clarify what she did not understand	1	2	3	4	
J24	Shows that she knows what she is good at	1	2	3	4	
J25	Receive praise or admiration from adults	1	2	3	4	
J26	Express her feelings (how she felt)	1	2	3	4	

		Very Unlikely	Somewhat Unlikely	Somewhat Likely	Very Likely	Don't Know
J27	Participate in class	1	2	3	4	
J28	Identify people who can help her to solve a problem if she needs help	1	2	3	4	
J29	Feel moved by another person's feelings	1	2	3	4	
J30	Make a suggestion or request in a polite way	1	2	3	4	
J31	Be kind to younger children	1	2	3	4	
J32	Be restless	1	2	3	4	
J33	Be afraid	1	2	3	4	
J34	Be worried	1	2	3	4	
J35	Be nervous in new situation	1	2	3	4	
J36	Be shy	1	2	3	4	

S. No.	Questions	Responses
001	School name	
002	Girl ID	|__|__|__|__|__|__|
003	Name of girl (copy from the roster)	
004	Date of interview	|__|__|/|__|__|/|__|__|
005	Start time of interview	|__|__| : |__|__| HH:MM
006	End time of interview	|__|__| : |__|__| HH:MM
007	Parent/guardian interviewed	0. Mother 1. Father 2. Female Guardian 3. Male Guardian

I am going to read a series of statements out loud to you. Please think about your daughter as I read them. We are going to use four answers options: Never, sometimes, often and always. (Enumerator show scale)

You will select "never" when the action that we describe never happens, you will select sometimes when the action happens rarely, you will select "often" when the action happens regularly, and you will select always when the action is happening all the time. For example, if I say, "your daughter wakes up after 5:30," you can reply "never" if your daughter never wakes up after 5:30. You will reply "sometimes" if she wakes up after 5:30 in the weekend, but she does not during the week. You will reply "often" if she wakes up after 5:30 more than half of the time (one day out of two). Finally, if she wakes up every day after 5:30 you will reply always. Do you understand?

Practice questions:

		Never	Sometimes	Often	Always	Don't Know/ Reject
J1	My daughter talks about her friends	1	2	3	4	
J2	My daughter goes to the market	1	2	3	4	
J3	My daughter speaks Chinese	1	2	3	4	
J4	My daughter eats rice for lunch	1	2	3	4	

		Never	Some-times	Often	Always	Don't Know/ NA	Not Under-stood
	Over the past 6 months (since the Khmer New Year), how often did she...						

		Never	Some-times	Often	Always	Don't Know/ NA	Not Under-stood
J1	...speak bravely in front of a an elder or a group of people	1	2	3	4		
J2	Keep trying when facing difficulties? *If not understood* Rephrase: "keep trying to complete any task even though she is facing a difficulty" Probe: Can you think about the difficult tasks that you daughter tries to do?	1	2	3	4		
J3	Say good things about herself?	1	2	3	4		
J4	Cope well with insults or mean comments? *If not understood* Probe: Can you think about a moment someone speaks harshly to your daughter?	1	2	3	4		
J5	Prepare herself when there is something she wants to do? *If not understood* Prepare oneself: get ready/organize oneself before doing something	1	2	3	4		
J6	Try to do her best? *If not understood* Rephrase: "How often did you daughter, doing anything, try to do her best?"	1	2	3	4		
J7	Try to learn new things?	1	2	3	4		
J8	Act respectfully (show respect) when she participated in a game or competition? *If not understood* Rephrase: .... when participating in games or competition with friends or relative, acted respectfully towards others	1	2	3	4		
J9	Respect another person's opinion? *If not understood* Rephrase: speak and act showing respect to others' opinions	1	2	3	4		

		Never	Some-times	Often	Always	Don't Know/ NA	Not Under-stood
J10	Contribute to group activities? *If not understood* Probe: Can you think about people at school, at home or in the community, gathering to do any activity?	1	2	3	4		
J11	Act as a leader to other girls her age?	1	2	3	4		
J12	Resolve a disagreement peacefully? *If not understood* Probe: can you think about anytime your daughter had a disagreement with someone?	1	2	3	4		
J13	Show creativity in completing a task? *If not understood* Definition of creativity: different/new ways of doing anything Probe: Can you think about moments where your daughter has something to do? Any task? Can you think about the way she does it?	1	2	3	4		
J14	Share what she has with others?	1	2	3	4		
J15	Try challenging tasks? *If not understood* Probe: Can you think about the last time your daughter tried something hard?	1	2	3	4		
J16	Say good things about the future? *If not understood* Future: next days, months, years, many years	1	2	3	4		
J17	Cooperate with peers (girls her own age) or siblings? *If not understood* Probe: Can you think about the moment where your daughter is with people her age?	1	2	3	4		
J18	Work hard on schoolwork?	1	2	3	4		

