

Psychometric Characteristics of Cognitive Development and Achievement Instruments in Round 3 of Young Lives

Santiago Cueto and Juan León

Contents

Summary	ii
The authors	ii
Acknowledgements	ii
1. Introduction	1
2. Measuring cognitive abilities in Round 3 of Young Lives	4
2.1 Pilot tests for Round 3	4
2.2 Tests used and definitions	5
2.3 Procedures for administering the tests	9
3. Data analysis and methods	9
3.1 Sample	9
3.2 Data collection	11
3.3 Data-cleaning process	12
4. Results	14
4.1 Psychometric characteristics of the tests used in Round 3	14
4.2 Distribution of raw and Rasch scores	15
4.3 Reliability indexes and standard error of measurement according to CTT and IRT among cognitive tests	17
4.4 Correlation among cognitive tests	20
4.5 Correlation between cognitive tests and socio-demographic and educational variables	21
4.6 Correlation among cognitive measures in Round 2 and Round 3	30
4.7 Comparison across cohorts over time (Younger Cohort 2009 and Older Cohort 2002)	31
5. Conclusions and final considerations	33
References	36
Appendix 1. Additional characteristics and sources of items used in the mathematics tests	38
Appendix 2. Pearson correlation between the Rasch scores using the method in Round 2 and the method used in Round 3	39
Appendix 3. Item statistics for tests administered in Round 3	40
Appendix 4. Raw and Rasch scores distributions	89

Summary

This technical note gives details of the reliability and validity of the tests administered in Round 3 of Young Lives. The tests administered for the Younger Cohort children (aged 7 to 8) were the Peabody Picture of Vocabulary Test (PPVT), Early Grade Reading Assessment (EGRA), and a mathematics test. For the Older Cohort children (aged 14 to 15) the tests administered were the PPVT, sentence comprehension (Cloze), and mathematics. This document gives details of the procedures conducted to clean and build the scores (raw and Rasch) for each of the tests. We give details of the Rasch modelling performed to generate standardised scores and to analyse for possible bias (by gender and language used in the test) in order to estimate and correct for differential item functioning (DIF). The correlation between raw and Rasch scores is estimated, together with demographic variables, in order to check for the contextual validity of these scores for each age group and country. The raw and corrected scores have been uploaded in the Young Lives dataset, which is publicly archived in the UK Public Data Archive. They are also published as a separate Appendix to accompany this document.

The authors

Santiago Cueto is Senior Researcher at GRADE (Group for the Analysis of Development) in Lima and Country Coordinator for Young Lives in Peru. Juan León is Associate Researcher at GRADE in Lima.

Acknowledgements

The authors would like to thank colleagues in the Young Lives team: Amare Asgedom in Ethiopia, K. Mayuri in India, Andrea Baertl, Gabriela Guerrero and Claudia Sugimaru in Peru, and Vu Thi Son in Vietnam for helping to conduct the pilot studies in each of their countries, and for providing valuable input for the development of the tests and procedures used in Round 3. Professor Walter Secada from the University of Miami provided valuable input for the development of the mathematics tests. Ismael Muñoz from GRADE provided valuable help in the initial stages of analysis for this report. The authors would also like to thank Caine Rolleston from the Young Lives team at the University of Oxford for his input on the analysis that needed to be performed.

1. Introduction

Young Lives is a long-term international study focused on childhood poverty which integrates cutting-edge research with local, national and international policy-analysis engagement. The project seeks (a) to improve understanding of the causes and consequences of childhood poverty and to examine how policies affect children’s well-being; and (b) to inform the development and implementation of future policies and practices that will reduce childhood poverty. The Young Lives project, coordinated by the University of Oxford, is tracking the development of 12,000 children in Ethiopia, India (in the state of Andhra Pradesh only), Peru and Vietnam through quantitative and qualitative research over a 15-year period. Young Lives has been following two cohorts of children per country since the beginning of the project: 2,000 children in each country who were born in 2001–02 (Younger Cohort); and 1,000 in each country who were born in 1994–95 (Older Cohort).

We administered three rounds of surveys to both cohorts. Table 1 gives the years in which these were administered, plus the scheduled times for future surveys, and the approximate ages of the children at the time.

Table 1. *Dates of household and child surveys administered (or planned)*

	Year	Younger Cohort	Older Cohort
Round 1 survey	2002	6 to 18 months	7 to 8 years
Round 2 survey	2006–7	4 to 5 years	11 to 12 years
Round 3 survey	2009	7 to 8 years	14 to 15 years
Round 4 survey	2013	11 to 12 years	18 to 19 years
Round 5 survey	2016	14 to 15 years	21 to 22 years

Among other areas explored through Young Lives surveys are children’s cognitive abilities and achievement, with a heavy emphasis on those that could be linked to formal education in schools (see the description of the instruments below). In other words, while we acknowledge that specific skills are very relevant for children in specific social and physical environments, these are not measured by Young Lives. Instead we focus on the level at which children have acquired knowledge and educational and cognitive skills in receptive language (verbal) and mathematics, which are commonly included in international evaluations of schooling (TIMSS, Trends in International Mathematics and Science Study) and skills for life (PISA, the Programme for International Student Assessment).

Measurements used in each round of the Young Lives study are presented in Table 2. Round 1 did not include any measures of children’s development for the Younger Cohort; but it did include three questions on achievement for the older children (one in mathematics, one in reading, and one in writing), and these questions were administered again in Round 2 to the Older Cohort and in Round 3 to the Younger Cohort, when they were about the same age as children from the Older Cohort in Round 1. Round 1 also included the administration of the Raven Progressive Matrices to the children in the Older Cohort (Raven et al. 2003). The Raven Matrices are a non-verbal, multiple-choice measure of reasoning. We used the coloured version, especially designed for children. This test was not used in subsequent rounds, given reports from the field that children found the procedures hard to understand and that it took a long time to be administered in some cases (especially in rural communities).

In preparation for Round 2 we conducted pilot tests of several cognitive development and achievement instruments in each country. As a result, the international team decided to administer the following range of tests: the Cognitive Development Assessment (CDA, of which we used only the Quantity sub-test); the Peabody Picture Vocabulary Test (PPVT¹) to both cohorts; the reading, writing and mathematics items from Round 1; and a mathematics achievement test to the Older Cohort. Given the importance of establishing evidence on the validity of these tests as a measure of the knowledge and skills of children, we conducted some psychometric analysis and produced raw and corrected data bases (the latter eliminating those items that seemed biased by gender or language or had poor psychometric characteristics; see Cueto et al. 2009). Furthermore, the PPVT was adapted in each country to each of the major languages spoken in the country, and we conducted analyses by language within country, encouraging researchers not to compare scores for this specific test across languages or countries. The main use of the test results, in our opinion, should be with reference to children within the same language group within country, and when comparing associations of test results with other variables (i.e. level of schooling or parental education with test scores).

Given the above, and after further pilots in preparation for a new round of data collection, in Round 3 the international team decided (1) to drop the CDA for the Younger Cohort, as it would be too easy for most children, given that they would be in primary school by the time of testing; (2) to keep the PPVT and the three achievement items; and (3) to produce mathematics and reading tests for both cohorts. See Table 2. The description of each test is included in the next section. To assess intra-household variation, the PPVT was also administered to the siblings of Young Lives children in Ethiopia, Peru and Vietnam.

Table 2. *Measures of abilities and achievement administered in Young Lives*

		Cognitive	Reading	Mathematics
Round 1	YC	NA	NA	NA
	OC	Raven's Progressive Matrices for children	One item on reading and one on writing	One multiplication item
Round 2	YC	PPVT	NA	CDA
	OC	PPVT	One item on reading and one on writing	a. One multiplication item. b. Maths test.
Round 3	YC	PPVT (a younger sibling of the Young Lives child was also tested in Ethiopia, Peru and Vietnam)	a. One item on reading and one on writing. b. Early Grade Reading Assessment (EGRA).	a. One multiplication item. b. Maths test.
	OC	PPVT (a younger sibling of the Young Lives child was also tested in Ethiopia)	Cloze test of reading comprehension	Maths test

NA: Not administered.
See descriptions of the tests in the next section.

All tests were norm-referenced; in other words, we were interested in seeing how each child's scores were placed within the total sample that was tested, and not whether the child

1 Version 3 of the PPVT was used in Ethiopia, India, and Vietnam. Because this version is in English, adaptations of the test were developed by members of the research team in each of the countries. In Peru we used the available Spanish version of the PPVT-R (revised version of the first version of the PPVT), which was adapted also for use with Quechua-speaking subjects. All children were invited to respond to the PPVT in the language in which they felt most comfortable, thus allowing them to perform at their best.

reached a certain threshold or not. Among other characteristics, norm-referenced tests are expected to produce normal distributions, which was a challenge, given that for the most part the items were the same across countries and there were large differences in children's abilities within and between countries. (We present data on the distributions of scores below.)

