

# Psychosocial Scales in the Young Lives Round 4 Survey

Selection, Adaptation and Validation

Louise Yorke and María José Ogando Portela



# Psychosocial Scales in the Young Lives Round 4 Survey: Selection, Adaptation and Validation

Louise Yorke and María José Ogando Portela

First published by Young Lives in May 2018

© Young Lives 2018

Printed on FSC-certified paper from traceable and sustainable sources.

## **About Young Lives**

Young Lives is an international study of childhood poverty, following the lives of 12,000 children in four countries (Ethiopia, India, Peru and Vietnam) over 15 years. [www.younglives.org.uk](http://www.younglives.org.uk)

Young Lives is core-funded by UK aid from the Department for International Development (DFID).

The views expressed are those of the author(s). They are not necessarily those of, or endorsed by, Young Lives, the University of Oxford, DFID or other funders.

Core-funded by



**Young Lives**, Oxford Department of International Development (ODID), University of Oxford,  
Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751 • Email: [younglives@younglives.org.uk](mailto:younglives@younglives.org.uk)

# Contents

<b>The authors</b>	<b>4</b>
<b>Acknowledgements</b>	<b>4</b>
<b>1. Introduction</b>	<b>5</b>
<b>2. Psychosocial skills</b>	<b>5</b>
2.1. What are psychosocial skills?	5
2.2. Why are they important?	5
2.3. Psychosocial skills administered in Young Lives	6
2.4. Challenges of measuring psychosocial constructs	7
<b>3. Selection of new scales for Round 4</b>	<b>8</b>
3.1. Overview of procedure for selecting new scales	8
3.2. Review of scales included in previous rounds	8
3.3. Identification of new psychosocial scales	9
<b>4. Piloting of the scales (psychometric properties)</b>	<b>10</b>
<b>5. Psychosocial scales administered in Round 4</b>	<b>11</b>
5.1. Adaptation and administration of scales	11
<b>6. Reliability and validity of psychosocial scales introduced in Round 4</b>	<b>12</b>
6.1. Assessing reliability: inter-item correlations and Cronbach's Alpha	13
6.2. Confirmatory factor analysis	15
6.3. Multi-group confirmatory factor analysis	15
6.4. Composite score: raw	16
<b>7. Reliability and validity of retained scales (Round 2 and Round 3) in Round 4</b>	<b>18</b>
7.1. Assessing reliability: inter-item correlations and Cronbach's Alpha coefficient	18
7.2. Composite score: raw	19
<b>8. Conclusion</b>	<b>20</b>
<b>9. References</b>	<b>21</b>
<b>Appendices</b>	<b>23</b>
Appendix A: Psychosocial skills measured in Round 2 and Round 3	23
Appendix B: CFA results from pilot across countries (revised fit statistics)	27
Appendix C: Adjusted reliability of the scales	27
Appendix D: Variables included in new scales for Round 4	28
Appendix E: Specified models for confirmatory factor analysis (Younger Cohort)	29
Appendix F: Fit statistics for confirmatory factor analysis (CFA)	30
Appendix G: Fit statistics for multi-group confirmatory factor analysis (MGCFA)	31
Appendix H. Results for partially constrained models	32
Appendix I: Variables from Round 3 included in retained scales for Round 4	33

## The authors

**Louise Yorke** is a Research Assistant at the Research for Equitable Access and Learning (REAL) Centre at the Faculty of Education, University of Cambridge. She has a BA (University College Dublin) and MSc in Psychology (Trinity College Dublin). She recently submitted her PhD at the School of Social Work and Social Policy, Trinity College Dublin, which focused on the lives, experiences and choices of rural girls as they negotiate different pathways to urban secondary schools in Ethiopia. Her research interests involve the use of quantitative qualitative and participatory research methods to understand educational access and equity, social and cultural factors influencing the educational experiences and trajectories of girls and young women, and school-to-work transitions.

**María José Ogando Portela** is a Young Lives Associate whose research focuses on psychosocial well-being and violence against children, with a particular emphasis on education. She spent two years in Rwanda as an ODI Fellow in the National Institute of Statistics of Rwanda (NISR). From 2011 to 2015 she worked with Young Lives as a Research Assistant and survey coordinator of the fourth round of the quantitative surveys in Ethiopia, and in the development of the fifth round. She holds a MSc in Statistics from the London School of Economics and a BA in Economics from Universidad Iberoamericana in Mexico City.

## Acknowledgements

We would like to thank Abhijeet Singh for his internal discussion note on the Round 2 and Round 3 Young Lives psychosocial scales which guided this work (Singh 2011), and Sofya Krutikova, Zoe James, Juan Leon, Caine Rolleston, Abhijeet Singh, and Martin Woodhead for their work in the development and selection of the final scales. We are grateful to Patricia Espinoza for her comments on earlier drafts of this paper, and to Santiago Cueto, Alula Panhurst, Renu Singh and Thang Nguyen for their in-country support.

# 1. Introduction

The importance of positive psychosocial functioning for children's current well-being and for their future outcomes is recognised across a range of disciplines, including psychology, economics and sociology. Nonetheless, research on the psychosocial development of individuals in developing countries is in its infancy. This technical note describes the procedure involved in the selection, adaptation and administration of the scales administered in the Young Lives Round 4 survey in 2013-14 to capture the development of psychosocial skills of 12-year-old children and young adults age 19 across all four study countries (Ethiopia, India (the states of Andhra Pradesh and Telangana), Peru and Vietnam). The note introduces what psychosocial skills are and why they are important, then provides an overview of some of the challenges of measuring psychosocial constructs, particularly in developing countries. After outlining the selection, adaptation and administration of psychosocial scales in Round 4 of Young Lives data collection, the basic properties of the items in each scale are assessed using traditional descriptive statistics including inter-item correlations, initial reliability using Cronbach's Alpha and the properties of the individual items. The psychometric properties of the scales are evaluated through a confirmatory factor analysis, and multi-group factor analysis is performed on the scales to investigate whether the scales are equivalent across countries.

## 2. Psychosocial skills

### 2.1. What are psychosocial skills?

Psychosocial skills refer to behavioural attributes of the individual, such as 'motivation', 'self-concept' and 'self-esteem'. Psychosocial skills are sometimes referred to as 'non-cognitive' skills (other terms such as 'competencies', 'soft skills' and 'life skills' are also used). However, the term 'non-cognitive' can sometimes be misleading as psychosocial skills do not exclude cognition. For this reason, it was decided to use the term 'psychosocial skills' or 'psychosocial functioning' to capture these attributes. Furthermore, the term 'psychosocial' captures the interactional nature of these constructs between the individual and the environment.

### 2.2. Why are they important?

Psychosocial skills are considered to be important for cognitive skill acquisition. Some studies suggest that psychosocial skills may be more important than cognitive skills, as high levels of cognitive skills do not guarantee success in life: individuals with high levels of cognitive skills and low levels of psychosocial skills often fail to succeed in life, while those with low cognitive levels and high psychosocial skills often have successful outcomes (Heckman and Rubinstein 2001).<sup>1</sup> Furthermore, the development of cognitive and psychosocial skills is found to be mutually reinforcing, whereby the acquisition of skill in one domain facilitates the acquisition of skills in others (Dercon and Krishnan 2009; Heckman 2007). For example, self-esteem can

---

<sup>1</sup> Cognitive skills indicate the basic mental abilities that are usually associated with intelligence and problem solving and directly impact the acquisition of literacy and numeracy skills (Adams 2011; Brunello and Schlotter 2010).

enhance academic outcomes which in turn enhance future self-esteem (Marsh and O'Mara 2008). Although it is difficult to discern a direction of causality between cognitive and psychosocial skills, evidence suggests that the influence of psychosocial skills on cognitive skills is stronger than the influence of cognitive skills on psychosocial skills (Borghans et al. 2008). For this reason, it has been suggested that the development of psychosocial skills may have the potential to remediate deficits in earlier cognitive development. This is strengthened by studies that have suggested that while cognitive skills are relatively crystallised by age 8, psychosocial skills are more malleable later in life. Thus, there is a wider window of opportunity for individuals to develop psychosocial skills, and they may help to remediate or compensate for shortcomings in other domains later in life (Borghans et al. 2008). As psychosocial skills are susceptible to external influences, it is thought that they can be developed and nurtured in children and young people.

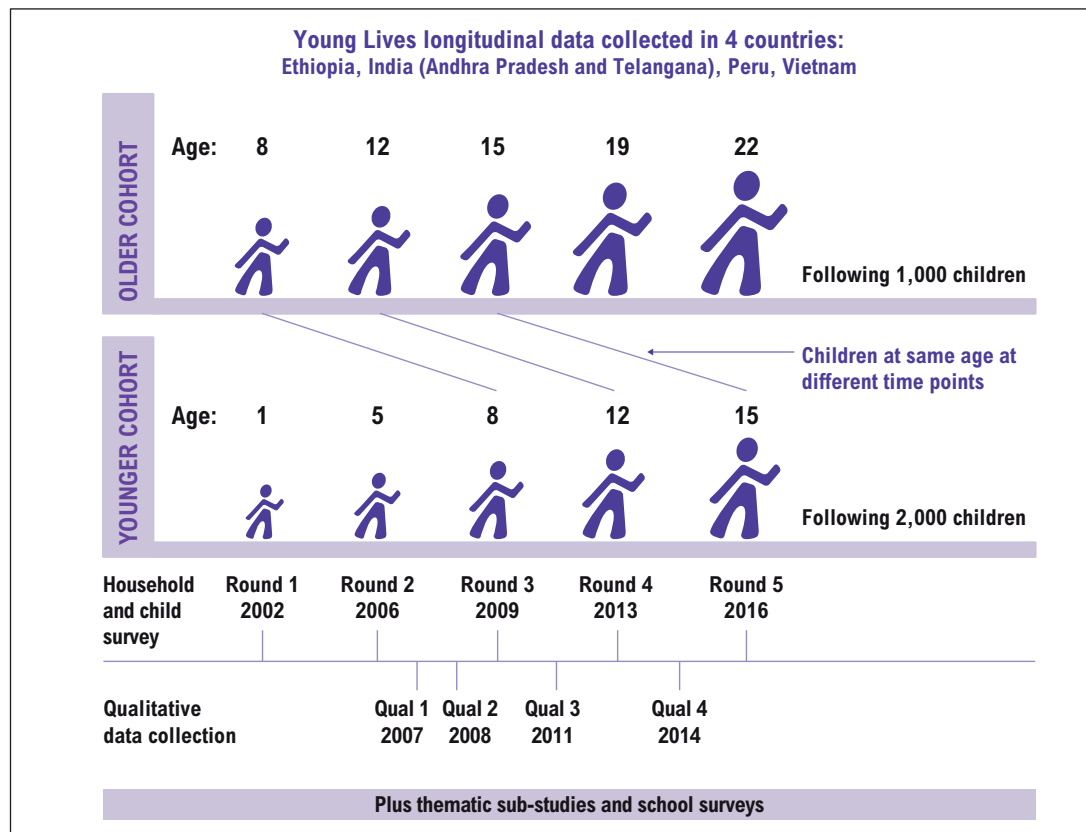
In terms of education, psychosocial skills also have a direct influence on the educational achievement of students, the level of education they complete, and their educational aspirations (Adams 2011; Brunello and Schlotter 2010; Dercon and Krishnan 2009). Therefore, educational approaches which foster psychosocial skills may have particular benefits. In addition, psychosocial skills may influence the entry of individuals into the labour market and there is evidence that young adults' psychosocial attributes predict future labour market outcomes (Caspi et al. 1998).