		Never	Some-times	Often	Always	Don't Know/ NA	Not Under-stood
J19	Forgive if somebody hurts her feelings or upset her? *If not understood* Probe: Can you think about any moment when someone was not nice to your daughter?	1	2	3	4		
J20	Express high expectations for herself? *If not understood* Probe: Can you think about moments when your daughter is talking about her own future?	1	2	3	4		
J21	Pay attention when doing any task? *If not understood* Rephrase: How often did she do something she had to do with focus?	1	2	3	4		
J22	Wait for others to finish speaking before she speaks? *If not understood* Rephrase: ...does not interrupt when someone talks? ...does she let someone speak before speaking herself?	1	2	3	4		
J23	Show appreciation of or thank others?	1	2	3	4		
J24	Speak about different things in a positive (happy) way?	1	2	3	4		
J25	Focus on a task despite a problem or distraction? *If not understood* Rephrase: focus with a job even though she is sad/disappointed or distracted by someone or something	1	2	3	4		
J26	Greet a person in a polite or respectful way?	1	2	3	4		

		Never	Some-times	Often	Always	Don't Know/ NA	Not Under-stood
J27	Receive praise or admiration from her peers?	1	2	3	4		
J28	Think before she acted?	1	2	3	4		
J29	Express concern for another person?	1	2	3	4		
J30	Accept another choice when her first choice was unavailable *If not understood* Rephrase: Accepts some things or does something instead of the first thing she wanted to get or to do. If she cannot get her the first thing/goal, she accepts the second thing/goal Probe: Can you think about moments when she wanted to do or obtain something?	1	2	3	4		
J31	Ask questions to clarify what she did not understand? *If not understood* Rephrase: if she doesn't understand something she ask to clarify/ask for clarification	1	2	3	4		
J32	Show an awareness of her personal strengths? *If not understood* Rephrase: .... show that she knows what she is good at	1	2	3	4		
J33	Stay calm when faced with a challenge? *If not understood* Rephrase: When having a problem, girl keeps silent or calm Probe: Can you think about the last time she had to face a challenge?	1	2	3	4		
J34	Receive praise or admiration from adults?	1	2	3	4		
J35	Express her feelings how she felt?	1	2	3	4		

		Never	Some-times	Often	Always	Don't Know/ NA	Not Under-stood
	If not understood Express: show/tell Definition of feeling: feeling is the way someone reacts to a situation the person can be sad, happy, angry, surprised						
J36	Give an opinion when asked? *If not understood* Rephrase: when asked something she responds and share her ideas	1	2	3	4		
J37	Decide between right and wrong before doing something *If not understood* Rephrase: Can she differentiate between right and wrong before she does something?	1	2	3	4		
J38	Identify people who can help her to solve a problem if she needs help? *If not understood* Rephrase: if she needs help, she knows who she can ask for help to solve the problem	1	2	3	4		
J39	Offer to help somebody?	1	2	3	4		
J40	Feel touched by another person's feelings? *If not understood* Explain feeling like we did above Use synonyms for "touched" Affected, moved, impressed ...	1	2	3	4		
J41	Make a suggestion or request in a polite way? *If not understood* Rephrase: When she uses comments or request she uses polite words Probe: Think about moments where your daughter wants something	1	2	3	4		

Annex 4: Additional Analyses

Factor Analysis: Example from Bangladesh Baseline

This example presents results from a confirmatory factor analysis across items within each life skills construct. We use results like these to explore which items are working, discuss why we think specific items are not working and make decisions about how to best construct each scale.

Self-Confidence

The low Cronbach's alpha is explained by the last three items of the scale which load null or negatively on the main factor. This is a difficulty identified in previous assessments which we are investigating.

Table 33. Self-Confidence Factor Loadings

	Item	Factor Loading
A1	In general, I am happy with the way I am.	0.36
A2	I feel that I have a lot of good qualities.	0.58
A3	I can do things as well as most other girls my age.	0.42
A4	I feel that I am just as important as anybody else.	0.35
A5	I can make decisions that will help me in my life.	0.52
A6	I know what I need to do to be a better person.	0.39
A7	I think a lot about how I want to live my life in the future.	0.48
A8	I feel positively about myself.	0.73
A9	I am afraid I am going to fail.	0.00
A10	I don't think I have many good qualities/skills.	-0.15
A11	Sometimes, I don't feel confident about myself.	-0.19

Expressing & Managing Emotions

Item B10 loads low on the main factor.

Table 34. Expressing & Managing Emotions Factor Loadings

	Item	Factor Loading
B4	I understand my feelings.	0.47
B5	I can understand the reasons behind my emotions.	0.50
B6	I understand why I feel differently in different situations.	0.49
B7	I am aware of how my feelings change my behavior.	0.41
B8	I know what makes me happy.	0.39
B9	I can describe the best and worst moments of yesterday.	0.56
B10	I share my feelings with adults in my family.	0.21
B11	When I am upset, I can manage my feelings before I take action	0.55

Empathy

Table 35. Empathy Factor Loadings

	Item	Factor Loading
C1	I feel bad when one of my friends is insulted or teased.	0.56
C2	I understand how the people I am close to feel.	0.36
C3	I can understand other people's situations.	0.37
C4	Even if I feel down, when I see other people feel excited, I start to feel excited too.	0.42
C5	I know how to make other people feel happier.	0.53
C6	When I see someone being picked on, I feel sorry for him or her.	0.61
C7	I get a strong urge to help when I see someone who is upset.	0.49
C8	It makes me sad when I see someone who is alone	0.58

Self-Control

All negative items load negatively on the main factor.