The main aim of this technical document is to present evidence on the reliability and validity of the instruments used in Round 3 of Young Lives. 'Reliability' refers to the degree to which the items measure a construct that is stable across the test (e.g. the inter-item correlation should be positive and high in all cases). 'Validity', a more complex characteristic, refers to the level at which the inferences made about the children's knowledge and abilities, as measured by the tests, are supported by the test data. However, a variety of researchers are likely to use these data bases for specific topics, and we cannot establish that all associations are valid. Thus, in order to present evidence of validity, we show below that the scores obtained for children are of acceptable levels of reliability, do not present systematic bias by gender or language, and correlate as theory would predict with other constructs. With these methods we produced scores, using both Classical Test Theory (CTT) and Item Response Theory (IRT) procedures. In simple terms, CTT produces a score as the result of considering all the responses (correct, incorrect, and null) from each individual, treating all items equally. IRT models do not treat all items equally but estimate the level of difficulty of each at the same time that the ability of the individual in the measured construct is estimated; hence, IRT models enable an estimation of the probability of success for each test-taker for each item, thus producing a total score with the overall ability of the person. In our case we used the Rasch model, which is the one-parameter model of IRT (Baker 2004; Hambleton et al. 1991). Below we present indicators of reliability and validity of the tests for both CTT and Rasch models.

In general our definitions and analysis have been guided by the standards for educational and psychological testing proposed by the American Educational Research Association (AERA), the American Psychological Association (APA) and the National Council on Measurement in Education (NCME) (1999). The scores that we produced, as well as the original raw scores, are available in the Young Lives data sets that are publicly available in the UK Public Data Archive.² We do not present here further definitions of reliability and validity, or CTT and IRT, as these are available in the Technical Document produced for Round 2 and in the standards mentioned above, as well as other references (e.g. Crocker and Algina 2006).

The analysis and caveats presented below for the data bases and their uses are only suggestions for researchers to consider. We welcome any suggestions for further methods to estimate the reliability and validity of the scores used in Round 3, and would like to know about further procedures performed by researchers with these data bases, in order to create a reflection of the children's knowledge and skills that is as accurate as possible.

The paper is organised into five chapters. In the second chapter we describe the instruments used in Round 3 of Young Lives and how these were developed. In the third we present the procedures used for analysing the data. Chapter Four presents the results, and finally Chapter 5 summarises the information and offers conclusions on the reliability and validity of the data. Appendix 3 and Appendix 4 (published in a separate document to accompany this Technical Note) present the raw and Rasch scores for each age group and country.

² See <http://www.younglives.org.uk/what-we-do/access-our-data>.

2. Measuring cognitive abilities in Round 3 of Young Lives

Below we present the steps that we took to develop the tests used in Round 3; and we describe the characteristics and scoring of the tests, and the procedures used to administer the tests in each of the countries.

2.1 Pilot tests for Round 3

With the help of local teams, we conducted pilot tests of instruments in the four Young Lives countries between March and April 2009. The analysis essentially included two stages. In the first, for each new test a large number of items were sent to the countries, with a request for the formation of a panel to judge their quality. Each panel was expected to be composed of at least five people plus the pilot coordinator, and to be diverse: for instance it should include males and females, people from different regions, and people with experience of teaching or working with children of ages similar to those in Young Lives Round 3. It was hoped that such panels would be competent to assess the fairness³ of each test for use in their respective countries. 'Unfairness' was defined as inadequate for the population in the country (e.g. the item might be judged as too 'Western', or to be measuring a cultural trait or item of knowledge not relevant to the skill that the test was trying to measure). The panel were also asked to assess the tests for gender bias (i.e. no item should be biased against either boys or girls within the country); for urban–rural bias (i.e. no item should be biased against either the urban or rural population in the country); and finally for majority/minority-group bias (i.e. no item should be biased against a specific cultural group within the country). Furthermore, the panel was asked to review again the items that had been selected to be used locally from the PPVT. (In Round 2 we had asked similar panels in Ethiopia, India and Vietnam to select the non-biased items from forms A and B of the PPVT III and to construct their own, national version of this instrument, but retaining one of the main characteristics of the test, which is to present the child with items of increasing levels of difficulty). Results from Round 2 for the Older Cohort (Cueto et al. 2009) suggested that the items chosen were not of increasing levels of difficulty as expected. Hence too many children were reaching the final items. For Peru we used the standard version of the PPVT-R in Spanish, so no changes were necessary. Also, we asked for the tests to be adapted locally following these general guidelines:

1. All the languages spoken by children in either cohort should be identified (in other words, not the family language but the language preferred by the child to answer the tests; this is especially crucial for the PPVT and reading-comprehension tests).
2. For each language, two translators who were proficient in English and had the target language as their mother-tongue should be identified and requested to translate independently all the texts and items to be administered to the children. (The instructions and general procedures of administration could be translated only into one, main, national language, so that they would be administered in a standard way.⁴)

³ For further details about test fairness, see Cueto et al. (2009).

⁴ In fact, not all tests were translated into local languages. This happened when there were only a few speakers of a given language, and it was left for the fieldworkers to translate the test during administration. Since this is likely to bias these administrations, for this report we analyse only the data for which there was a standard translation (and administration) for a test.

3. After each completing their translation, the two translators were asked to meet to compare their translations and agree on a best, final translation to be used in the pilot.
4. It was acceptable to change first names, mathematics symbols, and similar elements to make the items more familiar to the children in the mathematics and Cloze items, as long as the changes did not affect the degree of difficulty of the item.⁵

The second stage of the pilot involved administering the instruments to a sample of around 48 children in each country. The sample included around 24 children from each age group (Younger and Older Cohorts, as expected for Round 3) and from different environments and ethnic groups, as relevant for the large Young Lives survey. The instruments were administered, data bases were formed and analysed by the GRADE team, and a final set of instruments and procedures was suggested, based on the opinions of the panel and the psychometric characteristics of the tests (in particular we chose items that increased reliability and helped to build a normal distribution for the total scores). The final instruments were designed to meet three criteria:

1. Instruments should provide data that could be compared with similar constructs from previous rounds. This was done sometimes with the same instruments or items, and at other times with instruments measuring the same domains.
2. Instruments should provide data that capture a wide variability in abilities. This was a challenge, as the levels of ability within and across countries varied a lot; yet with essentially the same instruments we tried to avoid ceiling and floor effects (i.e. difficulties measuring the higher or lower end of children's abilities).
3. The expected length of total time for all tests administered to children should be less than one hour. This was to avoid children getting tired or distracted and hence not performing at their best.

From the pilot tests we selected the best items, but in some cases we had to construct new ones, for example to increase the length of some tests which had relatively poor reliability. Unfortunately we did not have a chance to carry out further pilots with the revised instruments, hence the need for the analysis below. More information on the pilot and the summary report from it is available from the authors.

2.2 Tests used and definitions

The final set of instruments that we used consisted of the following elements.

- A. Common achievement items across rounds.** As mentioned above, since Round 1 we have administered three achievement items to children. The first focused on reading, the second on writing, and the third on mathematics. The cohorts and dates when these items were administered were presented in Table 1 (above). The original reading item consisted of three letters ('T, A, H'), one word ('hat'), and one sentence ('The sun is hot'). For the original writing item children were asked to write 'I like dogs'. However, for both the reading and the writing tests, countries adapted these sentences to specific languages and cultural contexts, so that the sentences are not always the same – but it is assumed that the levels of difficulty are comparable. Finally for the original mathematics item children were asked to multiply 2 x 4.

⁵ The cognitive equivalence of the English version with the translated local version of each language was checked, in order to ensure that they were of the same level of difficulty.

B. PPVT. The Peabody Picture Vocabulary Test is a widely used test of receptive vocabulary. It was originally developed in English in 1959 and has been updated several times. In this study we used version III (204 items; Dunn et al. 1997) in Ethiopia, India and Vietnam; this was the version available for Round 2 and we wanted to repeat the same instrument in Round 3 (currently there is a version IV of the PPVT). In Spanish we used the PPVT-R (125 items) adapted for Latin America (Dunn et al. 1986). Several studies have found that the PPVT has a positive strong correlation with some commonly used intelligence measures, such as the Wechsler and the McCarthy Scales (e.g. Campbell et al. 2001; Gray et al. 1999; Campbell 1998). However, most of the studies have been carried out for English-speaking populations, and here we wanted to use a combination of the plates used in both versions of the PPVT (which consist of four pictures presented on a board for the child to choose the one that corresponds to the word read to him/her)⁶ to create local versions of the PPVT.