Taken together, this evidence suggests that increased understanding of the development of psychosocial skills could lead to a better and more holistic understanding of children's lives. Nonetheless, it is important to note that because the development of psychosocial skills is dependent on the interaction between the individual and the environment, individual children should not be seen as 'successful' or 'unsuccessful' based on the level of psychosocial skills they have. Rather, attention should be given to the environment in which children are located, particularly those living in situations of poverty and disadvantage. Psychosocial skills should not distract attention away from other structural issues that create conditions of poverty and disadvantage, but should be seen as one part of an integrated approach to understanding and improving children's lives.

### **2.3. Psychosocial skills administered in Young Lives**

In developing countries, large-scale survey data that include measures of psychosocial constructs are rare and as such, little is known about the process of the development of these skills in these contexts. Young Lives has contributed to closing this knowledge gap through extensive data collection, including a range of measures of psychosocial skill domains, administered in schools and households. These scales were first administered in Round 2 to Older Cohort children, and subsequently to both cohorts in Round 3 (Figure 1).

**Figure 1.** *Young Lives study*



This information, along with the psychosocial scales introduced in Round 4 (described here), will help to answer questions such as how early childhood conditions of poverty affect the development of these skills, and also the importance of these skills for children’s current and future outcomes. For example, they may facilitate an exploration of factors which impact the development of psychosocial skills at different stages of children’s trajectories in different contexts (using previous data), while also providing insight into the relationship between psychosocial skills and children’s future life outcomes, such as academic outcomes and entry into the labour market (using data from upcoming rounds).

## 2.4. Challenges of measuring psychosocial constructs

Measuring psychosocial constructs in any context is difficult as they are not directly observable and therefore cannot be measured directly. When researchers want to measure psychosocial constructs they must operationally define the construct in terms of behaviours believed to represent it. These behaviours then serve as measurable indicators of the underlying construct. Therefore, a critical starting point for research on psychosocial skills and functioning is the development of psychometrically sound instruments that are theoretically supported (Marsh 2005). However, most research on psychosocial skills has originated in Western contexts, and as psychosocial constructs are influenced by cultural factors there are many complications in administering psychosocial scales designed for one context to another, particularly in developing country contexts. Strong evidence is needed to suggest that a scale originating in one context is measuring the same construct in a different context or that it is culturally relevant (Bryne 2000). Thus, each time a scale is adapted, its reliability and validity should be reassessed to demonstrate that the same qualities are being measured in the new context (Douglas and Nijssen 2003; Geisinger 1994; Reise, Waller and Comrey 2000).

## 3. Selection of new scales for Round 4

### 3.1. Overview of procedure for selecting new scales

The psychosocial scales included in Round 4 came from two sources: (1) scales and items included in previous rounds of Young Lives;<sup>2</sup> and (2) other questionnaires with previously validated and theoretically grounded scales that are relevant to children's lives. The Psychosocial Skills Working Group (PSWG) was developed to oversee the selection of the final scales to be included in Round 4.<sup>3</sup> A systematic procedure was followed. First, the scales and items administered in previous rounds to measure psychosocial skills were reviewed. Second, a review of evidence from the literature concerning relevant psychosocial skills for the Young Lives survey was undertaken. Third, based on these reviews, an initial set of scales relevant to the objectives of Young Lives was selected and adapted for use. Fourth, these scales were piloted and their psychometric properties assessed, including the reliability and validity of the scales in each country. Lastly, the final set of scales to be administered were chosen and relevant adaptations made before administering the scales to the Young Lives children.

### 3.2. Review of scales included in previous rounds

Consideration was given to the psychosocial components that were included in previous rounds, particularly the constructs of 'pride/self-esteem', 'self-efficacy/agency', 'trust' and 'sense of inclusion' (listed in Appendix A). This involved identifying the origin of the scales and reviewing the scales in relation to the literature, reviewing the psychometric properties and performance of the scales, and discussing the importance of these scales. Table 1 provides a summary of the reliability of the psychosocial scales included in previous rounds for the Older Cohort.

**Table 1.** *Reliability ( $\alpha$ ) for Round 2 and Round 3*

Scale	Based on	Cronbach's Alpha ( $\alpha$ )		
		Round 2	Round 3	Round 4
Pride and self-esteem	Rosenberg (1965).	0.59 - 0.70	0.41 - 0.74	Yes
Agency/self-efficacy	Based on Rotter (1966); Bandura (1993).	0.22 - 0.39	0.29 - 0.39	Yes
Inclusion	-	0.29 - 0.57	0.41 - 0.59	No
Trust	-	0.44 - 0.53	0.22 - 0.54	No
Subjective well-being	Cantril's Ladder (Cantril 1965).	n/a	n/a	Yes

2 See Appendix A for a description of these.

3 Members of this group include Abhijeet Singh, Caine Rolleston, Martin Woodhead, and Sofya Krutikova.



Based on the review of the previous items and scales administered the following recommendations were made:

- **Pride and self-esteem:** While it was thought to be important to retain these items in future rounds, it was suggested to include a more general measure of self-esteem or a measure of self-esteem relating to a particular domain of functioning. As a result, this scale was retained to keep panel consistency, but complemented with an additional measure of self-esteem (reviewed in detail in Section 3.3 and Section 6).
- **Agency/self-efficacy:** Self-efficacy is defined as one's belief in one's capabilities to produce given attainments and plays a key direct and indirect role in human functioning (Bandura 2006). In Round 2 and Round 3 the items pertaining to agency/self-efficacy captured different domains such as school, work and time use. Agency and self-esteem seem to have performed well in Round 2 and Round 3. Thus, it was recommended to include these items in Round 4 in a comparable form to the previous rounds. It was also suggested that a more comprehensive and theoretically grounded measure of general self-efficacy at the broadest domain be included.
- **Sense of inclusion and trust:** The items pertaining to children's sense of inclusion and trust gave an overview of young people's support networks but did not provide information on the type or degree of support that they received. These items seemed to have low reliability and predictive power in previous rounds, and it was suggested to exclude these items from Round 4.
- **Subjective well-being:** The Ladder of Life Question (Cantril's Ladder) was administered in Round 2 and Round 3 to assess current satisfaction with life. It has been used widely in country reports and papers, is a low burden on participants, and was therefore recommended for inclusion in Round 4.

### 3.3. Identification of new psychosocial scales

In addition to the existing scales, there was agreement to include new items that may be relevant for the Young Lives sample, recognising that different constructs may become more relevant to children's lives as they grow older. It was decided at the outset that the scales should be theoretically grounded, inexpensive, freely available and have demonstrated good reliability and validity in previous studies. A review of the literature in relation to children's psychosocial functioning was undertaken, paying close attention to the possible implications of different constructs for the Young Lives children and the importance of contextual relevance. The two main aims were to identify constructs that may build upon previous findings and to identify new constructs which may be useful for inclusion to understand children's life trajectories. As a result, the following five scales were selected to be administered in the pilot:

**1. General self-efficacy (Schwarzer and Jerusalem 1995):** Self-efficacy is defined as one's belief in one's capabilities to produce given attainments and to cope with adversity. It facilitates goal-setting, effort investment, persistence in the face of barriers, and recovery from setbacks. The general self-efficacy scale (Schwarzer and Jerusalem 1995) was created to assess a general sense of perceived self-efficacy rather than specific behaviour. The scale is created for adolescents (12<) and the adult population. It has been adapted to many countries (including Hong Kong, India, Indonesia, Iran, Korea, Peru and Syria) and the findings from such studies confirm that the measure is reliable and unidimensional across cultures (Scholz et al. 2002). An example of a question from this scale is: 'I can always manage to solve difficult problems if I try hard enough'.

The remaining scales (general self-esteem, emotional stability, parent relations, and peer relations) were taken from the self-description questionnaires (I, II and III). The self-description questionnaires are based on the multidimensional structure rooted in the Shavelson, Hubner and Stanton (1976) theoretical models of self-concept. Self-concept is referred to as an individual's perceptions of the self which is formed in conjunction with the environment and interactions with others (Marsh and Shavelson 1985; Shavelson et al. 1976). The scales are considered to be some of the most validated self-concept measures available (Bryne 1996), having been extensively validated to establish their psychometric soundness. The instruments are widely used and freely distributed.

**2. General self-esteem (SDQ I):** Self-esteem refers to an individuals' judgement of their own self-value or self-worth. Global self-esteem is strongly related to measures of well-being. An example of a question measuring general self-esteem is: 'In general, I like being the way I am'.

**3. Emotional stability (SDQ III):** Indicates an individual's ability to regulate/balance their emotions and is considered an important aspect of human development and functioning. An example of an item measuring emotional stability is: 'I am usually pretty calm and relaxed'.

**4. Parent relations (SDQ II):** Indicates positive relations with parents. An example of an item measuring positive relations with parents is: 'My parents and I spend a lot of time together'.

**5. Peer relations (SDQ I):** Relates to positive relations with peers. An example of an item measuring positive peer relations is: 'I get along with other kids easily'.

## 4. Piloting of the scales (psychometric properties)

The scales were adapted and translated into the target languages and back-translated into English. The five scales were piloted in the four Young Lives countries: Ethiopia (n=314), India (n=332), Peru (n=286) and Vietnam (n=100). The data were prepared by recoding any negatively phrased items and values that were identified as 'missing at random' were deleted using listwise deletion. The internal consistency of each of the scales was tested using Cronbach's Alpha, while the unidimensionality of the scales was tested using confirmatory factor analysis (CFA).

CFA helps to test for construct validity by identifying whether the items in the scale contribute to an underlying trait which they are designed to measure. A good-fitting model indicates the items are contributing to an underlying factor (unidimensionality), while a bad-fitting model indicates that all the items are not contributing to the scale (and may indicate multi-dimensionality). CFA was run for each scale in each country to test the unidimensionality of the scale. Any item that was not working well in the scales was removed on an iterative basis (standardised regression weights below .3), and CFA was re-run for the adjusted models.<sup>4</sup> Reliability for the adjusted scales was then tested.<sup>5</sup>

Overall, the *general self-efficacy* scale and the *peer relations* scale were found to be working well across all four countries and demonstrated moderate-high reliability. The *emotional*

---

4 See Appendix B for fit statistics across country.

5 See Appendix C for results.

*stability* scale did not work well across the countries. The *general self-esteem* and *parent relations* scales did not work well across countries; however, it was suggested that this was due to the influence of the reverse-coded items in the scales (which potentially complicated the process of translation and interpretation), whereas those scales that were all positively coded were found to function well.