Table 36. Self-Control Factor Loadings

	Item	Factor Loading
D1	I try to focus on the lesson in class.	0.38
D2	It is hard to wait for something I want.	-0.19
D3	I make time to study.	0.53
D4	I am polite to adults and classmates, even when I do not agree with what they are saying.	0.44
D5	I have a hard time finishing what I started.	-0.20
D6	I can control my anger even if someone is not nice to me.	0.57
D7	I find it hard to resist when my friends want me do to things that are wrong or dangerous.	-0.21
D8	I can remain calm even when someone hurts my feelings or teases me.	0.47
D9	I am able to stop myself from panicking when I am nervous before exams.	0.43
D10	When I feel nervous or scared, I know how to calm or relax myself.	0.56
D11	My parents tell me I do things without thinking.	-0.20

Decision-Making

All negative items load low on the main factor.

Table 37. Decision-Making Factor Loadings

	Item	Factor Loading
E1	When I have problems at school, I am good at finding ways to solve them.	0.33
E2	When I make a decision, I think about what might happen as a result.	0.47
E3	If I have a big problem, I try different solutions to solve it.	0.45
E4	I can admit when I have made a mistake.	0.26
E5	I hide my true opinions from my friends because I worry they will make fun of me.	-0.33
E6	When I am around my friends, I change the way I act so that I will fit in.	-0.24
E7	If someone I trust asks me to do something I know is wrong, it is hard for me to say "no."	0.42
E8	If I need to buy something, I know how to save my money.	0.31
E9	I know what I want to be in the future.	0.45
E10	I have a plan to achieve my goals.	0.33

Perseverance

Loadings are fine and high enough, low alpha may be due to a limited number of items.

Table 38. Perseverance Factor Loadings

	Item	Factor Loading
F1	I make sure that I finish whatever I start to do before starting something new	0.52
F2	I try very hard even when I don't see the results of my efforts.	0.43
F3	Even if I feel like quitting, I keep working hard	0.58
F4	Big difficulties do not discourage me. I try again.	0.53
F5	When I do something, I work hard at it	0.66

Communication

G4 and G9 are loading very low on the main factor.

Table 39. Communication Factor Loadings

	Item	Factor Loading
G1	I ask my teachers for help when I need it.	0.50
G2	If I get angry with a friend, I can talk to them and make things better.	0.46
G3	I can work well with others.	0.45
G4	It is easy for me to join a conversation that other students have already started.	0.36
G5	I get along well with my classmates.	0.41
G6	I feel confident enough to talk in front of a group.	0.52
G7	When a friend is talking to me, I make eye contact (look at his/her eyes).	0.43
G8	I ask questions when I do not understand what someone is telling me.	0.45
G9	I know how to get support/help if I have a problem.	0.36

Relationship-Building

Loadings are mostly high and this is coherent with the construct's high alpha.

Table 40. Relationship-Building Factor Loadings

	Item	Factor Loading
	I have a friend who...	
H1	Helps me when I am unhappy.	0.75
H2	Will give me good advice.	0.76
H3	Helps me to do the things I want to do.	0.65
H4	Tells me about my good qualities.	0.66
H5	Cares about me.	0.66
H6	I can do fun things with.	0.63
H7	I can count on to be there for me if I have a problem or if I am sad.	0.74
H8	I can talk to about a problem at school or at home.	0.65
	I have an adult in my life who ...	
H9	Helps me when I am unhappy.	0.60
H10	Will give me good advice.	0.76
H11	Helps me to do the things I want to do.	0.72
H12	Makes me feel good about myself.	0.61
H13	Cares about me.	0.63
H14	I can count on to be there for me if I have a problem or if I am sad.	0.60
H15	I can talk to about a problem at school or at home.	0.56
H16	I can go to when I don't feel safe.	0.61
H17	I can talk to about my future career.	0.53
H18	I can talk to about who and when I want to marry.	0.40

Test-Retest Correlation Results in Pilot 5

The following are item-level results of analysis from Pilot 5's small sample test-retest.

Table 41. Test-Retest Correlation Results by Item

Item	Item Content	Coefficient of Correlation	Sign
A1	In general, I am happy with the way I am.	0.54	*
A2	I feel that I have a lot of good qualities.	0.48	*
A3	I can do things as well as most other girls my age.	0.51	*
A4	I feel that I am just as important as anybody else.	0.42	*
A5	I can make decisions that will help me in my life.	0.10	
A6	I know what I need to do to be a better person.	0.35	*
A7	I think a lot about how I want to live my life in the future.	0.10	
A8	I feel positively about myself.	0.62	*
A9	I am afraid I am going to fail.	0.18	
A10	I don't think I have many good qualities/skills.	0.56	*
A11	Sometimes, I don't feel confident about myself.	0.56	*
A12	I am shy	0.09	
B2	number of emotions given per girl	0.24	
B3	how hard was it for the girl	0.05	
B4	I understand my feelings.	0.06	
B5	I can name the feelings I have	0.29	
B6	I can understand the reasons behind my emotions.	0.48	*
B7	I understand why I feel differently in different situations.	0.08	
B8	I am aware of how my feelings change my behavior.	0.54	*
B9	I know what makes me happy.	0.03	
B10	I can describe the best and worst moments of yesterday.	-0.06	
B11	I share my feelings with adults in my family.	0.33	
B12	When I am upset, I can manage my feelings before I take action	0.06	
B13	Knowing when my feelings are making it hard for me to focus.	0.28	