The test is individually administered, orally administered, untimed, and norm-referenced. The task of the test taker is to select the picture that best represents the meaning of a stimulus word presented orally by the examiner. All items in the test are not expected to be administered. Instead, the fieldworker has to administer enough items to establish a ceiling and a baseline. For the Spanish version of the PPVT-R, the basal is formed by the highest eight consecutive correct responses, and the ceiling is formed by the lowest eight consecutive responses containing six errors. In the case of the PPVT-III, the basal set rule is one error, or no errors, in a set of 12 items, and the ceiling set rule is eight or more errors in a set. Non-administered items below the baseline are automatically given a score of 1, given that they are expected to be easier, while items above the ceiling are given a score of 0, given that they are more difficult. The raw score is formed by all the items given a score of 1 (i.e. answered correctly or below the basal item).

As mentioned above, the PPVT-III was used to evaluate both cohorts in the four Young Lives countries in Round 2.⁷ In order to use the PPVT, copies for each country were bought directly from the publisher. Although the test provides standard (age-converted) scores, we did not use them, because we wanted to generate scores that had acceptable levels of reliability and validity for each language in each country. The results of Round 2 showed some items that had poor psychometric indicators (i.e. poor item-total correlations and/or bias by gender). These items were discarded from the data sets (i.e. for Round 2 we had original raw scores, corrected raw scores that eliminated poor items, and Rasch-converted scores of the better items). These data bases are available in the public archive.

Our results for Round 2 showed that some children in the Older Cohort reached or almost reached the most difficult items, which is something that should not have happened, given that the test is intended to measure vocabulary up to adulthood. This is probably due to the adaptation carried out in each country. For the adaptation we took advantage of the fact that the PPVT-III is available in two parallel forms,

6 See for example <http://www.nlsinfo.org/childya/nlsdocs/guide/assessments/PPVT.htm>

7 Enough copies of the PPVT were purchased for use in field work, and one of the authors of the test was contacted to obtain permission for adaptation of the test to different languages.

designated as Form III-A and Form III-B. For Round 2, as mentioned above, we asked for a team to be assembled to judge the items, translate them as necessary, and produce a local version of a vocabulary test that followed the general format of the PPVT (e.g. the items were expected to be of increasing difficulty). If an item from Form A, which is the one that teams started working with, did not appear relevant for the local context, or its translation significantly affected the difficulty of the item, it should have been replaced with an item from approximately the same position in Form B. Apparently, however, as suggested above, the level of difficulty of the test was reduced in the process. For example, in some languages a single word was not available as a translation, hence local teams decided to use several words as an equivalent. However, we still advocate the reliability of the results, given that the administration of the test was standardised for all children speaking the same language within a country (i.e. translation of the PPVT for most local languages was generated, and the language used by the child in answering the tests was registered).

Given the above, for Round 3 we asked local teams in Ethiopia, India and Vietnam to again assemble panels of experts to select items from Forms A and B, test them in a pilot, and construct their own version of a test, which did not need to be exactly the same as in Round 2 but would measure the same construct. The plates had to be the same for both cohorts. In Peru we continued with the Spanish version of the PPVT-R, and with the Quechua version developed in Round 2 for those children who preferred to answer in this language. These are the results analysed below.

C. EGRA. The Early Grade Reading Assessment is a test that was developed with the support of USAID. It is 'an oral assessment designed to measure the most basic foundation skills for literacy acquisition in the early grades: recognising letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension'.⁸This test is freely available.⁹ It is administered individually and has several sub-tests, of which we adapted three for Young Lives: (1) *familiar-word identification*: in this sub-test the child is presented with 60 words (randomly ordered) and asked to read them in order; the task is to read as many words as possible in 60 seconds; (2) *passage reading*: in this sub-test the child is presented with a small text and asked to read it, the fieldworker marks how many words were read by the child after 60 seconds, then the child is given more time to read the text silently, after which eight questions on reading comprehension are posed orally by the fieldworker; (3) *listening comprehension*: in this sub-test the fieldworker reads a short text to the child, after which the child is asked to respond to six questions posed orally by the fieldworker. The justification for these sub-tests is that on the one hand a certain level of automaticity is needed to read (demonstrated in the number of words read in a minute); and that on the other hand we wanted children to demonstrate that beyond automaticity they could comprehend the text from reading or listening to it. For the analysis below we used four scores: number of correct words read in a minute in familiar-word identification; number of correct words read in a minute in passage reading; number of correct responses to reading-comprehension items in passage reading; and number of correct responses to

⁸ Taken from <https://www.eddataglobal.org/documents/index.cfm?fuseaction=showdir&ruid=1&statusID=3> on 18 August 2011.

⁹ Available at <https://www.eddataglobal.org/documents/index.cfm?fuseaction=showdir&ruid=1&statusID=3>.

reading comprehension and listening comprehension. The words and texts were adapted, in many cases making them more demanding than the original EGRA tests. This is because the pilot test showed that we needed more difficult items if we wanted to assess the children's range of reading skills. After this, local teams adapted the EGRA to local languages. Finally, RTI International, whose researchers have done most of the technical work behind EGRA, has been carrying out an intensive research agenda with this instrument, translating and adapting it to many countries, and producing research reports of different kinds, which are available on the web page mentioned above.

D. Cloze. As suggested above, our aim was to measure reading-comprehension abilities with the Older Cohort, for which the EGRA would be too simple a test. We pilot-tested traditional alternatives, such as those used in international evaluation programmes of reading comprehension of school-aged children such as PISA, but the format of medium to long texts followed by a few questions was too time-consuming and did not result in enough variability of scores. Hence the team at GRADE developed a Cloze test, which is a type of measure commonly used to measure verbal skills. In general, Cloze tests are measures of reading-comprehension abilities in which the person is asked to read a sentence or a short paragraph that lacks one or more words. The person has to demonstrate knowledge of vocabulary and comprehension of the situation in order to write down a word that would complete the meaning of the sentence or paragraph. The score that we used was the correct number of responses provided by the child. We developed 24 items, of increasing difficulty, and a set of acceptable answers to each one; the correct codes were filled in not by the fieldworkers but by a set of trained enumerators when all surveys were collected, so as to standardise procedures of coding. The test was taken by the child at his/her own pace, but gently discontinued after 10 minutes. As with the other verbal tests, there was a translation and adaptation process for each relevant local language, and we registered the language in which the child preferred to answer the test.

E. Mathematics tests. While we administered a mathematics test to the Older Cohort in Round 2 and ideally we should have used the same instrument in Round 3, data from Round 2 suggested that there was a ceiling effect for many children. Hence using it again would not have provided enough variability. Thus we decided to use existing items from national and international testing programmes that had been published freely, and developed a few new ones based on existing measures that are commonly used to assess mathematics skills. For both cohorts the maths items were divided into two sections.

For the Younger Cohort the first section aimed to measure basic quantitative and number notions. It included nine items on counting, knowledge of numbers, number discrimination, and using basic operations (including the 2×4 item mentioned above). These questions were read by the fieldworker with the aid of cards, so that no interference would result from poor reading skills. The second section aimed to measure ability to perform basic mathematics operations with numbers. It included 20 items using numbers for addition, subtraction, multiplication and division. The items were ordered in increasing levels of difficulty (according to the pilot test), all with whole numbers; each child took the test at his or her own pace, but the test was gently discontinued after eight minutes.

For the Older Cohort we also had two sections. The first aimed at the same goals as part 2 of the YC test. It included 20 items dealing with addition, subtraction, multiplication, division and square roots, using both whole numbers and fractions, and children were also allowed eight minutes to complete this part. Some of these items were the same as those used with the Older Cohort and some were repeated from Round 2. (Appendix 1 provides the details of the common items across cohorts and rounds.) Section 2 contained 10 items on mathematics problem solving; all of them were released items that were publicly available from TIMMS and PISA (Appendix 1 gives information about the source of these items). These topics measured with the items were (1) data interpretation, (2) number problem solving, (3) measurement, and (4) basic knowledge of geometry. This was the only maths section that included both open responses and multiple-choice items. The time allowed for this section was 10 minutes. The total score for children in mathematics was obtained from adding the correct responses. Children were not allowed to use calculators or any other help for any of the maths tests.

All items administered to children and the format of responses and manuals for fieldworkers, are available on the Young Lives web page.¹⁰

2.3 Procedures for administering the tests

The procedures were aimed at standardising the administration across fieldworkers so that children's best performance could be elicited. A detailed manual was prepared for test administration and fieldworkers were trained in the use of the cognitive and achievement instruments. The tests were to be administered in a space at home that was as quiet as possible. Children were given instructions on what to do, with examples, so that it was clear that the task was understood (if not, the test was not administered). The fieldworkers had to have a stop watch so that the time of test administration could be registered and children stopped at specific times, depending on the tests. Other specific procedures have been mentioned above or are described in the manual.

3. Data analysis and methods

3.1 Sample

In Round 1, the sample size was around 2,000 children per country for the Younger Cohort and 1,000 children for the Older Cohort. Table 3 summarises the attrition patterns for each cohort and country. The attrition rate for both age groups is less than 5 per cent on average. Outes-Leon and Dercon (2008) found that the attrition rate is lower than other longitudinal studies, even in developed regions, and there is no evidence of attrition bias.