## 5. Psychosocial scales administered In Round 4

### 5.1. Adaptation and administration of scales

Based on the results of the pilot study, the PSWG decided to retain the following scales for the Round 4 survey:

- Generalised self-efficacy (Schwarzer and Mathias 1979)
- General self-esteem (self-description questionnaire I)
- Parent relations (self-description questionnaire II)
- Peer relations (self-description questionnaire I)

These scales were administered to children from the Younger Cohort (12 years old) and Older Cohort (19 years old). A subset of 24 items, excluding those items on the self-efficacy scale, was also administered to the sibling of the Younger Cohort child.<sup>6</sup> Appendix D lists the items (34 in total) administered to measure each competency in all four countries: 10 items on self-efficacy and eight each on self-esteem, parents relations, and peer relations.

**Table 2.** *Number of children included in the Round 4 survey*

	Older Cohort	Younger Cohort	Total
Ethiopia	908	1,873	2,781
India	952	1,915	2,867
Peru	635	1,902	2,537
Vietnam	887	1,931	2,818
Total	3,382	7,621	11,003

Children were asked to respond to each item on a Likert scale ranging from '1=strongly disagree' to '4=strongly agree'. Most items were applicable to all children, with the exception of those belonging to the parent relations scale, which were skipped in cases where both parents were dead at the time of the interview.<sup>7</sup> Additional codes for item refusals and 'don't know' responses were allowed in all countries, with the exception of Vietnam. Item refusals or 'don't know' replies together were very low for both cohorts, accounting between 0-3 per

6 In Round 3 (2009), information was collected on the next close-in-age sibling, in the domains of anthropometrics, cognitive and psychosocial development, education and time use.

7 In the Older Cohort, this was the case for 4.4 per cent of the children in Ethiopia, 2.5 per cent in Vietnam, 1.9 per cent in India, and 0.3 per cent in Peru; while in the Younger Cohort it was only the case for 1.3 per cent of the children in Vietnam, 1 per cent in Ethiopia, and 0.5 per cent in India and Peru.

cent of the responses in all items in Ethiopia and Peru, and in 29 of the 34 items in India. Among the five remaining items in India, 'don't know' answers together with refusals (although in a much lower proportion) accounted between 3-5 per cent of the responses in three items, and between 7.5-10.5 per cent of the responses in the following items (but for the Younger Cohort sample only): 'It is easy for me to stick to my aims and accomplish my goals' (10.5 per cent), and 'Overall, I have a lot to be proud of' (7.7 per cent).

All items were positively worded and measured in the same direction, whereby a greater value in the coded response of any item indicates more of a good outcome (i.e. more self-efficacy, more self-esteem, better parent relations or better peer relations, respectively). Item refusals and 'don't know' answers were recoded to missing values. The scales were adapted and translated into the target languages and back-translated into English, and then administered in the four countries.

## 6. Reliability and validity of psychosocial scales introduced in Round 4

In this section, we assess the individual functioning of items as well as the reliability of the scales by looking at the consistency of responses across the items included in each of the scales. Clark and Watson (1995) argue that the ultimate goal of test developers is to achieve unidimensionality – that is, whether the items in the scale are measuring a single underlying construct—rather than a good degree of internal consistency per se. Following their recommendation, we assess both the internal consistency and stability of responses across all items designed to measure each psychosocial competence, and the unidimensionality of the scale. For the former, we use the inter-item correlations within each index and the closely related Cronbach Alpha. For the latter, we perform confirmatory factor analysis.

The idea behind looking at inter-item correlations is simple: once all the questions are recoded in the same direction, such that 'a greater value represents a better outcome', the directions of paired correlation between all the items should all be positive. For example, if 'being proud of your clothes' and 'proud of your shoes' are both positive things for a measure of self-esteem, we would usually expect these to be positively correlated. A low inter-item correlation may suggest that two items are not sufficiently related, whereas high inter-item correlation may suggest items are very similar and that the scale is not parsimonious. Inter-item correlations can also be used to evaluate our cultural assumptions, about what is good in particular contexts and whether these assumptions are borne by the data. For example, a positive response to the statement 'Other people in my family make all the decisions about how I spend my time' might be a negative thing in an individualistic culture, but not in cultures where obedience is given a very high premium and letting others decide how you spend your time might be a result of active choice and agency. Therefore, examining inter-item correlation can help to check our assumptions about how items work in a particular context.

The Cronbach's Alpha ( $\alpha$ ) is a commonly used measure of scale reliability for scales composed of scores on individual items in a battery of indicators. The reliability  $\alpha$  is defined as the square of the correlation between the measured scale and the underlying factor. In general, a rule-of-thumb value of 0.7 (or more) is used to accept a scale as reliable (Nunnally and Bernstein 1994: 264-265). Nonetheless, a high  $\alpha$  value does not indicate

unidimensionality. Therefore, confirmatory factor analysis (CFA) is used to test the unidimensionality of the scales.

Section 6.1 shows the results from both statistics. Sections 6.2 and 6.3 present the results of CFA performed to test for construct validity and unidimensionality.<sup>8</sup>

## 6.1. Assessing reliability: inter-item correlations and Cronbach's Alpha

Table 3 presents the average inter-item correlations for all four psychosocial competencies introduced in Round 4 and administered to both cohorts and the sibling of the Younger Cohort. All average inter-item correlations fall in the recommended range of 0.15-0.50 (Clark and Watson 1995), going from 0.187 (Younger Cohort self-efficacy scale in Peru) to 0.368 (Older Cohort parents relations scale in Peru). This range is recommended to capture a variety of constructs of different scopes – narrower constructs are expected to have higher mean inter-item correlation than broader constructs – while at the same time avoiding item redundancy stemming from very high individual correlations.

**Table 3.** *Average inter-item correlations*

		Ethiopia	India	Peru	Vietnam
<b>Self-efficacy scale</b>	<i>Older Cohort</i>	0.254	0.315	0.253	0.196
	<i>Younger Cohort</i>	0.286	0.311	0.187	0.195
	<i>Sibling (Younger Cohort)</i>	-	-	-	-
<b>Self-esteem scale</b>	<i>Older Cohort</i>	0.257	0.206	0.246	0.227
	<i>Younger Cohort</i>	0.297	0.268	0.231	0.226
	<i>Sibling (Younger Cohort)</i>	0.293	0.292	0.228	0.204
<b>Parent relations scale</b>	<i>Older Cohort</i>	0.278	0.357	0.368	0.349
	<i>Younger Cohort</i>	0.297	0.358	0.327	0.344
	<i>Sibling (Younger Cohort)</i>	0.309	0.367	0.366	0.332
<b>Peer relations scale</b>	<i>Older Cohort</i>	0.308	0.322	0.350	0.269
	<i>Younger Cohort</i>	0.318	0.360	0.265	0.247
	<i>Sibling (Younger Cohort)</i>	0.342	0.320	0.246	0.243

Of the four psychosocial competencies, one can argue that parental and peer relations are a narrower constructs than self-efficacy and self-esteem. While the former are assessed by more tangible questions ('I like my parents'; 'I have a lot of friends') the latter comprises more appraisal-type measures ('When I am confronted with a problem, I can usually find several solutions'; 'Overall, I have a lot to be proud of'). It is not surprising then that the average inter-item correlations for these competencies are higher than those for the self-efficacy and self-esteem scales.

Even though the average inter-item correlation cannot alone be used as measure of unidimensionality, looking at each of the individual inter-item correlations and their range can help assess this. For example, it is possible to have more than one construct measured in the same scale yielding a moderate average inter-item correlation as a result of having a fraction of the items in the scale correlating very highly with each other and the remaining items correlating very low. Clark and Watson (1995: 316) therefore recommend that for

<sup>8</sup> See Section 4 for a description of the confirmatory factor analysis.

unidimensionality to be possibly ensured, 'all of the individual inter-item correlations should be moderate in magnitude and should cluster narrowly around the mean value'.

Table 4 summarises the number of out-of-range individual inter-item correlations for each cohort and competencies, where the range considered is 0.15-0.50.<sup>9</sup> Very few correlations lie outside the range, with the exception of the self-efficacy scale in Peru, the self-efficacy and self-esteem scale in Vietnam, and the self-esteem scale in India (Older Cohort), where out-of-range correlations account for a fifth to one third of total correlations. Most of these correlations lie below the range and those which lie above rarely go above 0.600 (the highest correlation – 0.652 – is found in the Younger Cohort sample in Ethiopia between the items 'I have lots of friends' and 'I have more friends than most other kids'). In Peru and Vietnam (both cohorts), the item 'If someone opposes me, I can find the means and ways to get what I want' correlates very poorly with most of the rest of the items in the scale and accounts for most of the below-range correlations; while in the self-esteem scale the same is true for the item 'In general, I like being the way I am' in India and Vietnam. Overall, the individual inter-item correlations do not seem to show a presence of potential multidimensionality in our scales (and of redundancy of items within scales).

**Table 4.** *Number of inter-item correlations out of range*

		Ethiopia	India	Peru	Vietnam
<b>Self-efficacy scale</b> (45 total inter-item correlations)	<i>Older Cohort</i>	1	0	9	12
	<i>Younger Cohort</i>	2	3	14	11
	<i>Sibling (Younger Cohort)</i>	-	-	-	-
<b>Self-esteem scale</b> (28 total inter-item correlations)	<i>Older Cohort</i>	1	7	0	5
	<i>Younger Cohort</i>	0	1	0	3
	<i>Sibling (Younger Cohort)</i>	0	0	2	6
<b>Parent relations scale</b> (28 inter-item total correlations)	<i>Older Cohort</i>	2	1	2	1
	<i>Younger Cohort</i>	0	0	0	0
	<i>Sibling (Younger Cohort)</i>	1	0	0	0
<b>Peer relations scale</b> (28 total inter-item correlations)	<i>Older Cohort</i>	2	1	0	0
	<i>Younger Cohort</i>	1	1	0	0
	<i>Sibling (Younger Cohort)</i>	1	1	0	1

Table 5 presents the Cronbach's Alpha coefficients for each of the competencies. As mentioned in Section 6, a coefficient of 0.7 is used as a rule of thumb for considering a scale reliable. The  $\alpha$  coefficient is a function of the inter-item correlations across items and the number of items in the scale, and so it is possible to achieve reliable measures by increasing either the number of items in the scale or by having very correlated items (Cortina 1993). We have shown the inter-item correlations in our scales are moderately low and our scales only comprise eight to 10 items; therefore we feel confident that our estimates in Table 5 represent a measure of internal consistency which is also parsimonious. All our scales lie above this value, except the self-esteem scale in the Older Cohort sample in India and in the sibling of the Younger Cohort in Vietnam which lie on an acceptable boundary close to 0.7.

<sup>9</sup> We do not provide full tables of the inter-item correlations.