Item	Item Content	Coefficient of Correlation	Sign
B14	Knowing the emotions I feel.	0.02	
B15	Knowing ways to make myself feel better when I'm sad.	0.23	
B16	Noticing what my body does when I am nervous.	0.39	
B17	Knowing when my mood affects how I treat others.	-0.25	
B18	Knowing ways I calm myself down.	0.12	
B19	Getting through something even when I feel frustrated.	0.10	
B20	Being patient even when I am really excited.	0.13	
B21	Staying calm when I feel stressed.	-0.01	
B22	Working on things even when I don't like them.	0.10	
C1	I feel bad when one of my friends is insulted or teased.	0.11	
C2	I understand how the people I am close to feel.	0.30	*
C3	I can understand other people's situations.	0.21	
C4	I can imagine what other people are feeling	0.55	*
C5	Even if I feel down, when I see other people feel excited, I start to feel excited too.	0.46	*
C6	I know how to make other people feel happier.	0.20	
C7	When I see someone being picked on, I feel sorry for him or her.	0.29	
C8	I get a strong urge to help when I see someone who is upset.	0.33	*
C9	It makes me sad when I see someone who is alone	0.37	*
Canch_1	x understands and share others' feeling	0.17	
Canch_2	y understands and share others' feeling	0.12	
Canch_3	z understands and share others' feeling	0.46	*
D1	I pay attention in class and try to focus on the lesson.	0.32	*
D2	I follow instructions given by my teacher and parents, even when I am not sure why they gave me these instructions.	0.28	
D3	I make time to study.	0.66	*
D4	I am polite to others/other people, even when I do not agree with what they are saying.	0.00	

Item	Item Content	Coefficient of Correlation	Sign
D5	I can control my anger even if someone is not nice to me.	0.10	
D6	I can remain calm even when someone hurts my feelings or teases me.	0.39	*
D7	I am able to stop myself from panicking when I am nervous before exams.	0.47	*
D8	When I feel nervous or scared, I know how to calm or relax myself.	0.04	
D9	My parents tell me I do things without thinking.	0.50	*
D10	If my friend wanted me to do something wrong or dangerous, I would find it hard to resist.	0.07	
D11	I have a hard time finishing what I started.	0.06	
D12	It is hard to wait for something I want.	0.35	
Danch_1	x can control herself [anchoring vignette]	-0.11	
Danch_2	y can control herself [anchoring vignette]	0.14	
Danch_3	z controls her behavior [anchoring vignette]	0.22	
E1	When I have problems at school, I am good at finding ways to solve them.	0.33	
E2	When I make a decision, I think about what might happen afterwards.	0.09	
E3	If I have a big problem, I try different solutions to solve it.	0.40	
E4	I can admit when I have made a mistake.	0.29	
E5	I say "no" to my friends when they want me to do things my parents don't want me to do.	0.01	
E6	If I need to buy something, I know how to save my money.	-0.05	
E7	I have a plan to achieve my goals in life	-0.37	
E8	Thinking about what might happen before making a decision.	0.10	
E9	Knowing what is right or wrong.	0.67	*
E10	Thinking of different ways to solve a problem.	-0.02	
E11	Saying "no" to a friend who wants to break the rules.	-0.09	
E12	Helping to make my school a better place.	-0.08	
ENUM1	This girl spoke loudly and clearly. It was easy to hear what she said.	0.29	
ENUM2	This girl made eye contact (looked in my eyes) during the interview.	0.21	
ENUM3	This girl asked questions when she did not understand.	0.04	

Item	Item Content	Coefficient of Correlation	Sign
ENUM4	This girl sat facing me during the interview; she did not turn to the side or away.	0.06	
F1	I make sure that I finish whatever I start to do before starting something new	0.34	*
F2	I try very hard even when I don't see the results of my efforts.	0.26	
F3	I stay committed to my goals	0.32	*
F4	Even if I feel like quitting, I keep working hard	0.61	*
F5	Big difficulties do not discourage me. I try again.	0.32	*
F6	When I do something, I work hard at it	0.38	*
G1	I ask my teachers for help when I need it.	0.34	
G2	If I get angry with a friend, I can talk to them and make things better.	0.52	*
G3	I can work well with others.	0.12	
G4	It is easy for me to join a conversation that other students have already started.	0.38	*
G5	I get along well with my classmates.	-0.07	
G6	I feel confident enough to talk in front of a group.	0.27	
G7	When a friend is talking to me, I make eye contact (look at his/her eyes).	0.45	*
G8	I ask questions when I do not understand what someone is telling me.	0.16	
G9	I know how to get support/help if I have a problem.	0.25	
H1	I have a friend who helps me when I am unhappy.	0.43	*
H2	I have a friend who will give me good advice.	0.30	
H3	I have a friend who helps me to do the things I want to do.	0.66	*
H4	I have a friend who tells me about my good qualities.	0.22	
H5	I have a friend who cares about me.	0.10	
H6	I have a friend who I can do fun things with.	0.38	*
H7	I have a friend who I can count on to be there for me if I have a problem or if I am sad.	0.31	*
H8	I have a friend who I can talk to about a problem at school or at home.	0.53	*
H9	I have an adult in my life who …Helps me when I am unhappy.	0.23	