¹⁰ See <http://www.younglives.org.uk/what-we-do/access-our-data>.

Table 3. *Percentage of children with data available in each round (following 3 rounds of data collection)*

	Round 1 only	Rounds 1 and 3 only	Rounds 1 and 2 only	Rounds 1, 2 and 3 (analytical sample)
Younger Cohort				
Ethiopia	4	0	1	94
N=	(85)	(0)	(28)	(1,886)
India	3	0	0	97
N=	(60)	(1)	(0)	(1,950)
Peru	3	1	2	93
N=	(61)	(28)	(48)	(1,915)
Vietnam	0	2	0	99
N=	(0)	(30)	(0)	(1,970)
Total	3	1	1	96
N=	(206)	(59)	(76)	(7,721)
Older Cohort				
Ethiopia	0	-	3	97
N=	(0)		(26)	(974)
India	1	-	1	98
N=	(14)		(9)	(985)
Peru	0	-	5	95
N=	(0)		(36)	(678)
Vietnam	1	-	1	98
N=	(10)		(14)	(976)
Total	1	-	2	97
N=	(24)		(85)	(3,613)

Source: Young Lives study data baseline, first and second follow-up.

Table 4 shows the demographic characteristics of children across countries and age cohorts. In all countries fathers appear to have had more education than mothers.

Table 4. *Main background characteristics (age, gender, ethnicity, parents' education and socio-economic status) of the panel sample by cohort in Round 3*

		Age (years)	Male (%)	Household size	Father: complete secondary or more (%)	Mother: complete secondary or more (%)
Younger Cohort						
Ethiopia	Mean	8.2	52.7	6.2	12.8	5.6
	Standard deviation	(0.3)	(49.9)	(2.0)	(33.4)	(23.1)
	N =	1,881	1,884	1,884	1,799	1,865
India	Mean	8.1	53.1	5.5	14.8	6.0
	Standard deviation	(0.3)	(49.9)	(2.2)	(35.5)	(23.7)
	N =	1,917	1,929	1,929	1,926	1,925
Peru	Mean	8.1	50.3	5.4	48.4	36.9
	Standard deviation	(0.3)	(50.0)	(1.9)	(50.0)	(48.3)
	N =	1,903	1,905	1,905	1,820	1,868
Vietnam	Mean	8.2	51.2	4.6	19.0	13.5
	Standard deviation	(0.3)	(50.0)	(1.4)	(39.3)	(34.2)
	N =	1,949	1,961	1,956	1,904	1,936
Older Cohort						
Ethiopia	Mean	15.2	51.3	6.4	8.9	4.6
	Standard deviation	(0.3)	(50.0)	(2.1)	(28.5)	(20.9)
	N =	961	973	973	908	958
India	Mean	15.2	49.1	5.1	11.1	4.0
	Standard deviation	(0.4)	(50.0)	(1.9)	(31.5)	(19.7)
	N =	976	976	976	971	970
Peru	Mean	15.2	53.2	5.4	48.6	33.2
	Standard deviation	(0.4)	(49.9)	(1.9)	(50.0)	(47.1)
	N =	671	671	671	632	662
Vietnam	Mean	15.3	49.5	4.5	19.6	13.5
	Standard deviation	(0.3)	(50.0)	(1.4)	(39.7)	(34.2)
	N =	964	972	962	937	963

Source: Young Lives study data baseline, first and second follow-up.

3.2 Data collection

The third round of data collection occurred between 2009 and the beginning of 2010, depending on the country. As in previous rounds, children were evaluated individually at home. As in Round 2, the time taken for administration for each cognitive test was registered by the fieldworkers. Table 5 presents information about this aspect for each cognitive test by age cohort. Peru shows the lowest time for the PPVT test, with 10 minutes on average. This is probably due to the PPVT in Peru having 125 items, compared with 204 in the other countries.

Table 5. Administration time in minutes by test and cohort

		Time of administration by test				
		PPVT	EGRA / Cloze	Maths (part 1)	Maths (part 2)	PPVT (siblings)
Younger Cohort						
Ethiopia	Mean	25.9	14.2	6.7	7.5	20.4
	Standard deviation	(11.2)	(10.5)	(3.9)	(1.8)	(10.8)
	N =	1,800	1,827	1,685	1,663	1,470
India	Mean	24.6	16.8	6.7	7.9	NA
	Standard deviation	(9.6)	(9.9)	(2.9)	(0.6)	
	N =	1,808	1,897	1,647	1,882	
Peru	Mean	10.3	11.9	3.4	8.0	NR
	Standard deviation	(4.3)	(5.8)	(1.6)	(0.8)	
	N =	1,818	1,852	1,861	1,848	
Vietnam	Mean	14.4	11.7	2.3	7.9	13.5
	Standard deviation	(7.7)	(5.6)	(1.4)	(1.0)	(6.1)
	N =	1,815	1,924	1,899	1,833	998
Older Cohort						
Ethiopia	Mean	28.1	10.0	8.0	9.9	27.5
	Standard deviation	(11.9)	(0.5)	(0.3)	(0.7)	(11.8)
	N =	908	884	917	892	701
India	Mean	26.5	9.7	7.9	9.6	NA
	Standard deviation	(9.9)	(1.1)	(0.8)	(1.4)	
	N =	921	858	896	883	
Peru	Mean	9.5	9.5	7.9	9.7	NA
	Standard deviation	(3.8)	(1.3)	(0.8)	(1.2)	
	N =	642	646	657	648	
Vietnam	Mean	15.5	8.8	7.7	9.5	NA
	Standard deviation	(6.8)	(1.9)	(1.2)	(1.4)	
	N =	933	927	940	901	

Note: The Cloze and Maths tests have a time limit, as indicated in the previous section. However, we considered times up to a limit of 14 minutes; any time above that limit was coded as missing.

Source: Young Lives study data baseline, first and second follow-up.

NR: not registered; NA: not administered

3.3 Data-cleaning process

The data sets provided by the four country teams were reviewed and cleaned before estimating the reliability and validity of the four achievement instruments. What follows is a description of the general steps implemented during the data-cleaning process.

- **Identification variable (ID) cleaning.** In Young Lives, data about children and their homes and families appear in all data files. Checking that the ID of children was correctly introduced in the different data sets was crucial, since this is the main variable used to merge the data sets.
- **Children's age validation.** In the context of an achievement assessment, the correct calculation of children's ages is very important, particularly in tests like the PPVT, where the child's chronological age determines his or her starting point in the test. For that reason, the age of the children was validated by contrasting each child's recorded age with his or her date of birth and date of administration.

- **Solving data inconsistencies.** Cross-tabulation of variables was undertaken to check for consistency among responses. All such inconsistencies that were detected were flagged and the national research-team leader was asked to correct the data. Those cases that could not be corrected, and cases where data made no sense, were coded as missing.

The cleaning process of the PPVT data set requires further explanation. The PPVT was the most difficult instrument to clean, because of the rules for test administration. Unlike the other tests, in the PPVT not all the items are administered to a given child, but only those within his or her critical range, i.e. those items assumed to be extremely easy or extremely hard for the individual were not administered. This requires the examiner to select the appropriate start item (according to the child's chronological age) and correctly establish the Basal Item Set and Ceiling Item Set for the individual. All the observations in the data set were reviewed to verify that Basal and Ceiling Sets were correctly established for each child. When any of these conditions was not met, we recalculated the basal and ceiling item; when it was not possible to recalculate, we coded the score for that child as missing. In some cases children were administered more PPVT items than required, i.e. the test was not discontinued when the children reached the Ceiling Item Set. This usually happened because the examiner did not mark an incorrect answer as an error and therefore failed to recognise that the child had already made enough errors to reach a ceiling. In such cases the additional items were not considered for score calculations; they were coded as missing and the raw score was recalculated, in order to be fair to other children who were administered the test with standard rules.

Also for the PPVT, in Round 2 we assumed that any item below the basal item would be answered correctly (as it was supposed to be of lower difficulty), and thus it was coded with 1 (correct). Any item above the ceiling was assumed to be too difficult for the child and thus coded 0 (incorrect). This procedure imputes the answer for each child according to the nature of the test. The relative positions of the items, however, were based on an adaptation done by local experts, which could have resulted in levels of difficulty different from what would be expected.

Given the above, in Round 3 we decided to test different methods to impute scores for the items below the basal item and above the ceiling item. A first approach employed was to estimate the children's ability by using only the set of common items that were answered by the entire sample; once this was estimated, we estimated the probability of correctly answering the items below the basal item and above the ceiling item. Once scores were imputed for those items, we ran the analysis using the whole set of items. A second approach was not to impute the answer of the items below the basal item or above the ceiling item; instead, we left the Rasch model to impute those values, using the information available for each of the items and the estimated children's ability (based on those items that were answered). We calculated the Rasch scores using the method in Round 2 and the two methods previously mentioned; the correlations were above 0.90, suggesting that all three methods provided very similar relative positions. For the analysis presented below, we decided to use the Rasch-generated scores.