**Table 5.** *Cronbach's Alpha coefficients*

		Ethiopia	India	Peru	Vietnam
<b>Self-efficacy scale</b>	<i>Older Cohort</i>	0.773	0.821	0.772	0.709
	<i>Younger Cohort</i>	0.800	0.819	0.697	0.708
	<i>Sibling (Younger Cohort)</i>	-	-	-	-
<b>Self-esteem scale</b>	<i>Older Cohort</i>	0.735	0.675	0.723	0.701
	<i>Younger Cohort</i>	0.772	0.745	0.707	0.701
	<i>Sibling (Younger Cohort)</i>	0.769	0.768	0.703	0.672
<b>Parent relations scale</b>	<i>Older Cohort</i>	0.755	0.816	0.823	0.811
	<i>Younger Cohort</i>	0.772	0.817	0.796	0.807
	<i>Sibling (Younger Cohort)</i>	0.781	0.823	0.822	0.799
<b>Peer relations scale</b>	<i>Older Cohort</i>	0.781	0.792	0.811	0.746
	<i>Younger Cohort</i>	0.788	0.818	0.743	0.724
	<i>Sibling (Younger Cohort)</i>	0.806	0.791	0.723	0.720

## 6.2. Confirmatory factor analysis

Confirmatory factor analysis (CFA) using maximum likelihood was used to test the validity of the four scales administered to the Younger Cohort in Ethiopia, India, Peru and Vietnam. This identifies whether the items in the scale contribute to an underlying trait which they purport to measure (unidimensionality).

First, listwise deletion was used to treat the missing data, with the remaining participants for each country as follows: Ethiopia (1,745), India (1,518), Peru (1,775) and Vietnam (695). The models for each scale were specified based on the one factor model in the literature.<sup>10</sup> Initial fit statistics were obtained for each of the models. There was some evidence of error co-variation among items and, where necessary, the model was re-specified to reflect this. The addition of these error co-variances are justified due to the item content overlap in the scales. CFA was re-run for the adjusted model, which improved the fit. Appendix F shows the final fit statistics for the Older Cohort, Younger Cohort, and Younger Cohort sibling. Overall, the scales were found to be working well in each of the four countries, suggesting that the models are unidimensional and thus that all items are contributing to the latent factor.

## 6.3. Multi-group confirmatory factor analysis

When an instrument is administered to more than one group it is important to investigate measurement invariance (MI) to determine whether the instrument is operating in the same way across groups and whether the underlying construct has the same theoretical structure for the groups under study (Byrne 2009; Geisinger 1994).<sup>11</sup> Failure to establish MI indicates that the meaning of the latent construct is shifting across groups and that interpretations of between-group differences are ambiguous (Cheung and Rensvold 2002). The general null hypothesis tested in multiple group models is that the data from each group are from the same population (Tabachnick and Fidell 2007). Measurement invariance investigates whether the underlying construct has the same theoretical structure across groups (Byrne 2004). If the scale does not meet MI, this indicates that different groups respond differently to the items and thus, as a consequence, factor means cannot be reasonably compared.

<sup>10</sup> See Appendix E for schematic representations of these models.

<sup>11</sup> Measurement invariance indicates that the same construct is being measured across groups.

A multi-group confirmatory factor analysis (MGCFA) was carried out in order to test the measurement invariance of each of the scales across the four countries.<sup>12</sup> Measurement invariance was investigated by first testing the configural invariance and then the metric invariance.<sup>13</sup> First, a baseline model was specified across groups. For all of the models, configural invariance was obtained, suggesting that the factor structure was equivalent across countries. A constrained model was then tested across groups. The chi-squared difference test was used to test whether there were differences across groups. For all the groups the test was significant, suggesting that there are differences between groups and that comparison across groups will not be valid. Nonetheless, the chi-square test can be problematic in large sample sizes and thus alternative methods of determining the equivalence of the scales across countries should be sought.

As metric invariance was not found between the groups, partial invariance was tested.<sup>14</sup> Although partial MI is considered as a compromise between full MI and complete lack of invariance (Byrne, Shavelson and Muthén 1989), once the invariant items comprise a majority of the items, or at least two loadings and intercepts are constrained to be equal across groups, then cross-group comparisons can still be made (Byrne, Shavelson and Muthén 1989). This was only carried out for scales in the Younger Cohort.<sup>15</sup> The results suggested that at least two items were invariant in each scale across countries. This suggests that while comparisons can be made across countries, analysis should be carefully interpreted.

#### **6.4. Composite score: raw**

To compute all psychosocial indices we followed an identical procedure across countries and cohorts:<sup>16</sup>

1. All item refusals and 'DK=don't know' responses were recoded into missing values.
2. Relevant questions were all normalised to z-scores within each country (subtract mean and divide by standard deviation).
3. An average of the relevant z-scores was taken across the non-missing values of the questions.

This is a common standardisation practice for scale construction when the researcher aims to explore factors influencing psychosocial outcomes or how psychosocial skills are associated with other outcomes (Fischer and Milfont 2010). It is referred to as within-group standardisation or z-score standardisation by variable, whereby each individual is positioned in relation to other individuals in the group for each item (Dawis 2000; Fischer and Milfont 2010). Then, the composite score is constructed by averaging the relevant z-scores.

---

12 See Appendix G for results.

13 Configural invariance indicates that the pattern of fixed and free factor is equivalent across groups.

14 Partial invariance indicates that only a subset of parameters in a model is constrained to be invariant while another subset of parameters is allowed to vary across groups (Byrne et al. 1989). Partial measurement invariance may allow appropriate cross-group comparisons even if full measurement invariance is not obtained.

15 See Appendix H for full results.

16 The procedure is very similar to that followed in previous rounds, but unlike Round 2 and Round 3, all of the items introduced in Round 4 are positively worded so there was no need to flip scales for specific items.



Step 1 and Step 3 together incorporate one assumption:

- In Round 4, all questions were administered to the full sample of children in both cohorts. In the case of children without parents, none of the items belonging to the parental relations scale were administered. Within each scale, missing z-scores out of the total list of questions are considered 'missing at random'. As noted in previous rounds, this raises issues of understanding how much the scores vary due to particular questions, given that the weight of each question within the scale varies according to the number of missing dimensions.

Additionally, Step 3 adds a second assumption:

- Equal weights are assigned across the questions in each of the indices without any further consideration. This implies, for instance, that responses to 'I get along with other kids easily' and 'I have more friends than most other kids' count equally in the measure of peer relations.

Table 6 shows the descriptive statistics of all psychosocial indices introduced in Round 4.

**Table 6.** Mean and standard deviations of indices introduced in Round 4

		Ethiopia			India			Peru			Vietnam		
		Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.
<b>Self-efficacy index</b>	<i>Older Cohort</i>	908	-0.00	0.57	951	-0.00	0.62	615	-0.00	0.57	885	-0.00	0.53
	<i>Younger Cohort</i>	1,867	-0.00	0.60	1,903	-0.01	0.63	1,877	-0.00	0.52	1,923	0.00	0.52
	<i>Sibling (Younger Cohort)</i>	-	-	-	-	-	-	-	-	-	-	-	-
<b>Self-esteem index</b>	<i>Older Cohort</i>	908	-0.00	0.59	952	-0.00	0.55	615	-0.00	0.58	885	-0.00	0.57
	<i>Younger Cohort</i>	1,868	-0.00	0.62	1,903	-0.00	0.60	1,879	-0.00	0.57	1,923	0.00	0.57
	<i>Sibling (Younger Cohort)</i>	1,485	-0.00	0.63	1,601	-0.02	0.63	767	-0.00	0.57	961	-0.00	0.55
<b>Parent relations index</b>	<i>Older Cohort</i>	908	-0.01	0.64	952	-0.01	0.68	612	-0.00	0.67	885	-0.00	0.66
	<i>Younger Cohort</i>	1,848	-0.00	0.62	1,894	-0.00	0.66	1,870	-0.00	0.64	1,899	-0.00	0.65
	<i>Sibling (Younger Cohort)</i>	1,477	0.00	0.63	1,596	-0.01	0.67	764	-0.01	0.67	950	0.00	0.64
<b>Peer relations index</b>	<i>Older Cohort</i>	908	-0.00	0.63	952	-0.00	0.64	615	-0.00	0.66	885	-0.00	0.60
	<i>Younger Cohort</i>	1,868	-0.00	0.64	1,903	-0.00	0.66	1,879	-0.00	0.60	1,923	0.00	0.58
	<i>Sibling (Younger Cohort)</i>	1,486	-0.00	0.65	1,603	-0.01	0.64	769	-0.00	0.58	961	0.00	0.58

## 7. Reliability and validity of retained scales (Round 2 and Round 3) in Round 4

In addition to the new scales introduced in Round 4, we retained two of the four scales previously administered in Rounds 2 and 3: the agency and the pride and self-esteem indexes (see Appendix A and I). These scales were administered in Round 4 in exactly the same way as Round 3 (i.e. wording was kept unchanged and children in all four countries were asked to respond on a scale ranging from '1=strongly disagree' to '5=strongly agree'). The number of items in the agency scale remained the same as in Round 3 for both the Younger and the Older Cohort and the sibling of the Younger Cohort. However, with the self-esteem scale, the two items related to education – 'I am never embarrassed because I do not have the right books, pencils and other equipment for school' and 'I am proud that I have the correct uniform' – were considered no longer relevant for the Older Cohort and thus removed for that sample. Appendix I presents the questions included in these retained scales from previous rounds.

Questions framed positively – the majority of the statements – are coded in a way that a greater value response is more of good outcome (i.e. more agency or more pride and self-esteem). Questions framed negatively (e.g. 'Other people in my family make all the decisions about how I spend my time' and 'I have no choice about the work I do – I must do this sort of work') are consistently recoded to positive.

### 7.1. **Assessing reliability: inter-item correlations and Cronbach's Alpha coefficient**

We assessed the reliability of the agency and the pride and self-esteem scales retained in Round 4 using two statistics: the inter-item correlations and Cronbach's Alpha coefficient. Regarding the agency scale, all average inter-item correlations shown in Table 7 across cohorts and countries, with the exception of the Younger Cohort sibling in Vietnam, are below the recommended range of 0.15-0.50 (Clark and Watson 1995). A deeper look at the individual inter-item correlations shows that not all of them are positive, as one would expect. Consistent with what was found for Rounds 2 and 3 (Singh 2011), the statement 'Other people in my family make all the decisions about how I spend my time' presents a consistent negative correlation with other items of the scale in 10 out of the 12 scales.<sup>17</sup> Similarly, the statement 'I have no choice about the work I do – I must do this sort of work' correlates negatively with other items of the scale in 10 out of 12 cases. As a result, the internal consistency of the scale is considerably low in all cases (Table 8).

On the other hand, all average inter-item correlations of the pride and self-esteem scale (Table 7) are moderately low, as one should expect with all individual items correlating positively across all scales. In comparison with Round 3 (Singh 2011), Cronbach's Alpha coefficients are slightly higher in all countries and also closer to the 0.7 rule of thumb in three out of the four countries.

---

<sup>17</sup> We have assumed this is a negative statement in terms of agency, but it might be positively perceived by children in the study countries.