Item	Item Content	Coefficient of Correlation	Sign
H10	I have an adult in my life who ...Will give me good advice.	0.18	
H11	I have an adult in my life who ...Helps me to do the things I want to do.	0.09	
H12	I have an adult in my life who ...Makes me feel good about myself.	0.10	
H13	I have an adult in my life who ...Cares about me.	0.06	
H14	I have an adult in my life who ...I can count on to be there for me if I have a problem or if I am sad.	0.08	
H15	I have an adult in my life who ...I can talk to about a problem at school or at home.	0.32	*
H16	I have an adult in my life who ...I can go to when I don't feel safe.	0.47	*
H17	I have an adult in my life who ...I can talk to about my future career.	-0.05	
H18	I have an adult in my life who ...I can talk to about who and when I want to marry.	0.45	
H19	I can choose friends who are nice to me.	0.54	
H20	It is important to me to help other students in my school.	0.11	
H21	I support my friends when they do the right thing.	-0.05	
H22	I encourage my friends to be the best they can be.	0.16	
H23	I help my close friends feel good about themselves.	0.37	*
H24	It is hard for me to be kind and caring to my friends.	0.20	*
H25	I am ready when a close friend needs me.	0.35	*
H26	I would stand up for my friends if another person was causing them trouble.	0.08	
H27	I find it hard to make new friends.	0.44	
H28	I have women in my life whom I admire.	0.48	*
H29	There is an adult woman with whom I can talk about everything	0.70	*
I1_2	Number of answers given in I1	0.38	*
I2_2	Number of answers given in I2	0.16	
I3_2	Number of answers given in I3	0.02	
SDQ1	I try to be nice to other people. I care about their feelings.	0.47	*
SDQ2	I am restless, I cannot stay still for long.	0.58	*

Item	Item Content	Coefficient of Correlation	Sign
SDQ3	I usually share with others (food, games, pens etc.).	0.24	
SDQ4	I worry a lot.	0.31	*
SDQ5	I am helpful if someone is hurt, upset or feeling ill.	0.29	
SDQ6	I often feel unhappy, down-hearted or like crying.	0.28	
SDQ7	I am easily distracted; I find it difficult to concentrate.	0.13	
SDQ8	I am nervous in new situations. I easily lose confidence.	0.34	*
SDQ9	I am kind to younger children.	0.37	*
SDQ10	I often volunteer (on my own) to help others (parents, teachers, children, etc.).	0.29	
SDQ11	I think before I do things.	0.18	
SDQ12	A lot of things scare me.	0.15	
SDQ13	I finish the work I am doing, paying close attention.	0.37	*
Z1	age	1	
Z2	female guard read	1	
Z3	female guard write	1	
Z4	male guard read	1	
Z5	male guard write	1	
Z6	academic tutoring	1	
Z7	religion	1	
Z8	cast	1	

Annex 5: Domain Analysis

We define life skills along three domains: 1) Self-efficacy with self-control, perseverance and decision making; 2) Self-awareness with empathy, expressing and managing emotions and self-confidence, and 3) Social awareness with communication, creative problem-solving and relationship building.

We ran a factor analysis on all Likert-scaled variables to try to identify our three domains. We expected to find the highest loadings of all items on the factor corresponding to the domain trait. For example, all items belonging to "empathy" should load high on the "self-awareness" factor.

We allowed factors to be correlated, as it seems likely that personality traits are correlated. Results are detailed below and summarized as follows:

- We found 5 underlying factors in the transformed (centered) data, and 4 in the "raw" data
- Not all items are retained from this analysis (as those that did not load high enough on either of the 4 factors were dropped).
- Not all items load on the expected factors.
- Negative items always load on a separate factor from the rest of the items.
- We identified a "self-efficacy" factor, and (less distinctly) a "social awareness" and "self-awareness" factor in the raw data.
- We identified similar dimensions in the transformed data, but here, social awareness seems to be divided between "conflict aversion" and "conflict resolution."
- Self-Confidence seems to pertain to more than one domain.

In the tables below, we predicted factors and saw how they correlated with various validation tools.