Table 6 shows the analytical sample for each of the cognitive tests by language used in the administration. The analytical sample is the sample used for the psychometric analysis of each test. This sample includes those children with complete information in each cognitive measure, as well as excluding those children with inadequate scores or test administration.

For the siblings data, one additional filter was used: we deleted from the analysis any sibling who did not meet the inclusion criteria stated in the field-work manuals.¹¹

Table 6. *Analytical sample by test, language and cohort*

	Younger Cohort				Older Cohort			
	PPVT	EGRA ¹	Maths ¹	PPVT-siblings	PPVT	Cloze	Maths ¹	PPVT-siblings
Ethiopia								
Amharic	778	778	778	255	461	461	460	288
Tigrigna	382	381	380	232	200	200	200	162
Oromiffa	352	348	353	195	171	171	172	133
Other	369	367	345	221	139	135	134	132
Total	1,881	1,874	1,856	903	971	967	966	715
India								
Telugu	1,604	1,468	1,494	-	895	859	863	-
English	263	407	380	-	59	81	79	-
Other	51	42	43	-	21	22	22	-
Total	1,918	1,917	1,917	-	975	962	964	-
Peru								
Spanish	1,751	1,743	1,777	459	657	660	659	-
Quechua	59	53	43	45	1	1	1	-
Other	79	73	57	36	8	4	3	-
Total	1,889	1,869	1,877	540	666	665	663	-
Vietnam								
Viet	1,924	1,920	1,891	401	960	948	936	-
H'roi	13	15	0	0	4	1	0	-
Other	14	12	14	7	2	2	2	-
Total	1,951	1,947	1,905	408	966	951	938	-

¹ The effective sample size for this test was taken from the first sub-test administered. For the EGRA test, the effective sample is from the reading comprehension sub-tests; while for the maths test, the effective sample was taken from the first sub-test administered in each cohort.

Source: Young Lives study data baseline, first and second follow-up.

4. Results

4.1 Psychometric characteristics of the tests used in Round 3

The psychometric analysis reported in this section was done separately for each age cohort within country. Following both classical test theory (CTT) and item response theory (IRT), we estimated the psychometric properties of each of the items used in the third round. In addition, we identified the items with Differential Item Functioning (DIF) across subgroups of examinees based on gender and language, adjusting for the total ability.

¹¹ The criterion was to evaluate a younger sibling of the index child, but one over the age of 4 years.

We used the same five criteria defined in the psychometric report of the second round to define the properties of items:

- A. Item–Test correlation lower than 0.10.
- B. In-fit out of the range 0.50 to 1.50.
- C. Out-fit out of the range 0.50 to 1.50.
- D. The difference by gender is significant at 5 per cent (gender DIF).
- E. The difference by language is significant at 5 per cent (language DIF).

In the second round, an item was deleted if it met criteria A, B, or D and the number of children who answered the item was 30 or higher; items with fewer than 30 children were not deleted, since the estimated statistics would not be stable, given the small sample size. For Round 3 an item was deleted if it met the A or B criteria. If an item met the D or E criteria, we split the item and generated one item for each subgroup (gender or language). This procedure allowed us to save more items (Pearson correlation between the scores using these methods was above 0.90 for all the countries and tests, see Appendix 2). Tables containing the item statistics for each of the cognitive tests can be found in Appendix 3. Therefore, the psychometric analysis below uses the final configuration of each of the tests after removing items with poor fit and after correcting by gender and language bias.

4.2 Distribution of raw and Rasch scores

As mentioned above, our aim was to generate a normal distribution of scores for each test in each country. Two characteristics generally used to describe a distribution are *skewness* (lack of symmetry) and *kurtosis* (tallness or flatness). In a normal distribution skewness is zero and kurtosis is 3. When a distribution is not normal, there might be problems with either one of these indicators, or both. There are two types of skewness: positive skew, when the right tail is longer and therefore the distribution is concentrated on the left side (right-skewed); and negative skew, when the left tail is longer and therefore the distribution is concentrated on the right side (left-skewed).

In terms of its kurtosis, a normal distribution is mesokurtic, while a distribution which is not normal can be leptokurtic when it has a more acute peak around the mean (kurtosis values are greater than 3), or platykurtic when it has a smaller peak around the mean (kurtosis values are smaller than 3).¹²

Using the final scores of each test, we graphed the distribution of raw and Rasch scores for each of the tests by cohort and language. Frequently the Rasch transformation contributes to the normalisation of test curves. Tables 7 and 8 summarise the characteristics of the distributions of raw and Rasch scores for each instrument by country and language. The graphs of all the distributions are included in Appendix 4.

¹² For more details, see Tabachnick and Fidell (1996) and Gujarati (2003).

Table 7. *Characteristics of the distribution of the raw and Rasch scores by language: Younger Cohort*

	EGRA		Maths		PPVT	
	Raw	Rasch	Raw	Rasch	Raw	Rasch
Ethiopia						
Amharic	PS PK	PK	PS PK	ND	PS PK	ND
Oromiffa	PS PK	ND	PS	PK	C	C
Tigrigna	PS LK	LK	PS LK	ND	C	C
India						
Telugu	PS PK	PS LK	PS PK	ND	PS LK	ND
English	PS PK	ND	PK	ND	C	C
Peru						
Spanish	NS PK	NS LK	PK	NS LK	NS LK	NS LK
Quechua	C	C	PS	ND	C	C
Vietnam						
Tieng Viet Nam	NS LK	LK	NS PK	LK	PS LK	NS LK

C: Not considered for the analysis, given the small sample size.

ND: normal distributed; PS: positively skewed; NS: negatively skewed; LK: leptokurtic; PK: platykurtic.

Table 8. *Characteristics of the distribution of the raw and Rasch scores by language: Older Cohort*

	Cloze		Maths		PPVT	
	Raw	Rasch	Raw	Rasch	Raw	Rasch
Ethiopia						
Amharic	PS	PK	PS LK	ND	NS	LK
Oromiffa	PS	PK	PS	ND	C	c
Tigrigna	PS LK	PS PK	PS LK	ND	C	c
India						
Telugu	PK	NS	PS PK	ND	ND	ND
English	C	c	PK	ND	C	c
Peru						
Spanish	NS	NS LK	ND	NS LK	NS LK	NS LK
Vietnam						
Tieng Viet Nam	NS LK	NS LK	NS PK	PS	NS LK	NS LK

C: Not considered for the analysis, given the small sample size and the number of items for the test.

ND: normal distributed; PS: positively skewed; NS: negatively skewed; LK: leptokurtic; PK: platykurtic.

As shown, normal distributions were not found for many tests and countries (even after Rasch transformations). This is an important consideration for researchers conducting studies with these scores, given that many statistical models assume a normal distribution of scores.

4.3 Reliability indexes and standard error of measurement according to CTT and IRT among cognitive tests

We calculated the reliability indexes of each of the tests according to Classical Test Theory (CTT) and Item Response Theory (IRT),¹³ in order to have a measure of the stability of the items. In the case of CTT, a reliability coefficient of 0.60 or more was considered acceptable for research purposes. In the case of IRT, a person-reliability index above 0.50 is considered adequate, meaning that the test could discriminate between higher and lower achievers (Linacre 2008).

As in Round 2, we calculated separate reliability indexes for each language within a country, provided that the language had enough observations given the number of items,¹⁴ and provided that we had information that the test was administered in a standard way. In addition to providing the reliability indexes, we also report the average standard error of measurement (SEM) across children as an indicator of the overall precision of the scores calculated. There is an inverse relationship between the SEM and the reliability index: the higher the error, the lower the reliability. Tables 9 to 12 contain information about the reliability indexes and the SEM for each of the four tests used in Round 3.

Table 9 contains information about the reliability indexes for the EGRA. In general terms, the reliability coefficients of the EGRA are acceptable according to CTT (Cronbach's Alpha) and/or IRT (person-reliability index). Table 10 shows the reliability indexes for the Maths achievement tests in both age cohorts.

Table 9. *Reliability index and standard error of measurement for the EGRA (reading-comprehension items), using CTT and IRT by country and language*

Country	Language	Items Deleted	CTT Reliability ¹	SEM	IRT ² Reliability	SEM
Ethiopia	Amharic		0.83	1.36	0.79	6.4
	Oromiffa	Oral: Item 4. Reading: Item 8	0.83	1.32	0.77	6.7
	Tigrigna		0.64	1.06	0.56	8.7
India	Telugu		0.85	1.29	0.80	6.1
	English		0.83	1.46	0.71	7.6
Peru	Spanish		0.75	1.59	0.69	7.9
Vietnam	Tieng Viet Nam		0.69	1.46	0.59	9.1

1 Using Cronbach's Alpha coefficient.

2 The Rasch scores were fixed with a mean of 300 and standard deviation of 15.

13 The IRT analysis was conducted within each country for maths measures, and for vocabulary and reading-comprehension measures also within each local language.