**Table 7.** *Average inter-item correlations*

		Ethiopia	India	Peru	Vietnam
<b>Agency scale</b>	<i>Older Cohort</i>	0.120	0.121	0.142	0.131
	<i>Younger Cohort</i>	0.118	0.020	0.063	0.105
	<i>Sibling (Younger Cohort)</i>	0.131	0.040	-0.006	0.166
<b>Pride scale</b>	<i>Older Cohort</i>	0.440	0.322	0.281	0.363
	<i>Younger Cohort</i>	0.316	0.240	0.248	0.189
	<i>Sibling (Younger Cohort)</i>	0.309	0.297	0.241	0.180

**Table 8.** *Cronbach's Alpha coefficients*

		Ethiopia	India	Peru	Vietnam
<b>Agency scale</b>	<i>Older Cohort</i>	0.407	0.410	0.454	0.429
	<i>Younger Cohort</i>	0.402	0.103	0.253	0.371
	<i>Sibling (Younger Cohort)</i>	0.430	0.181	NA*	0.500
<b>Pride scale</b>	<i>Older Cohort</i>	0.760	0.655	0.610	0.695
	<i>Younger Cohort</i>	0.735	0.654	0.665	0.582
	<i>Sibling (Younger Cohort)</i>	0.729	0.717	0.656	0.570

Note: \* Alpha coefficient was not estimated.

## 7.2. Composite score: raw

Both psychosocial indices are produced through an identical procedure:

1. All relevant questions are recoded to be positive outcomes (Appendix C).
2. All item refusals and 'DK=don't know' responses are recoded into missing values.
3. Relevant questions are all normalised to z-scores within each country (subtract mean and divide by standard deviation).
4. An average of the relevant z-scores is taken across the non-missing values of the questions.

Descriptive statistics of these indices are presented in Table 9 .

**Table 9.** *Mean and standard deviations of retained indices*

		Ethiopia			India			Peru			Vietnam		
		Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.
<b>Agency index</b>	<i>Older Cohort</i>	908	0.00	0.55	952	0.01	0.56	615	0.00	0.56	885	0.02	0.58
	<i>Younger Cohort</i>	1,868	0.00	0.55	1,901	-0.00	0.50	1,876	-0.00	0.51	1,923	-0.00	0.55
	<i>Sibling (Younger Cohort)</i>	1,486	-0.00	0.56	1,579	-0.02	0.54	760	-0.00	0.45	960	-0.01	0.59
<b>Pride index</b>	<i>Older Cohort</i>	908	0.00	0.77	952	-0.00	0.71	615	0.00	0.68	885	0.00	0.73
	<i>Younger Cohort</i>	1,869	0.00	0.66	1,902	0.01	0.61	1,878	-0.00	0.61	1,923	-0.00	0.58
	<i>Sibling (Younger Cohort)</i>	1,486	-0.00	0.66	1,598	0.00	0.65	764	-0.00	0.61	960	-0.01	0.59

## 8. Conclusion

Research on the psychosocial development of individuals in developing countries is in its infancy. Young Lives has contributed to closing this knowledge gap through extensive data collection on a range of measures of psychosocial skill domains, administered in schools and households in developing countries, where large-scale survey data that include measures of psychosocial constructs are rare.

Overall, the scales introduced in Round 4 do not seem to show a presence of potential multidimensionality or redundancy of items within a scale. Most estimates across scales and cohorts represent a measure of internal consistency. With respect to scales from Round 2 and Round 3, the pride and self-esteem scale is close to the 0.7 rule of thumb for accepting a scale as reliable when looking at the Cronbach's Alphas. However, the agency scale does not perform to an acceptable level in any of the countries.

This technical note, describing the procedure followed at Young Lives for developing and testing the scales, can act as a guide for researchers interested in using Young Lives scales with the confidence that they have been tested thoroughly, and for other studies interested in filling the current knowledge gap, whether using the same or a different set of items.

## 9. References

- Adams, A.V. (2012) 'The role of skills development in overcoming social disadvantage', background paper prepared for the Education for all Global Monitoring Report 2012, Youth and Skills: Putting Education to Work.
- Bandura, A. (1993) 'Perceived Self-efficacy in Cognitive Development and Functioning', *Educational Psychologist* 28.2: 117-148.
- Bandura, A. (2006) 'Guide for Constructing Self-efficacy Scales', *Self-efficacy Beliefs of Adolescents* 5.1: 307-337.
- Byrne, B.M. (2004) 'Testing for Multigroup Invariance Using AMOS graphics: A Road Less Traveled', *Structural Equation Modeling* 11.2: 272-300.
- Byrne, B.M. (2009) *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Abingdon: Routledge.
- Byrne, B.M., R.J. Shavelson, and B. Muthén (1989) 'Testing for the Equivalence of Factor Covariance and Mean Structures: The Issue of Partial Measurement Invariance', *Psychological Bulletin* 105.3: 456.
- Brunello, G., and M. Schlotter (2010) 'The Effect of Non-cognitive Skills and Personality Traits on Labour Market Outcomes'. Analytical Report for the European Commission prepared by the European Expert Network on Economics of Education (EENEE).  
[www.epis.pt/downloads/dest\\_15\\_10\\_2010.pdf](http://www.epis.pt/downloads/dest_15_10_2010.pdf) (accessed 2 February 2018).
- Borghans, L., A.L. Duckworth, J.J. Heckman, and B. Ter Weel (2008) 'The Economics and Psychology of Personality Traits', *Journal of Human Resources* 43.4: 972-1059.
- Caspi, A., B.R.E. Wright, T.E. Moffitt, and P.A. Silva (1998) 'Early Failure in the Labor Market: Childhood and Adolescent Predictors of Unemployment in the Transition to Adulthood', *American Sociological Review* 63.3: 424-451.
- Cantril, H. (1965) *The Pattern of Human Concerns*, New Brunswick: Rutgers University Press.
- Clark, L. and D. Watson (1995) 'Constructing Validity: Basic Issues in Objective Scale Development', *Psychological Assessment* 7: 309-319.
- Cheung, G.W., and R.B. Rensvold (2002) 'Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance', *Structural Equation Modeling* 9.2: 233-255.
- Cortina, J.M. (1993) 'What Is Coefficient Alpha? An Examination of Theory and Applications', *Journal of Applied Psychology* 78.1: 98-104.
- Dawis, R.V. (2000) 'Scale Construction and Psychometric Considerations', in H.E. Tinsley and S.D. Brown (eds.) *Handbook of Applied Multivariate Statistics and Mathematical Modelling*, 65-94, San Diego: Academic Press.
- Dercon, S., and P. Krishnan (2009) 'Poverty and the Psychosocial Competencies of Children: Evidence from the Young Lives Sample in Four Developing Countries', *Children, Youth and Environments*, 19.2: 138-163.
- Douglas, S.P., and E.J. Nijssen (2003) 'On the Use of "Borrowed" Scales in Cross-National Research: A Cautionary Note', *International Marketing Review* 20.6: 621-642.
- Fischer, R. and T.L. Milfont (2010) Standardization in Psychological Research, *International Journal of Psychological Research* 3.1: 88-96.

- Geisinger, K.F. (1994) 'Cross-cultural Normative Assessment: Translation and Adaptation Issues Influencing the Normative Interpretation of Assessment Instruments', *Psychological Assessment* 6.4: 304.
- Heckman, J.J. (2007) 'The Economics, Technology, and Neuroscience of Human Capability Formation', *Proceedings of the National Academy of Sciences* 104.33: 13250-13255.
- Heckman, J.J., and Y. Rubinstein (2001) 'The Importance of Noncognitive Skills: Lessons From the GED Testing Program', *American Economic Review* 91.2: 145-149.
- Marsh, H.W. (2005) 'Big-Fish-Little-Pond Effect on Academic Self-concept: A Reply to Responses', *Zeitschrift für Pädagogische Psychologie* 19.3: 141-144.
- Marsh, H.W., and A. O'Mara (2008) 'Reciprocal Effects Between Academic Self-concept, Self-esteem, Achievement, and Attainment over Seven Adolescent Years: Unidimensional and Multidimensional Perspectives of Self-concept', *Personality and Social Psychology Bulletin* 34.4: 542-552.
- Marsh, H.W., and R. Shavelson (1985) 'Self-concept: Its Multifaceted, Hierarchical Structure', *Educational Psychologist* 20.3: 107-123.
- Nunnally, J.C., and I.H. Bernstein (1994) *Psychometric Theory* (3<sup>rd</sup> Edition), New York: McGraw-Hill.
- Reise, S.P., N.G. Waller, and A.L. Comrey (2000) 'Factor Analysis and Scale Revision', *Psychological Assessment* 12.3: 287.
- Rosenberg, M. (1965) *Society and the Adolescent Self-image*, Princeton, NJ: Princeton University Press.
- Rotter, J.B. (1966) 'Generalized Expectancies for Internal Versus External Control of Reinforcement', *Psychological Monographs: General and Applied* 80.1: 1-28.
- Shavelson, R.J., J.J. Hubner, and G.C. Stanton (1976) 'Self-concept: Validation of Construct Interpretations', *Review of Educational Research* 46.3: 407-441.
- Schwarzer, R., and M. Jerusalem (1995) 'Optimistic Self-beliefs as a Resource Factor in Coping with Stress', in S.E. Hobfoll and M.W. de Vries (eds.) *Extreme Stress and Communities: Impact and Intervention*, 159-177, Dordrecht: Springer.
- Scholz, U., B.G. Doña, S. Sud, and R. Schwarzer (2002) 'Is General Self-efficacy a Universal Construct? Psychometric Findings from 25 countries', *European Journal of Psychological Assessment* 18.3: 242-251.
- Singh, A. (2011) 'Review of Psychosocial Skills in Round 2 and Round 3', internal discussion note, Oxford: Young Lives.
- Tabachnick, B.G., and L.S. Fidell (2007) *Using Multivariate Statistics*, Needham Heights, MA: Allyn & Bacon/Pearson Education.
- Tinsley, H.E., and S.D. Brown (eds.) (2000) *Handbook of Applied Multivariate Statistics and Mathematical Modeling*, San Diego, CA: Academic Press.

# Appendices

## Appendix A:

### Psychosocial skills measured in Round 2 and Round 3

The psychosocial information in Young Lives was established in Round 2 (2006). In each country, the Round 2 questionnaire asked for the child's opinion on a range of statements that were designed to capture four psychosocial competencies – agency (five items), pride and self-esteem (nine items), trust (four items) and inclusion (five items). Children in Ethiopia, India and Vietnam were asked to respond on a scale ranging from '1=strongly agree' to '4=strongly disagree'. In Peru children were asked instead to respond between '0=no', '1=yes' and '2=more or less/doubt'. Not all questions were applicable to all children. Those relating specifically to school or work were only administered to those children who were enrolled in school, or were undertaking work.<sup>18</sup>

In Round 3, certain questions that caused distress or difficulties were refocused or dropped following in-depth discussions in each country (Morrow 2013). As a result, in some countries several statements were changed from being negatively worded to positively worded (e.g. from 'I am ashamed of my clothes' to 'I am proud of my clothes') and several questions were dropped. These changes were country specific; Table A.1 reflects the precise number of questions used for each scale in each country. While we may hope that a double negative in the first statement corresponds exactly to a positive in the second, such symmetry is unlikely to be exact as the wording of the question would also change the responses.