Social Awareness (Factor Scores)	Comm - SM	Comm - Teacher	Relate - Teacher 1	Enumerator	# of qualities cited by girl good communi-cator	# of solutions given in conflict-resolution exercise
Raw Data	-0.10	-0.08	-0.09	0.26***	0.13*	0.09
Conflict Aversion Centered Data	0.03	-0.13	-0.14	0.23***	0.11	-0.01
Conflict Resolution Centered Data	-0.07	0.05	0.04	0.06	0.04	0.13*

t statistics in parentheses   * p<0.05; ** p<0.01; *** p<0.001

Social Awareness (Factor Scores)	Self Confidence SM	General Average	Self-Confidence Teacher	# of qualities cited by girl	# of leader qualities cited by girl	Enumerator
Raw Data	-0.03	0.10	-0.10	0.01	0.11	0.37***
Centered Data	-0.07	0.26***	0.10	0.14*	0.18**	0.30***

t statistics in parentheses   * p<0.05** p<0.01*** p<0.001

Social Awareness (Factor Scores)	General Average	Expressing & Managing Emotions SM	Persever. Teacher	Anxiety Teacher	Enumerator	Voice
Raw Data	0.22***	-0.15*	0.14	-0.09	0.27***	0.05
Centered Data	0.19**	-0.11	0.31**	0.28**	0.05	0.11

t statistics in parentheses   * p<0.05** p<0.01*** p<0.001

Factor and Domain Structure

Negative items

Self-efficacy

Self-awareness?

Social awareness?

Social awareness?

Results from factor analysis: data were "centered"; factors are allowed to be correlated; results were rotated for better fit.

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
CE1			0.3857			0.7714
CE2						0.9691
CE3		0.4699				0.6820
CE4		0.3598				0.7834
CE5						0.8273
CE6						0.8869
CE7	0.5950					0.5815
CE8		0.3949				0.6992
CE9	0.6612					0.4857
CE10		0.5689				0.5406
CE11		0.5815				0.6170
CF1		0.3609				0.7617
CF2		0.5792				0.6449
CF3		0.4265				0.7981
CF4		0.4503				0.6742
CF5		0.3795				0.7682
CD1		0.4048				0.6971
CD2		0.3710				0.8281
CD3	0.7053					0.5074
CD4						0.7789
CD5						0.7819
CD6	0.7537					0.4249
CD7						0.9150
CD8				0.4526		0.7927
CD9				0.3946		0.8056
CD10						0.8258
CD11						0.7644
CD12						0.9018
CD13	0.6866					0.4910
CB4			0.4346			0.6778
CB5				0.3764		0.7147
CB6						0.8516
CB7			0.5216			0.7258
CB8						0.8236
CB9						0.7885
CB10		0.3564				0.8156
CB11						0.7566
CC1						0.7964
CC2			0.5125			0.7257
CC3			0.4714			0.7267
CC4						0.8808
CC5			0.5006			0.7495
CC6					0.4054	0.7451
CC7					0.4404	0.7328
CC8						0.8570
CH1			0.4895			0.6681
CH2						0.7236
CH3				0.4791		0.7570
CH4					0.5158	0.6749
CG8					0.5932	0.5915
CA1						0.7652
CA2						0.8713
CA3						0.8964
CA4						0.8533
CA5						0.7980
CA6						0.8038
CA7						0.7596
CA8						0.7394
CA9						0.8862
CA10				0.3899		0.7975
CA11	0.7696					0.4177
CA12	0.8120					0.3164
CA13	0.7395					0.4414

Rotated factor loadings (pattern matr[...] [varian]ces

Self-efficacy

Social awareness

Self-awareness

Negative items

Results from factor analysis: data were not centered; factors are allowed to be correlated; results were rotated for better fit.

| Variable	Factor1	Factor2	Factor3	Factor4	Uniqueness
E1		0.3805			0.7522
E2					0.9332
E3		0.5729			0.6950
E4					0.7733
E5					0.8222
E6		0.4629			0.7037
E7					0.9243
E8		0.3906			0.8277
E9				0.3664	0.8471
E10		0.5455			0.7148
E11		0.4813			0.6346
F1		0.4453			0.6987
F2		0.6443			0.6664
F3		0.3965			0.7510
F4		0.4495			0.6794
F5	0.3797	0.4089			0.6331
D1		0.5390			0.6598
D2		0.5393			0.6763
D3					0.8894
D4		0.4274			0.6424
D5					0.6838
D6					0.8129
D7					0.8001
D8					0.7248
D9					0.8184
D10					0.8554
D11					0.7060
D12					0.7696
D13				0.3990	0.8296
B4		0.4489			0.6199
B5			0.3564		0.6503
B6			0.3740		0.8171
B7			0.4341		0.6989
B8	0.3845				0.7962
B9					0.7343
B10		0.4035			0.7911
B11	0.3625				0.6934
C1			0.4894		0.6898
C2			0.4282		0.6378
C3			0.5547		0.6285
C4			0.4486		0.7317
C5					0.6414
C6	0.4619				0.7040
C7	0.6312				0.5987
C8	0.3795				0.7060
H1	0.4006				0.6083
H2			0.3513		0.6768
H3					0.6234
H4	0.6331				0.5892
H5	0.4495				0.6531
G6			0.3584		0.7357
G7			0.6376		0.5999
G8	0.6326				0.5832
A1	0.4914				0.7734
A2					0.7256
A3			0.3714		0.6876
A4			0.5210		0.7429
A5	0.3728				0.6303
A6			0.3791		0.6626
A7	0.4004				0.7301
A8	0.5743				0.6319
A9					0.8223
A10			0.3611		0.6815
A11				0.6094	0.6170
A12				0.6513	0.5714
A13				0.5330	0.7044

(blanks represent abs(loading)< .35)

Self-Confidence

- A1   In general, I am happy with the way I am.
- A2   I feel that I have a lot of good qualities.
- A3   I am not afraid that I will fail.
- A4   I can do things as well as most other girls my age.
- A5   I feel that I am just as important as anybody else.
- A6   I can make decisions that will help me in my life.
- A7   I know what I need to do to be a better person.
- A8   I feel confident about myself.
- A9   I think a lot about how I want to live my life in the future.
- A10   I feel positively about myself.
- A11   I am afraid I am going to fail.
- A12   I don't think I have many good qualities.
- A13   I don't feel confident about myself.