14 The rule of thumb is four observations of children for each item in the test, in order to ensure stability in the analysis.

Table 10. *Reliability index and standard error of measurement for the Maths achievement, using CTT and IRT by country, cohort and language*

Cohort	Items Deleted	CTT Reliability ¹	SEM	IRT ² Reliability	SEM
Younger					
Ethiopia	Item 28	0.91	1.6	0.88	4.8
India		0.93	1.7	0.93	3.8
Peru		0.90	1.8	0.91	4.3
Vietnam	Item 4	0.90	1.8	0.88	4.9
Older					
Ethiopia	Items 24 and 26	0.68	2.7	0.75	7.3
India	Item 26	0.69	3.6	0.85	5.6
Peru		0.47	4.2	0.85	5.6
Vietnam		0.61	4.8	0.88	5.0

1 Using Cronbach's Alpha coefficient.

2 The Rasch scores were fixed with a mean of 300 and standard deviation of 15.

Table 11 contains the reliability indexes for the Cloze achievement tests. The indexes for both CTT and IRT are, in most cases, acceptable. The only case with a low reliability index, according to IRT, is that of the children who took the test in Tigrigna (Ethiopia); however, this could be associated with (1) the small sample size for this group (n=160) of children in Ethiopia, which could be causing the poor correlation among items; or (2) the test's level of difficulty for these children (i.e. the distribution of the scores was positively skewed and the majority of the children were clustered at the lower end of the score distribution, with small differences among children).

Table 11. *Reliability index and standard error of measurement for the Cloze achievement test, using CTT and IRT by country and language*

Country	Language	Items Deleted	CTT Reliability ¹	SEM	IRT ² Reliability	SEM
Ethiopia	Amharic		0.87	1.75	0.76	7.2
	Oromiffa	Item 24	0.84	1.68	0.69	8.2
	Tigrigna	Items 13, 19 and 22	0.79	1.15	0.39	11.5
India	Telugu		0.91	1.91	0.86	5.5
Peru	Spanish		0.89	1.88	0.85	5.5
Vietnam	Tieng Viet Nam		0.88	1.72	0.72	7.9

1 Using Cronbach's Alpha coefficient.

2 The Rasch scores were fixed with a mean of 300 and standard deviation of 15.

Finally, Table 12 displays the reliability indexes for the PPVT test administered to both cohorts. Using CTT or IRT, the reliability indexes are above adequate levels, showing that there is a high correlation among the items in this test. Also, we calculated the reliability indexes using the split-half method and the results are identical with those obtained using Cronbach's alpha. Regarding the IRT results, the reliability indexes calculated for the PPVT in both cohorts are very similar to those obtained in Round 2.

Table 12. *Reliability index and standard error of measurement for the PPVT (reading-comprehension items), using CTT and IRT by country, cohort and language*

Cohort	Language	No. of items	CTT Reliability ¹	SEM	IRT ² Reliability	SEM
Younger						
Ethiopia	Amharic	204	0.93	9.3	0.95	3.3
India	Telugu	204	0.98	4.3	0.88	4.9
Peru	Spanish	125	0.97	2.9	0.95	3.4
Vietnam	Tieng Viet Nam	204	0.98	4.0	0.92	4.0
Older						
Ethiopia	Amharic	204	0.98	5.0	0.94	3.4
India	Telugu	197	0.99	4.0	0.93	4.0
Peru	Spanish	125	0.97	3.0	0.91	4.4
Vietnam	Tieng Viet Nam	203	0.98	3.9	0.91	4.2

1 Using Cronbach's Alpha coefficient.

2 The Rasch scores were fixed with a mean of 300 and standard deviation of 15.

Box 1. *How dispersed are the cognitive scores across languages in reading comprehension?*

One issue that we wanted to explore is the dispersion of the reading comprehension scores across languages. For that purpose, we use the EGRA and Cloze Rasch scores. The mean scores were centred on 300 points and the standard deviation in 15 points for each country and each language. Thus, the purpose of the graphs presented below is not to compare the mean or standard deviations directly, but to compare the dispersion indexes, in order to see differences in the distribution of the reading scores across languages and countries.

The dispersion indexes used are (1) range: dispersion index indicating the spread of total scale or variable; the difference between the maximum and the minimum number in each scale; and (2) interquartile range (IQR): dispersion index calculated as the difference between the third and first quartile of the scale distribution. Thus the IQR is not affected by extreme values.

The table below shows the basic statistics for EGRA scores by country and language. In the case of the range, Peru and Ethiopia have the highest value or spread of scores. However, the IQR index shows that Ethiopia and India are the countries with the highest difference across quartiles. In Ethiopia and India, the difference between the third and first quartile is 19 or more points for all languages, while in Peru and Vietnam this difference is 18 points or less.

EGRA (reading comprehension) test scores descriptive statistics

	Mean	SD	Max.	Min.	Range	IQR
Ethiopia						
Amharic	300	15	341	261	80	24
Oromiffa	300	15	337	260	76	21
Tigrigna	300	15	357	259	98	20
India						
Telugu	300	15	343	261	82	19
English	300	15	340	268	73	24
Peru						
Spanish	300	15	337	243	94	18
Vietnam						
Tieng Viet Nam	300	15	333	231	102	17

Box 1. *How dispersed are the cognitive scores across languages in reading comprehension? continued*

The table below shows the score dispersion for the Cloze test. As with the EGRA, Peru and Vietnam have the greatest spread of data, with a range above 80 points. However, once outliers are taken into account, Ethiopia and Vietnam are the countries with the highest spread of data or scores. In Ethiopia, children who took the test in Tigrigna have the highest spread of data, with a difference of 27 points between the first and third quartile, followed by Amharic and Oromiffa children. Finally, Peru is the country with the smallest spread between quartiles.

Cloze test scores descriptive statistics

	Mean	SD	Max.	Min.	Range	IQR
Ethiopia						
Amharic	300	15	339	272	66	21
Oromiffa	300	15	341	274	67	19
Tigrigna	300	15	344	284	59	27
India						
Telugu	300	15	344	266	78	20
Peru						
Spanish	300	15	339	250	89	16
Vietnam						
Tieng Viet Nam	300	15	324	237	87	22

4.4 Correlation among cognitive tests

The correlations between raw and Rasch scores were calculated separately for each test, language and cohort. As shown in Table 13, all correlations are positive, statistically significant, and close to 1 as expected, given that Rasch scores are monotonically 1–1 with the raw scores; the only difference is the normalisation model used to calculate the scores.

Table 13. *Correlation between raw and Rasch cognitive measures by test, language and cohort*

	EGRA	Maths – YC	PPVT – YC	Cloze	Maths – OC	PPVT – OC
Ethiopia						
Amharic	0.98 ***	0.98 ***	0.93 ***	0.96 ***	0.96 ***	0.91 ***
Oromiffa	0.94 ***	0.98 ***	-	0.95 ***	0.96 ***	-
Tigrigna	0.96 ***	0.95 ***	-	0.96 ***	0.95 ***	-
India						
Telugu	0.97 ***	0.99 ***	0.90 ***	0.97 ***	0.96 ***	0.95 ***
English	0.97 ***	0.99 ***	-	-	0.99 ***	-
Peru						
Spanish	0.93 ***	0.99 ***	0.99 ***	0.98 ***	0.99 ***	0.98 ***
Quechua	-	0.98 ***	-	-	-	-
Vietnam						
Tieng Viet Nam	0.96 ***	0.98 ***	0.90 ***	0.96 ***	0.98 ***	0.93 ***
H'mong	-	-	-	-	-	-

***p<.001, **p<.01, *p<.05

Table 14 shows the correlations between the tests administered in each cohort (vocabulary, maths and reading comprehension). The correlations were calculated by language within each country. As shown, the correlation coefficients in all the cases are high, positive and statistically significant as would be expected, with Vietnam the country with the lowest correlations.

Table 14. *Correlation between original and corrected cognitive measures (raw and Rasch scores) by test, language and cohort*

	Younger Cohort			Older Cohort		
	PPVT – EGRA	PPVT – Maths	EGRA – Maths	PPVT – Cloze	PPVT – Maths	Cloze – Maths
Ethiopia						
Amharic	0.60 ***	0.69 ***	0.69 ***	0.61 ***	0.51 ***	0.61 ***
Oromiffa	-	-	0.66 ***	-	-	0.54 ***
Tigrigna	-	-	0.43 ***	-	-	0.56 ***
India						
Telugu	0.54 ***	0.57 ***	0.65 ***	0.64 ***	0.66 ***	0.67 ***
English	-	-	0.69 ***	-	-	-
Peru						
Spanish	0.57 ***	0.62 ***	0.53 ***	0.70 ***	0.58 ***	0.62 ***
Quechua	-	-	-	-	-	-
Vietnam						
Tieng Viet Nam	0.48 ***	0.52 ***	0.47 ***	0.44 ***	0.56 ***	0.55 ***
H'mong	-	-	-	-	-	-

Note: the correlations were calculated only for those children who took both tests in the same language. We excluded the cases where a child took one test in one language and the second one in another language.