**Table A.1.** *Number of questions administered in Round 3*

Competency	Ethiopia	India	Peru	Vietnam
Trust	3	2	3	2
Agency	5	5	5	5
Pride	5	5	5	5
Inclusion	3	3	3	3

In all four countries, the questions in Round 3 were asked on a 5-point 'happy-faces' scale that ranged from '1=strongly disagree' to '5=strongly agree', that is, in the opposite direction to the Round 2 scales and with an additional point in the scale. This shift in response format between rounds does not seem to have been essential, though it may affect the stability of the measure and can easily lead to coding errors.

### *Computation of psychosocial indices*

All psychosocial indices are produced through an identical procedure:

1. All relevant questions are recoded to be positive outcomes
2. Relevant questions are all normalised to z-scores within each country (subtract mean and divide by standard deviation)

<sup>18</sup> Questions relating to work circumstances were administered differently across countries. For example, at age 12 (Round 2) in Peru, India and Vietnam, work-related items were asked to children who reported doing 'something to get money or things for themselves or their family' (filter question); while in Ethiopia all children were asked these items.

3. An average of the relevant z-scores is taken across the non-missing values of the questions.<sup>19</sup>

Step 1 is intended to recode all statements such that all components are coded into positive outcomes, that is, more on any one question means a normatively positive improvement on the index and more of the index is more of a good outcome. This is conceptually simple but in practice prone to errors due to the changes of scales and phrasing between rounds. Table A.2 explains the conversion needed.

**Table A.2.** *Required conversion in individual items*

	Round 2	Round 3
Direction of original scale	Positive to negative (1=strongly agree to 4=strongly disagree)	Negative to positive (1=strongly disagree to 5=strongly agree)
Conversion required to change scales into positive outcomes	Flip scale symmetrically and flip negative statements; or flip scores on positive statements	Flip negative statements

Steps 2 and 3 both imply certain assumptions worth keeping in mind. Regarding Step 2, scores are normalised *within a particular round*. While this gives consistent *relative* rankings in each question in both rounds, it is not necessarily adequate for a comparison of scores *across rounds*. Each index is just a simple average across z-scores on different questions whose distributions have moved differently between rounds, so it is hard to see the effect of this normalisation on the comparison of an entire index over time. Step 3 incorporate two assumptions:

- Missing z-scores out of the total list of questions are considered ‘missing at random’. However, since at least few questions are missing due to non-eligibility (i.e. not enrolled or not engaged in work), this assumption is questionable. Furthermore, missing z-scores raises issues of understanding how much the scores are being shifted due to particular questions. For example, if one child answers four questions and a second child answers five, then the same question counts for a quarter of the index of the first child and a fifth of the index of the second child.
- Equal weights are assigned across the questions in each of the indices. This implies, for instance, that responses to ‘When I am at shops/market I am usually treated by others with fairness and with respect’, places where the child may go only rarely, and ‘The other children in my class treat me with respect’, children with whom he/she spends many hours every day, count equally in the measure of inclusion.

The list of questions used for each scale in both Round 2 and Round 3 is given in Table A.3.

<sup>19</sup> Some questions have differed in coverage across rounds and cohorts; thus the indices are calculated slightly differently.



**Table A.3.** *Items included in the psychosocial scales in Round 2 and Round 3*

Scales	Items
<b>Trust index</b>	<p>Younger Cohort Round 3 (8 years old):</p> <ol style="list-style-type: none"> <li>1. Most people in my neighbourhood can be trusted</li> <li>2. I believe the government does what is right for people like me</li> <li>3. I feel safe when I go out of the house on my own</li> </ol> <p>Older Cohort Round 2 (12 years old):</p> <ol style="list-style-type: none"> <li>1. Most people in my neighbourhood are basically honest</li> <li>2. Most people in my neighbourhood can be trusted</li> <li>3. I believe the government does what is right for people like me</li> <li>4. I feel safe when I go out of the house on my own</li> </ol> <p>Older Cohort Round 3 (15 years old):</p> <ol style="list-style-type: none"> <li>1. Most people in my neighbourhood can be trusted</li> <li>2. I feel safe when I go out of the house on my own</li> <li>3. I believe the government does what is right for people like me</li> </ol>
<b>Agency index</b>	<p>Younger Cohort Round 3 (8 years old):</p> <ol style="list-style-type: none"> <li>1. If I try hard, I can improve my situation in life</li> <li>2. Other people in my family make all the decisions about how I spend my time</li> <li>3. I like to make plans for my future studies and work</li> <li>4. If I study hard at school, I will be rewarded by a better job in the future</li> <li>5. I have no choice about the work I do – I must do this sort of work</li> </ol> <p>Older Cohort Round 2 (12 years old):</p> <ol style="list-style-type: none"> <li>1. If I try hard, I can improve my situation in life</li> <li>2. Other people in my family make all the decisions about how I spend my time</li> <li>3. I like to make plans for my future studies and work</li> <li>4. If I study hard at school I will be rewarded by a better job in future</li> <li>5. I have no choice about the work I do – I must work</li> </ol> <p>Older Cohort Round 3 (15 years old)</p> <ol style="list-style-type: none"> <li>1. If I try hard I can improve my situation in life</li> <li>2. Other people in my family make all the decisions about how I spend my time</li> <li>3. I like to make plans for my future studies and work</li> <li>4. If I study hard at school I will be rewarded by a better job in the future</li> <li>5. I have choice about the work I do</li> </ol>
<b>Pride and self-esteem index</b>	<p>Younger Cohort Round 3 (8 years old):</p> <ol style="list-style-type: none"> <li>1. I am proud of my shoes or of having shoes</li> <li>2. I am proud of my clothes</li> <li>3. I am never embarrassed because I do not have the right books, pencils or other equipment</li> <li>4. I am proud that I have the correct uniform</li> <li>5. I am proud of the work I have to do</li> </ol> <p>Older Cohort Round 2 (12 years old):</p> <ol style="list-style-type: none"> <li>1. I feel proud to show my friends or other visitors where I live</li> <li>2. I am ashamed of my clothes</li> <li>3. I am ashamed of my shoes</li> <li>4. I feel proud of the job my [RELATIONSHIP OF HOUSEHOLD HEAD TO THE CHILD] does</li> <li>5. I am often embarrassed because I do not have the right books, pencils and other equipment for school</li> <li>6. I am worried that I don't have the correct uniform</li> <li>7. I am proud of my achievements at school</li> <li>8. I am embarrassed by/ashamed of the work I have to do</li> </ol>

PSYCHOSOCIAL SCALES IN THE YOUNG LIVES ROUND 4 SURVEY:  
SELECTION, ADAPTATION AND VALIDATION

Scales	Items
	<p>9. The job I do makes me feel proud Older Cohort Round 3 (15 years old)</p> <p>1. I am proud of my shoes or of having shoes 2. I feel my clothing is right for all occasions 3. I am proud of my clothes 4. I am often proud because I have the right books, pencils and other equipment for school 5. I am proud that I have the correct uniform 6. I am proud of the work I have to do</p>
<b>Inclusion</b>	<p>Older Cohort Round 2 (12 years old)</p> <p>7. When I am at shops/market I am usually treated by others with fairness and with respect 8. Adults in my STREET/VILLAGE treat me worse than other children my age (recoded to positive) 9. The other children in my class treat me with respect 10. Other pupils in my class tease me at school (recoded to positive) 11. My teachers treat me worse than other children (recoded to positive)</p> <p>Older Cohort Round 3 (15 years old)</p> <p>1. The other children in my class treat me with respect 2. Pupils in my class never tease me at school 3. Adults in my community treat me as well as they treat other children my age</p>

## Appendix B: CFA results from pilot across countries (revised fit statistics)

	Model	$\chi^2$	$\chi^2/df$	TLI	CFI	RMSEA	LO 90	HI 90
	Threshold	p>0.05	<5	>.90	>.95	<0.05	0.02	0.08
<b>Ethiopia</b>	<i>Self-efficacy</i>	p > 0.05	1.16	.98	.98	.02	.00	.05
	<i>Self-esteem</i>	p < 0.05	5.52	.33	.57	.12	.10	.14
	<i>Parent relations</i>	p < 0.05	3.39	.78	.88	.09	.07	.11
	<i>Peer relations</i>	p < 0.05	2.10	.86	.92	.06	.03	.08
	<i>Emotional stability</i>	p < 0.05	3.94	.20	.49	.10	.08	.11
<b>India</b>	<i>Self-efficacy</i>	p < 0.05	2.17	.61	.75	.06	.04	.08
	<i>Self-esteem</i>	p < 0.05	4.17	.32	.57	.10	.08	.12
	<i>Parent relations</i>	p < 0.05	5.15	.66	.81	.11	.09	.13
	<i>Peer relations</i>	p < 0.05	1.63	.95	.97	.04	.01	.07
	<i>Emotional stability</i>	p < 0.05	2.03	.65	.78	.06	.04	.07
<b>Peru</b>	<i>Self-efficacy</i>	p < 0.05	2.01	.74	.83	.06	.04	.08
	<i>Self-esteem</i>	p < 0.05	3.66	.42	.63	.10	.08	.12
	<i>Parent relations</i>	p < 0.05	2.37	.79	.88	.07	.04	.10
	<i>Peer relations</i>	p < 0.05	1.71	.84	.91	.05	.02	.08
	<i>Emotional stability</i>	p < 0.05	1.76	.48	.67	.05	.03	.07
<b>Vietnam</b>	<i>Self-efficacy</i>	p > 0.05	1.30	.63	.76	.06	.00	.10
	<i>Self-esteem</i>	p < 0.05	2.43	.53	.70	.12	.09	.15
	<i>Parent relations</i>	p > 0.05	1.34	.89	.94	.06	.00	.11
	<i>Peer relations</i>	p < 0.05	3.23	.57	.76	.15	.11	.19
	<i>Emotional stability</i>	p > 0.05	.98	1.00	1.00	.00	.00	.18