Emotional Awareness

- B4   I know my feelings.
- B5   I can understand the reasons behind my emotions.
- B6   I understand why I feel differently in different situations.
- B7   I am aware of how my feelings change my behavior.
- B8   I know what makes me happy.
- B9   I can tell the best and worst moments of yesterday.
- B10   I share my feelings with adults in my family.
- B11   When I am upset, I can manage my feelings before I take action

Empathy

- C1   I feel bad when one of my friends is insulted or teased.
- C2   I understand how the people I am close to feel.
- C3   I can understand other people's situations.
- C4   When I see someone depressed, I can show my understanding and support to him or her.
- C5   I know how to make other people feel happier.
- C6   When I see someone being picked on, I feel sorry for him or her.
- C7   I get a strong urge to help when I see someone who is upset.
- C8   It makes me sad when I see someone who is lonely

Self-Control

- D1   I try to focus on the lesson in class.
- D2   My parents tell me I am thoughtful when completing a task.
- D3   It is hard to wait for something I want.
- D4   I make time to study.
- D5   I am polite to adults and classmates, even when I do not agree with what they are saying.
- D6   Many times I lose focus before completing a task.
- D7   It is not hard to wait for something I want.
- D8   I can control my anger even if someone is not nice to me.
- D9   I can remain calm even when someone hurts my feelings or teases me.
- D10 I am able to calm myself from panicking when I am nervous before exams.
- D11 When I feel nervous or scared, I know how to calm or relax myself.
- D12 I am able to complete tasks without losing focus.
- D13 My parents tell me I do things without thinking.

Decision-Making

- E1   When I have problems at school, I am good at finding ways to solve them.
- E2   If someone I trust asks me to do something I know is wrong, it is easy for me to say "no."
- E3   When I make a decision, I think about what might happen as a result.
- E4   If I have a big problem, I consider different solutions and then choose the best one.
- E5   I can admit when I have made a mistake.
- E6   When I am around my friends, I change the way I act so that I will fit in.
- E7   If someone I trust asks me to do something I know is wrong, it is hard for me to say "no."
- E8   If I need to buy something, I know how to save my money.
- E9   I do not change the way I act around my friends just to fit in.
- E10 I know what I want for my career in the future.
- E11 I have a plan to achieve my goals.

Perseverance

- F1   I finish whatever I begin.
- F2   I try very hard even when I don't see the results of my efforts.
- F3   Even if I feel like quitting, I keep working hard.
- F4   Difficulties do not discourage me. I try again.
- F5   When I do something, I work hard at it.
- F6   I enjoy working on challenging tasks.
- F7   I study even if nobody makes me do it.
- F8   If I fail I try to find out why.

Communication

- G1   I ask my teachers for help when I need it.
- G2   If I get angry with a friend, I can talk to them and make things better.
- G3   I can collaborate well with others.
- G4   I get along well with my classmates.
- G5   I feel confident enough to talk in front of a group.
- G6   When a friend is talking to me, I make eye contact (look at his/her face).
- G7   I ask questions when I do not understand what someone is telling me.
- G8   I know how to get support/help if I have a problem.

Relationship-Building

- H1   I can choose my friends well.
- H2   It is easy for me to always be kind and caring with my friends.
- H3   It is important to me to help other students in my school.
- H4   I support my friends when they do the right thing.
- H5   I find it easy to make friends.
- H6   I encourage my friends to be the best they can be.
- H7   I help my close friends feel good about themselves.
- H8   It is hard for me to always be kind and caring with my friends.
- H9   I am always ready when a close friend needs me.
- H10   I would stand up for my friends if another person was causing them trouble.
- H11   I find it hard to make friends.
- H12   There are female role models in my life whom I admire.

Variable Name	Item	Factor identified	
		Raw Data	Transformed Data
E1	When I have problems at school, I am good at finding ways to solve them.	Self-Efficacy	
E2	If someone I trust asks me to do something I know is wrong, it is easy for me to say "no."	Self-Efficacy	
E3	When I make a decision, I think about what might happen as a result.		Self-Efficacy
E4	If I have a big problem, I consider different solutions and then choose the best one.		Self-Efficacy
E6	When I am around my friends, I change the way I act so that I will fit in.	Self-Efficacy (note this item was not understood)	
E7	If someone I trust asks me to do something I know is wrong, it is hard for me to say "no."		Negative Item