***p<.001, **p<.01, *p<.05

One additional analysis performed in the previous psychometric report was to calculate the correlation between the original raw score and the corrected raw scores, which discounted several biased items. However, as we indicated previously, for Round 3 biased items were not deleted; instead we split the items which presented gender (verbal tests) or language (maths test) bias, generating new items, since the item difficulty was different between those groups. This method meant that only a few items were dropped from our analysis; and the original and the corrected raw scores are almost the same (see Tables 9 to 12).

4.5 Correlation between cognitive tests and socio-demographic and educational variables

Tables 15 and 16 show the correlation between each test (raw and Rasch scores) and children's socio-demographic characteristics by language within each country. As expected, given the characteristics of the tests, the wealth index is positively associated with most tests. Child's age is positively associated with some measures, which is also probably associated with differences in grade at school (see Table 19); this suggests that age in months would be a good control in statistical analysis using these tests. On the other hand, gender is correlated with just a few tests for the Younger Cohort and a few more for the Older Cohort.

Table 15. Correlation of cognitive measures with background variables (age, gender and socio-economic status) by language in Younger Cohort

	Raw scores			Rasch scores		
	Age	Male	Wealth index	Age	Male	Wealth index
EGRA (reading comprehension)						
Ethiopia						
Amharic	0.07	-0.06	0.51 ***	0.07	0.05	0.52 ***
Oromiffa	0.15 ***	0.06	0.34 ***	0.20 ***	0.06	0.34 ***
Tigrigna	-0.03	-0.07	0.26 ***	-0.04	-0.07	0.26 ***
India						
Telugu	0.13 ***	0.01	0.12 ***	0.12 ***	0.04	0.13 ***
English	0.12 *	-0.05	0.08	0.11 *	-0.03	0.07
Peru						
Spanish	0.10 ***	-0.01	0.41 ***	0.09 ***	-0.02	0.34 ***
Vietnam						
Tieng Viet Nam	0.19 ***	-0.05 *	0.32 ***	0.18 ***	-0.07 **	0.30 ***
Maths						
Ethiopia						
Amharic	0.13 ***	0.00	0.63 ***	0.12 ***	-0.01	0.63 ***
Oromiffa	0.21 ***	0.03	0.52 ***	0.21 ***	0.04	0.50 ***
Tigrigna	0.05	0.06	0.38 ***	0.06	0.06	0.37 ***
India						
Telugu	0.15 ***	0.00	0.15 ***	0.15 ***	-0.01	0.16 ***
English	0.21 ***	-0.01	0.14 **	0.19 ***	-0.02	0.14 **
Peru						
Spanish	0.20 ***	0.06 **	0.43 ***	0.20 ***	0.06 **	0.43 **
Quechua	0.25	0.05	-0.10	0.25	0.08	-0.09
Vietnam						
Tieng Viet Nam	0.33 ***	-0.01	0.39 ***	0.33 ***	-0.01	0.39 ***
PPVT						
Ethiopia						
Amharic	0.21 ***	0.00	0.62 ***	0.17 ***	0.00	0.62 ***
India						
Telugu	0.11 ***	0.11 ***	0.20 ***	0.12 ***	0.11 ***	0.22 ***
Peru						
Spanish	0.11 ***	0.05 *	0.54 ***	0.11 ***	0.05	0.55 ***
Vietnam						
Tieng Viet Nam	0.26 ***	0.02	0.36 ***	0.25 ***	0.00	0.44 ***

***p<.001, **p<.01, *p<.05

Table 16. *Correlation of cognitive measures with background variables (age, gender and socio-economic status) by language in Older Cohort*

	Raw scores			Rasch scores		
	Age	Male	Wealth index	Age	Male	Wealth index
Cloze						
Ethiopia						
Amharic	0.02	-0.05	0.49 ***	0.02	-0.01	0.48 ***
Oromiffa	0.02	0.15 *	0.21 **	0.04	0.14	0.19 **
Tigrigna	-0.03	0.11	0.18 *	-0.05	0.17 *	0.21 **
India						
Telugu	0.01	0.13 ***	0.23 ***	0.04	0.14 ***	0.20 ***
Peru						
Spanish	0.03	-0.01	0.47 ***	0.01	-0.04	0.46 ***
Vietnam						
Tieng Viet Nam	0.01	-0.13 ***	0.33 ***	0.00	-0.13 ***	0.30 ***
Maths						
Ethiopia						
Amharic	0.07	0.14 **	0.30 ***	0.07	0.14 **	0.31 ***
Oromiffa	0.04	0.20 **	0.35 ***	0.06	0.24 ***	0.35 ***
Tigrigna	-0.03	0.07	0.24 **	-0.03	0.11	0.22 **
India						
Telugu	-0.02	0.21 ***	0.30 ***	-0.01	0.21 ***	0.25 ***
English	0.13	-0.03	0.30 **	0.13	-0.06	0.29 **
Peru						
Spanish	0.04	0.01	0.39 ***	0.03	-0.01	0.38 ***
Vietnam						
Tieng Viet Nam	0.07 *	-0.13 ***	0.32 ***	0.07 *	-0.14 ***	0.31 ***
PPVT						
Ethiopia						
Amharic	0.02	0.08	0.59 ***	0.01	0.12 *	0.61 ***
India						
Telugu	-0.05	0.23 ***	0.31 ***	-0.03	0.23 ***	0.34 ***
Peru						
Spanish	0.08 *	0.08 *	0.48 ***	0.07	0.08 *	0.46 ***
Vietnam						
Tieng Viet Nam	0.11 ***	-0.02	0.44 ***	0.12 ***	-0.02	0.45 ***

***p<.001, **p<.01, *p<.05

The last association explored in this section is between the children's scores and parents' schooling. We expected that the association would be positive and statistically significant. Tables 17 and 18 show the correlation coefficient for the Younger and Older Cohorts respectively. Indeed, in most cases children with better-educated parents have higher cognitive skills in all countries and across languages. The exception is Ethiopia for the Oromiffa children from the Older Cohort, where the relationship between children's scores and parents' schooling is not statistically significant.

Table 17. *Correlation of cognitive measures with parents' educational variables (complete secondary or more) by language in Younger Cohort*

	Raw scores		Rasch scores	
	Father	Mother	Father	Mother
EGRA (reading comprehension)				
Ethiopia				
Amharic	0.28 ***	0.22 ***	0.28 ***	0.21 ***
Oromiffa	0.19 ***	0.10 *	0.19 ***	0.12 *
Tigrigna	0.29 ***	0.19 ***	0.30 ***	0.19 ***
India				
Telugu	0.16 ***	0.09 **	0.16 ***	0.07 **
English	0.24 ***	0.32 ***	0.24 ***	0.30 ***
Peru				
Spanish	0.32 ***	0.33 ***	0.27 ***	0.30 ***
Vietnam				
Tieng Viet Nam	0.20 ***	0.18 ***	0.22 ***	0.20 ***
Maths				
Ethiopia				
Amharic	0.37 ***	0.29 ***	0.36 ***	0.27 ***
Oromiffa	0.21 ***	0.12 *	0.20 ***	0.11 *
Tigrigna	0.35 ***	0.23 ***	0.31 ***	0.19 ***
India				
Telugu	0.21 ***	0.16 ***	0.20 ***	0.15 ***
English	0.25 ***	0.28 ***	0.25 ***	0.27 ***
Peru				
Spanish	0.33 ***	0.36 ***	0.32 ***	0.35 ***
Quechua	0.33 ***	C	0.29 ***	C
Vietnam				
Tieng Viet Nam	0.25 ***	0.22 ***	0.26 ***	0.25 ***
PPVT				
Ethiopia				
Amharic	0.32 ***	0.29 ***	0.33 ***	0.30 ***
India				
Telugu	0.21 ***	0.18 ***	0.23 ***	0.18 ***
Peru				
Spanish	0.40 ***	0.43 ***	0.40 ***	0.43 ***
Vietnam				
Tieng Viet Nam	0.28 ***	0.29 ***	0.29 ***	0.29 ***

C: no available information

***p<.001, **p<.01, *p<.05

Table 18. *Correlation of cognitive measures with parents' educational variables (complete secondary or more) by language in Older Cohort*