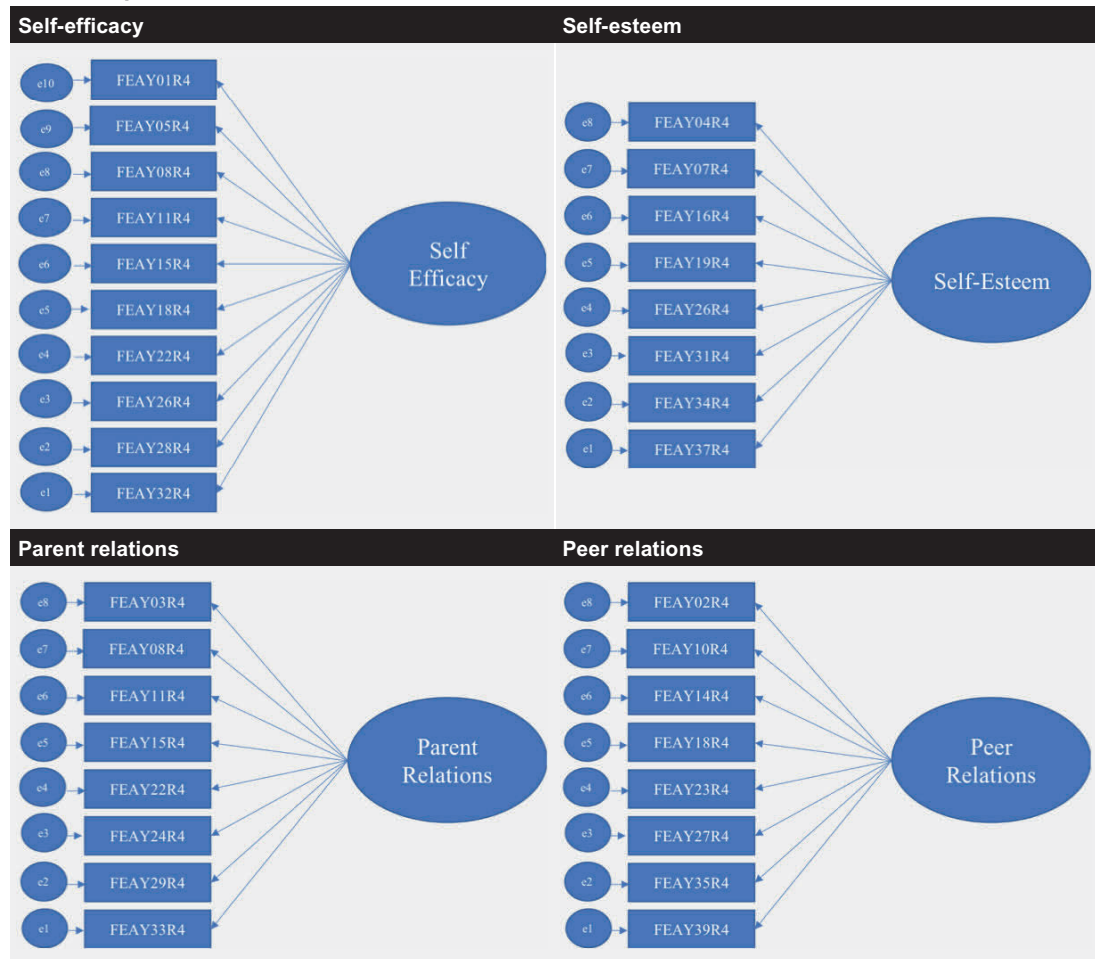
## Appendix C: Adjusted reliability of the scales

		<b>Ethiopia</b>	<b>India</b>	<b>Peru</b>	<b>Vietnam</b>
<b>Generalised self-efficacy</b>	<i>Younger Cohort</i>	0.797	0.818	0.69	0.701
	<i>Older Cohort</i>	0.771	0.817	0.766	0.70
<b>General self-esteem</b>	<i>Younger Cohort</i>	0.763	0.731	0.705	0.697
	<i>Older Cohort</i>	0.722	0.66	0.72	0.695
<b>Parent relations</b>	<i>Younger Cohort</i>	0.767	0.811	0.79	0.806
	<i>Older Cohort</i>	0.743	0.806	0.818	0.807
<b>Peer relations</b>	<i>Younger Cohort</i>	0.785	0.815	0.741	0.723
	<i>Older Cohort</i>	0.777	0.787	0.807	0.742

## Appendix D: Variables included in new scales for Round 4

Scales	Items
<b>Self-efficacy scale</b>	Younger Cohort Round 4 (12 years old) and Older Cohort Round 4 (19 years old) <ol style="list-style-type: none"> <li>1. I can always manage to solve difficult problems if I try hard enough</li> <li>2. If someone opposes me, I can find the means and ways to get what I want</li> <li>3. It is easy for me to stick to my aims and accomplish my goals</li> <li>4. I am confident that I could deal efficiently with unexpected events</li> <li>5. Thanks to my resourcefulness, I know how to handle unforeseen situations</li> <li>6. I can solve most problems if I invest the necessary effort</li> <li>7. I can remain calm when facing difficulties because I can rely on my coping abilities</li> <li>8. When I am confronted with a problem, I can usually find several solutions</li> <li>9. If I am in trouble, I can usually think of a solution</li> <li>10. I can usually handle whatever comes my way</li> </ol>
<b>Self-esteem scale</b>	Younger Cohort Round 4 (12 years old), Younger Cohort sibling and Older Cohort Round 4 (19 years old) <ol style="list-style-type: none"> <li>1. I do lots of important things</li> <li>2. In general, I like being the way I am</li> <li>3. Overall, I have a lot to be proud of</li> <li>4. I can do things as well as most people</li> <li>5. Other people think I am a good person</li> <li>6. A lot of things about me are good</li> <li>7. I'm as good as most other people</li> <li>8. When I do something, I do it well</li> </ol>
<b>Parents relations scale</b>	Younger Cohort Round 4 (12 years old), Younger Cohort sibling and Older Cohort Round 4 (19 years old) <p><i>Only asked to children with one or both parents alive</i></p> <ol style="list-style-type: none"> <li>1. My parents understand me</li> <li>2. I like my parents</li> <li>3. My parents like me</li> <li>4. If I have children of my own, I want to bring them up like my parents raised me</li> <li>5. My parents and I spend a lot of time together</li> <li>6. My parents are easy to talk to</li> <li>7. I get along well with my parents</li> <li>8. My parents and I have a lot of fun together</li> </ol>
<b>Peer relations scale</b>	Younger Cohort Round 4 (12 years old), Younger Cohort sibling and Older Cohort Round 4 (19 years old) <ol style="list-style-type: none"> <li>1. I have lots of friends</li> <li>2. I make friends easily</li> <li>3. Other kids want me to be their friend</li> <li>4. I have more friends than most other kids</li> <li>5. I get along with other kids easily</li> <li>6. I am easy to like</li> <li>7. I am popular with kids of my own age</li> <li>8. Most other kids like me</li> </ol>

**Appendix E:**  
**Specified models for confirmatory factor analysis (Younger Cohort)**



## Appendix F: Fit statistics for confirmatory factor analysis (CFA)

**Table F.1.** *Fit statistics for CFA for the Older Cohort across countries*

	Model Threshold	$\chi^2$ p>0.05	$\chi^2/df$ <5	TLI >.90	CFI >.95	RMSEA <0.05	LO 90 0.02	HI 90 0.08
<b>Ethiopia</b>	<i>Self-efficacy</i>	p>0.05	4.053	.899	.923	.061	.051	.072
	<i>Self-esteem</i>	p>0.05	6.66	.837	.884	.083	.070	.097
	<i>Parent relations</i>	p>0.05	5.614	.887	.923	.075	.062	.089
	<i>Peer relations</i>	p>0.05	3.242	.956	.970	.052	.038	.067
	<i>Emotional stability</i>	p>0.05	2.94	.965	.982	.049	.021	.079
<b>India</b>	<i>Self-efficacy</i>	p>0.05	4.055	.926	.942	.062	.051	.072
	<i>Self-esteem</i>	p>0.05	1.547	.976	.983	.026	.000	.043
	<i>Parent relations</i>	p>0.05	3.727	.952	.967	.058	.044	.073
	<i>Peer relations</i>	p>0.05	4.104	.938	.958	.062	.048	.077
	<i>Emotional stability</i>	p>0.05	4.690	.924	.962	.068	.042	.097
<b>Peru</b>	<i>Self-efficacy</i>	p>0.05	2.950	.916	.937	.058	.045	.071
	<i>Self-esteem</i>	p>0.05	.759	1.012	1.00	.000	.000	.025
	<i>Parent relations</i>	p>0.05	3.558	.946	.965	.066	.049	.084
	<i>Peer relations</i>	p>0.05	4.481	.914	.939	.077	.061	.094
<b>Vietnam</b>	<i>Self-efficacy</i>	p>0.05	3.121	.913	.936	.050	.039	.061
	<i>Self-esteem</i>	p>0.05	2.695	.943	.959	.044	.030	.059
	<i>Parent relations</i>	p>0.05	2.006	.984	.989	.034	.018	.050
	<i>Peer relations</i>	p>0.05	2.665	.958	.971	.044	.030	.059
	<i>Emotional stability</i>	p>0.05	4.738	.932	.966	.066	.041	.094

**Table F.2.** *Fit statistics for CFA for the Younger Cohort across countries*

	Model Threshold	$\chi^2$ p<0.05	$\chi^2/df$ <5	TLI >.90	CFI >.95	RMSEA <0.05	LO 90 0.02	HI 90 0.08
<b>Ethiopia</b>	<i>Self-efficacy</i>	p<0.05	5.907	.936	.953	.053	.046	.060
	<i>Self-esteem</i>	p<0.05	5.683	.948	.963	.052	.043	.061
	<i>Parent relations</i>	p<0.05	1.794	.991	.994	.021	.008	.033
	<i>Peer relations</i>	p<0.05	2.869	.983	.988	.033	.023	.043
<b>India</b>	<i>Self-efficacy</i>	p<0.05	3.793	.966	.973	.043	.035	.051
	<i>Self-esteem</i>	p<0.05	3.572	.961	.972	.041	.031	.052
	<i>Parent relations</i>	p<0.05	3.501	.976	.984	.041	.030	.051
	<i>Peer relations</i>	p<0.05	3.828	.974	.983	.043	.033	.054
<b>Peru</b>	<i>Self-efficacy</i>	p<0.05	3.163	.946	.961	.035	.027	.043
	<i>Self-esteem</i>	p<0.05	3.011	.966	.975	.034	.024	.044
	<i>Parent relations</i>	p<0.05	5.545	.960	.977	.051	.041	.061
	<i>Peer relations</i>	p<0.05	3.458	.968	.978	.037	.028	.047
<b>Vietnam</b>	<i>Self-efficacy</i>	p<0.05	4.249	.933	.951	.041	.035	.049
	<i>Self-esteem</i>	p<0.05	4.052	.951	.965	.040	.031	.049
	<i>Parent relations</i>	p<0.05	3.269	.982	.988	.035	.025	.044
	<i>Peer relations</i>	p<0.05	3.574	.965	.976	.037	.028	.047

**Table F.3.** *Fit statistics for CFA for the Younger Cohort sibling across countries*

	Model Threshold	$\chi^2$ $p < 0.05$	$\chi^2/df$ <5	TLI >.90	CFI >.95	RMSEA <0.05	LO 90 0.02	HI 90 0.08
<b>Ethiopia</b>	<i>Self-esteem</i>	$p < 0.05$	4.887	.946	.963	.053	.042	.064
	<i>Parent relations</i>	$p < 0.05$	3.268	.971	.980	.040	.029	.052
	<i>Peer relations</i>	$p < 0.05$	3.434	.974	.915	.042	.031	.053
<b>India</b>	<i>Self-esteem</i>	$p < 0.05$	3.254	.963	.974	.041	.030	.053
	<i>Parent relations</i>	$p < 0.05$	2.180	.988	.992	.030	.017	.042
	<i>Peer relations</i>	$p < 0.05$	3.615	.964	.976	.045	.034	.056
<b>Peru</b>	<i>Self-esteem</i>	$p < 0.05$	1.764	.965	.975	.033	.013	.051
	<i>Parent relations</i>	$p < 0.05$	2.703	.966	.975	.050	.034	.066
	<i>Peer relations</i>	$p < 0.05$	1.727	.970	.979	.032	.012	.050
<b>Vietnam</b>	<i>Self-esteem</i>	$p < 0.05$	2.387	.949	.963	.038	.024	.052
	<i>Parent relations</i>	$p < 0.05$	1.667	.989	.992	.027	.007	.042
	<i>Peer relations</i>	$p < 0.05$	2.882	.949	.965	.045	.031	.059