Variable Name	Item	Factor identified	
		Raw Data	Transformed Data
E8	If I need to buy something, I know how to save my money.	Self-Efficacy	Self-Efficacy
E9	I do not change the way I act around my friends just to fit in.	Negative Item (note this item was not understood)	
E10	I know what I want for my career in the future.	Self-Efficacy	Self-Efficacy
E11	I have a plan to achieve my goals.	Self-Efficacy	Self-Efficacy
F1	I finish whatever I begin.	Self-Efficacy	Self-Efficacy
F2	I try very hard even when I don't see the results of my efforts.	Self-Efficacy	Self-Efficacy
F3	Even if I feel like quitting, I keep working hard.	Self-Efficacy	Self-Efficacy
F4	Difficulties do not discourage me. I try again.	Self-Efficacy	Self-Efficacy
F5	When I do something, I work hard at it.	Self-Efficacy	Self-Efficacy
D1	I try to focus on the lesson in class.	Self-Efficacy	Self-Efficacy
D2	My parents tell me I am thoughtful when completing a task.	Self-Efficacy	Self-Efficacy
D3	It is hard to wait for something I want.		Negative Item
D4	I make time to study.	Self-Efficacy	
D6	Many times I lose focus before completing a task.		Negative Item
D8	I can control my anger even if someone is not nice to me.		Social Awareness (Conflict Aversion)
D9	I can remain calm even when someone hurts my feelings or teases me.		Social Awareness (Conflict Aversion)
D13	My parents tell me I do things without thinking.	Negative Item	Negative Item
B4	I know my feelings.	Self-Efficacy	Self-Awareness
B5	I can understand the reasons behind my emotions.	Self-Awareness	Social Awareness (Conflict Aversion)

Variable Name	Item	Factor identified	
		Raw Data	Transformed Data
B6	I understand why I feel differently in different situations.	Self-Awareness	
B7	I am aware of how my feelings change my behavior.	Self-Awareness	Self-Awareness
B8	I know what makes me happy.	Social Awareness / Self-Confidence	
B10	I share my feelings with adults in my family.	Self-Efficacy	Self-Efficacy
B11	When I am upset, I can manage my feelings before I take action.	Social Awareness / Self-Confidence	
C1	I feel bad when one of my friends is insulted or teased.	Self-Awareness	
C2	I understand how the people I am close to feel.	Self-Awareness	Self-Awareness
C3	I can understand other people's situations.	Self-Awareness	Self-Awareness
C4	When I see someone depressed, I can show my understanding and support to him or her.	Self-Awareness	
C5	I know how to make other people feel happier.		Self-Awareness
C6	When I see someone being picked on, I feel sorry for him or her.	Social Awareness / Self-Confidence	Social Awareness 2 (Conflict Resolution)
C7	I get a strong urge to help when I see someone who is upset.	Social Awareness / Self-Confidence	Social Awareness 2 (Conflict Resolution)
C8	It makes me sad when I see someone who is lonely.	Social Awareness / Self-Confidence	
H1	I can choose my friends well.	Social Awareness / Self-Confidence	Self-Awareness
H2	It is easy for me to always be kind and caring with my friends.	Self-Awareness	
H3	It is important to me to help other students in my school.		Social Awareness (Conflict Aversion)

Variable Name	Item	Factor identified	
		Raw Data	Transformed Data
H4	I support my friends when they do the right thing.	Social Awareness / Self-Confidence	Social Awareness 2 (Conflict Resolution)
H5	I find it easy to make friends.	Social Awareness / Self-Confidence	Self-Awareness
H6	I encourage my friends to be the best they can be.	Social Awareness / Self-Confidence	Social Awareness 2 (Conflict Resolution)
H7	I help my close friends feel good about themselves.		Social Awareness (Conflict Aversion)
H8	It is hard for me to always be kind and caring with my friends.	Negative Item	Negative Item
H9	I am always ready when a close friend needs me.	Social Awareness / Self-Confidence	
H11	I find it hard to make friends.	Negative Item	Negative Item
G2	If I get angry with a friend, I can talk to them and make things better.	Social Awareness / Self-Confidence	
G3	I can collaborate well with others.		Self-Awareness
G4	I get along well with my classmates.	Social Awareness / Self-Confidence	
G5	I feel confident enough to talk in front of a group.	Social Awareness / Self-Confidence	
G6	When a friend is talking to me, I make eye contact (look at his/her face).	Self-Awareness	
G7	I ask questions when I do not understand what someone is telling me.	Self-Awareness	
G8	I know how to get support/help if I have a problem.	Social Awareness / Self-Confidence	Social Awareness 2 (Conflict Resolution)
A1	In general, I am happy with the way I am.	Social Awareness / Self-Confidence	

Variable Name	Item	Factor identified	
		Raw Data	Transformed Data
A3	I am not afraid that I will fail.	Self-Awareness	
A4	I can do things as well as most other girls my age.	Self-Awareness	
A5	I feel that I am just as important as anybody else.	Social Awareness / Self-Confidence	
A6	I can make decisions that will help me in my life.	Self-Awareness	
A7	I know what I need to do to be a better person.	Social Awareness / Self-Confidence	
A8	I feel confident about myself.	Social Awareness / Self-Confidence	
A10	I feel positively about myself.	Self-Awareness	Social Awareness (Conflict Aversion)
A11	I am afraid I am going to fail.	Negative Item	Negative Item
A12	I don't think I have many good qualities.	Negative Item	Negative Item
A13	I don't feel confident about myself.	Negative Item	Negative Item