**Appendix G:**  
**Fit statistics for multi-group confirmatory factor analysis (MGCFA)**

**Table G.1.** *Fit statistics for MGCFA for the Younger Cohort across countries*

	Model Threshold	$\chi^2$	$\chi^2/df$ <5	TLI >.90	CFI >.95	RMSEA <0.05	LO 90 0.02	HI 90 0.08	$\chi^2$ diff
<b>Self- efficacy</b>	<i>Unconstrained</i>	572.26 (134)	4.271	.947	.961	.022	.020	.024	-
	<i>Fully constrained</i>	794.09 (161)	4.932	.936	.843	.024	.022	.025	0.000
	<i>Fully constrained (adj.)</i>	650.98 (155)	4.200	.948	.955	.021	.020	.023	0.000
	<i>Partially constrained</i>	583.819 (140)	4.170	.949	.960	.021	.020	.023	0.073
<b>Self- esteem</b>	<i>Unconstrained</i>	326.372 (80)	4.080	.956	.968	.021	.019	.023	-
	<i>Fully constrained</i>	381.215 (101)	3.877	.960	.964	.020	.018	.022	.000
	<i>Partially constrained</i>	333.757 (83)	4.021	.957	.968	.021	.019	.023	.061
<b>Parent relations</b>	<i>Unconstrained</i>	249.634 (72)	3.467	.977	.985	.019	.016	.021	-
	<i>Fully constrained</i>	320.600 (93)	3.447	.978	.981	.019	.017	.021	0.000
	<i>Partially constrained</i>	264.611 (81)	3.267	.979	.985	.018	.016	.021	0.092
<b>Peer relations</b>	<i>Unconstrained</i>	261.242 (76)	3.437	.974	.982	.019	.016	.021	-
	<i>Fully constrained</i>	332.737 (97)	3.430	.476	.527	.084	.082	.086	0.000
	<i>Partially constrained</i>	267.761 (79)	3.389	.974	.982	.019	.016	.021	0.089

Note: The null hypothesis of the chi-square difference test is that there is no difference between groups.

**Table G.2.** *Fit statistics for MGCFA for the Older Cohort across countries*

	Model	$\chi^2$	$\chi^2/df$		TLI	CFI	RMSEA	LO 90	HI 90	$\chi^2$ diff
	Threshold		<5	>.90						
<b>Self-efficacy</b>	<i>Unconstrained</i>	483.041 (136)	3.552	.914	.935	.029	.026	.032	-	
	<i>Fully constrained</i>	826.12 (163)	3.853	.904	.913	.031	.028	.033	.000	
	<i>Fully constrained (adj.)</i>	570.062 (161)	3.541	.915	.924	.029	.026	.031	.000	
<b>Self-esteem</b>	<i>Unconstrained</i>	233.13 (96)	2.914	.929	.949	.025	.021	.029	-	
	<i>Fully constrained</i>	320.307 (101)	3.171	.920	.928	.027	.023	.030	.000	
	<i>Fully constrained (adj.)</i>	317.789 (100)	3.178	.919	.928	.027	.023	.030	.000	
<b>Parent relations</b>	<i>Unconstrained</i>	279.656 (75)	3.729	.947	.965	.030	.026	.034	-	
	<i>Fully constrained</i>	383.139 (96)	3.991	.942	.951	.031	.028	.035	.000	
	<i>Fully constrained (adj.)</i>	293.807 (93)	3.159	.958	.965	.027	.023	.030	.000	
<b>Peer relations</b>	<i>Unconstrained</i>	279.845 (77)	3.634	.942	.960	.029	.026	.033	-	
	<i>Fully constrained</i>	333.027 (98)	3.398	.947	.954	.028	.025	.031	.000	
<b>Emotional stability</b>	<i>Unconstrained</i>	61.866 (15)	4.124	.941	.971	.036	.027	.045	-	
	<i>Fully constrained</i>	78.835 (23)	3.428	.954	.965	.031	.024	.039	.030	

Note: The null hypothesis of the chi-square difference test is that there is no difference between groups.

**Table G.3.** *Fit statistics for MGCFA for the Younger Cohort sibling across countries*

	Model	$\chi^2$	$\chi^2/df$		TLI	CFI	RMSEA	LO 90	HI 90	$\chi^2$ diff
	Threshold		<5	>.90						
<b>Self-esteem</b>	<i>Unconstrained</i>	240.965 (79)	3.050	.955	.968	.022	.019	.025	-	
	<i>Fully constrained</i>	293.48 (100)	2.935	.958	.962	.021	.018	.024	.000	
	<i>Partially constrained</i>	270.699 (99)	2.734	.962	.966	.020	.017	.023	.000	
<b>Parent relations</b>	<i>Unconstrained</i>	189.250 (77)	2.458	.979	.986	.018	.015	.022	-	
	<i>Fully constrained</i>	267.940 (98)	2.734	.975	.978	.020	.017	.023	.000	
<b>Peer relations</b>	<i>Unconstrained</i>	279.593 (78)	3.585	.955	.969	.024	.021	.027	-	
	<i>Fully constrained</i>	374.961 (99)	3.787	.951	.957	.025	.023	.028	.000	

Note: The null hypothesis of the chi-square difference test is that there is no difference between groups.

## Appendix H: Results for partially constrained models

**Table H.1.** *Partially constrained items for self-efficacy, Younger Cohort*

Item	Item 01	Item 05	Item 08	Item 11	Item 15	Item 18	Item 22	Item 26	Item 28	Item 32
$\chi^2$ difference	(.003)	(.460)	(.001)	(.000)	(.024)	(.000)	(.034)	(.000)	(.000)	(.034)
Invariant	N	Y	N	N	N	N	Y	N	N	N

**Table H.2.** *Partially constrained items for self-esteem, Younger Cohort*

Item	Item 04	Item 06	Item 14	Item 17	Item 23	Item 27	Item 30	Item 33
$\chi^2$ difference	(.267)	(.082)	(.115)	(.001)	(.001)	(.077)	(.107)	(.107)
Invariant	Y	N	Y	N	N	N	N	N



**Table H.3.** *Partially constrained items for parent relations, Younger Cohort*

Item	Item 03	Item 07	Item 10	Item 13	Item 19	Item 21	Item 25	Item 29
$\chi^2$ difference	(0.034)	(0.000)	(0.017)	(0.001)	(.288)	(.153)	(.171)	(.171)
Invariant	N	N	N	N	Y	Y	Y	Y

**Table H.4.** *Partially constrained items for peer relations, Younger Cohort*

Item	Item 02	Item 09	Item 12	Item 16	Item 20	Item 24	Item 31	Item 34
$\chi^2$ difference	(.005)	(.000)	(.036)	(.001)	(.000)	(.000)	(.089)	(.089)
Invariant	N	N	N	N	N	N	Y	Y

## Appendix I: Variables from Round 3 included in retained scales for Round 4

Scales	Items
<b>Agency scale</b>	<p>Younger Cohort Round 4 (12 years old), Younger Cohort sibling and Older Cohort Round 4 (19 years old)</p> <ol style="list-style-type: none"> <li>1. If I try hard, I can improve my situation in life</li> <li>2. Other people in my family make all the decisions about how I spend my time (recoded to positive)</li> <li>3. I like to make plans for my future studies and work</li> <li>4. If I study hard at school I will be rewarded by a better job in future</li> <li>5. I have no choice about the work I do – I must do this sort of work (recoded to positive)</li> </ol>
<b>Pride and self-esteem scale</b>	<p>Younger Cohort Round 4 (12 years old) and Younger Cohort sibling</p> <ol style="list-style-type: none"> <li>1. I am proud of my shoes or of having shoes</li> <li>2. I feel my clothing is right for all occasions</li> <li>3. I am proud of my clothes</li> <li>4. I am never embarrassed because I do not have the right books, pencils and other equipment for school</li> <li>5. I am proud that I have the correct uniform</li> <li>6. I am proud of the work I have to do</li> </ol> <p>Older Cohort Round 4 (19 years old)</p> <ol style="list-style-type: none"> <li>1. I am proud of my shoes or of having shoes</li> <li>2. I feel my clothing is right for all occasions</li> <li>3. I am proud of my clothes</li> <li>4. I am proud of the work I have to do</li> </ol>

# Psychosocial Scales in the Young Lives Round 4 Survey: Selection, Adaptation and Validation

The importance of positive psychosocial functioning for children's current well-being and for their future outcomes is recognised across a range of disciplines, yet research on the psychosocial development of individuals in developing countries is in its infancy.

Young Lives has contributed to closing this knowledge gap through extensive data collection on a range of measures of psychosocial skill domains, administered in schools and households in developing countries, where large-scale survey data that include measures of psychosocial constructs are rare.

This technical note describes the procedure involved in the selection, adaptation and administration of the scales administered in the Young Lives Round 4 survey in 2013-14 to capture the development of psychosocial skills of 12-year-old children and young adults age 19 across all four study countries (Ethiopia, India (the states of Andhra Pradesh and Telangana), Peru and Vietnam).

The note can act as a guide for researchers interested in using Young Lives scales with the confidence that they have been tested thoroughly, and for other studies interested in filling the current knowledge gap, whether using the same or a different set of items.



An International Study of Childhood Poverty

## About Young Lives

Young Lives is an international study of childhood poverty, involving 12,000 children in four countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the four study countries: Ethiopia, India, Peru and Vietnam.

Through researching different aspects of children's lives, we seek to improve policies and programmes for children.

---

## Young Lives Partners

Young Lives is coordinated by a small team based at the University of Oxford, led by Professor Jo Boyden.

- *Ethiopian Development Research Institute, Ethiopia*
- *Pankhurst Development Research and Consulting plc, Ethiopia*
- *Centre for Economic and Social Studies, Hyderabad, India*
- *Sri Padmavathi Mahila Visvavidyalayam (Women's University), Andhra Pradesh, India*
- *Grupo de Análisis para el Desarrollo (GRADE), Peru*
- *Instituto de Investigación Nutricional (IIN), Peru*
- *Centre for Analysis and Forecasting, Vietnamese Academy of Social Sciences, Vietnam*
- *General Statistics Office, Vietnam*
- *Oxford Department of International Development, University of Oxford, UK*

---

## Contact:

### Young Lives

Oxford Department of  
International Development,  
University of Oxford,  
Mansfield Road,  
Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751

Email: [younglives@younglives.org.uk](mailto:younglives@younglives.org.uk)

Website: [www.younglives.org.uk](http://www.younglives.org.uk)