	Raw scores		Rasch scores	
	Father	Mother	Father	Mother
Cloze				
Ethiopia				
Amharic	0.32 ***	0.17 ***	0.29 ***	0.16 ***
Oromiffa	0.04	-0.05	0.01	-0.07
Tigrigna	0.17 *	0.20 *	0.19 *	0.16 *
India				
Telugu	0.19 ***	0.11 ***	0.15 ***	0.12 ***
Peru				
Spanish	0.40 ***	0.36 ***	0.38 ***	0.35 ***
Vietnam				
Tieng Viet Nam	0.21 ***	0.20 ***	0.21 ***	0.23 ***
Maths				
Ethiopia				
Amharic	0.22 ***	0.16 ***	0.22 ***	0.16 ***
Oromiffa	0.01	0.02	0.00	0.03
Tigrigna	0.29 ***	0.26 ***	0.24 **	0.20 **
India				
Telugu	0.19 ***	0.16 ***	0.16 ***	0.14 ***
English	0.32 **	0.39 ***	0.31 **	0.38 ***
Peru				
Spanish	0.39 ***	0.33 ***	0.38 ***	0.32 ***
Vietnam				
Tieng Viet Nam	0.26 ***	0.28 ***	0.27 ***	0.31 ***
PPVT				
Ethiopia				
Amharic	0.26 ***	0.18 ***	0.29 ***	0.20 ***
India				
Telugu	0.19 ***	0.14 ***	0.20 ***	0.16 ***
Peru				
Spanish	0.43 ***	0.39 ***	0.41 ***	0.37 ***
Vietnam				
Tieng Viet Nam	0.20 ***	0.20 ***	0.24 ***	0.25 ***

***p<.001, **p<.01, *p<.05

Finally, we explore the association of the cognitive scores with children's school attendance (yes or no) and child's grade (only for those who are attending school). The expectation is that both associations would be positive and statistically significant. Tables 19 and 20 show the results. Overall, the results are as expected. In some cases the association is not statistically significant, which is probably due to a low distribution in the score and/or the school variables presented below. This is an important consideration for future analysis using these scores.

Table 19. Correlation of cognitive measures with children’s educational variables (child attends school and child’s grade) by language in Younger Cohort

	Raw scores		Rasch scores	
	Goes to school	Grade	Goes to school	Grade
EGRA (reading comprehension)				
Ethiopia				
Amharic	0.25 ***	0.36 ***	0.24 ***	0.35 ***
Oromiffa	0.21 ***	0.24 ***	0.20 ***	0.26 ***
Tigrigna	0.21 ***	0.37 ***	0.19 ***	0.33 ***
India				
Telugu	0.14 ***	0.37 ***	0.12 ***	0.35 ***
English	0.09	0.39 ***	0.08	0.38 ***
Peru				
Spanish	0.10 ***	0.32 ***	0.02	0.26 ***
Vietnam				
Tieng Viet Nam	0.19 ***	0.29 ***	0.14 ***	0.25 ***
Maths				
Ethiopia				
Amharic	0.30 ***	0.32 ***	0.32 ***	0.31 ***
Oromiffa	0.20 ***	0.28 ***	0.26 ***	0.28 ***
Tigrigna	0.38 ***	0.58 ***	0.44 ***	0.55 ***
India				
Telugu	0.15 ***	0.42 ***	0.16 ***	0.43 ***
English	0.08	0.48 ***	0.09	0.47 ***
Peru				
Spanish	0.13 ***	0.46 ***	0.14 ***	0.46 ***
Quechua	C	0.39 **	C	0.42 **
Vietnam				
Tieng Viet Nam	0.09 ***	0.49 ***	0.09 ***	0.46 ***
PPVT				
Ethiopia				
Amharic	0.27 ***	0.21 ***	0.29 ***	0.20 ***
India				
Telugu	0.10 ***	0.17 ***	0.12 ***	0.20 ***
Peru				
Spanish	0.10 ***	0.34 ***	0.11 ***	0.33 ***
Vietnam				
Tieng Viet Nam	0.14 ***	0.28 ***	0.21 ***	0.30 ***

C: no available information

***p<.001, **p<.01, *p<.05

Table 20. Correlation of cognitive measures with children’s educational variables (child attends school and child’s grade) by language in Older Cohort

	Raw scores		Rasch scores	
	Goes to school	Grade	Goes to school	Grade
Cloze				
Ethiopia				
Amharic	0.25 ***	0.54 ***	0.23 ***	0.56 ***
Oromiffa	0.23 ***	0.48 ***	0.23 ***	0.48 ***
Tigrigna	0.15 *	0.48 ***	0.18 *	0.47 ***
India				
Telugu	0.45 ***	0.32 ***	0.31 ***	0.33 ***
Peru				
Spanish	0.21 ***	0.53 ***	0.21 ***	0.51 ***
Vietnam				
Tieng Viet Nam	0.38 ***	0.34 ***	0.33 ***	0.30 ***
Maths				
Ethiopia				
Amharic	0.22 ***	0.45 ***	0.20 ***	0.49 ***
Oromiffa	0.27 ***	0.41 ***	0.24 ***	0.42 ***
Tigrigna	0.15 *	0.50 ***	0.18 *	0.51 ***
India				
Telugu	0.42 ***	0.23 ***	0.36 ***	0.23 ***
English	0.06	0.48 ***	0.06	0.49 ***
Peru				
Spanish	0.28 ***	0.47 ***	0.31 ***	0.49 ***
Vietnam				
Tieng Viet Nam	0.42 ***	0.07 *	0.39 ***	0.06
PPVT				
Ethiopia				
Amharic	0.28 ***	0.49 ***	0.26 ***	0.46 ***
India				
Telugu	0.39 ***	0.32 ***	0.40 ***	0.32 ***
Peru				
Spanish	0.20 ***	0.54 ***	0.18 ***	0.52 ***
Vietnam				
Tieng Viet Nam	0.35 ***	0.10 *	0.36 ***	0.07

***p<.001, **p<.01, *p<.05

Box 2. *Vocabulary performance among index children and siblings*

In Round 3, the Young Lives children and a sibling were evaluated with the PPVT in Ethiopia, Peru and Vietnam. For the Younger Cohort a sibling was evaluated if s/he was younger than the index child but older than 4 years. For the present report, we equated the scores of the index child and his/her sibling so that they would be both on the same scale. In Ethiopia a younger sibling of the child in the Older Cohort was also evaluated. The table below shows the mean scores for the PPVT. As expected, given the characteristics of the test, siblings have lower raw and Rasch scores than index children; the means are statistically different.

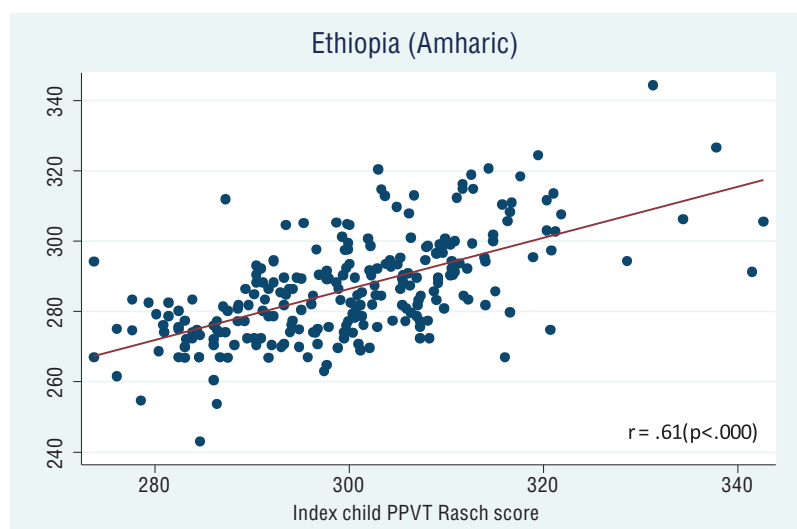
EGRA (reading comprehension) test scores descriptive statistics

	Index child	Sibling
Younger Cohort		
Ethiopia (Amharic)	300.0 ^a	286.4 ^b
Standard deviation	(12.3)	(14.6)
Peru (Spanish)	302.2 ^a	282.5 ^b
Standard deviation	(12.0)	(13.9)
Vietnam (Tieng Viet nam)	301.9 ^a	283.0 ^b
Standard deviation	(12.4)	(18.7)
Older Cohort		
Ethiopia (Amharic)	302.8 ^a	294.1 ^b
Standard deviation	(13.5)	(15.3)

Note: Means with different superscripts are statistically different (significance set at 5%).

Figure 1 shows the distribution of the scores and the correlation coefficient between the index child and his/her sibling. As expected, the correlation between these scores is positive and statistically significant. In Vietnam, the correlation coefficient is low (0.29). A similar result is observed for the Older Cohort. Figure 2 shows that in Ethiopia the correlation between the scores of the index child and his/her sibling is 0.69.

Figure 1. Scatter plot and line fit of the PPVT Rasch scores for the index child and sibling (Younger Cohort)



Box 2. *Vocabulary performance among index children and siblings continued*

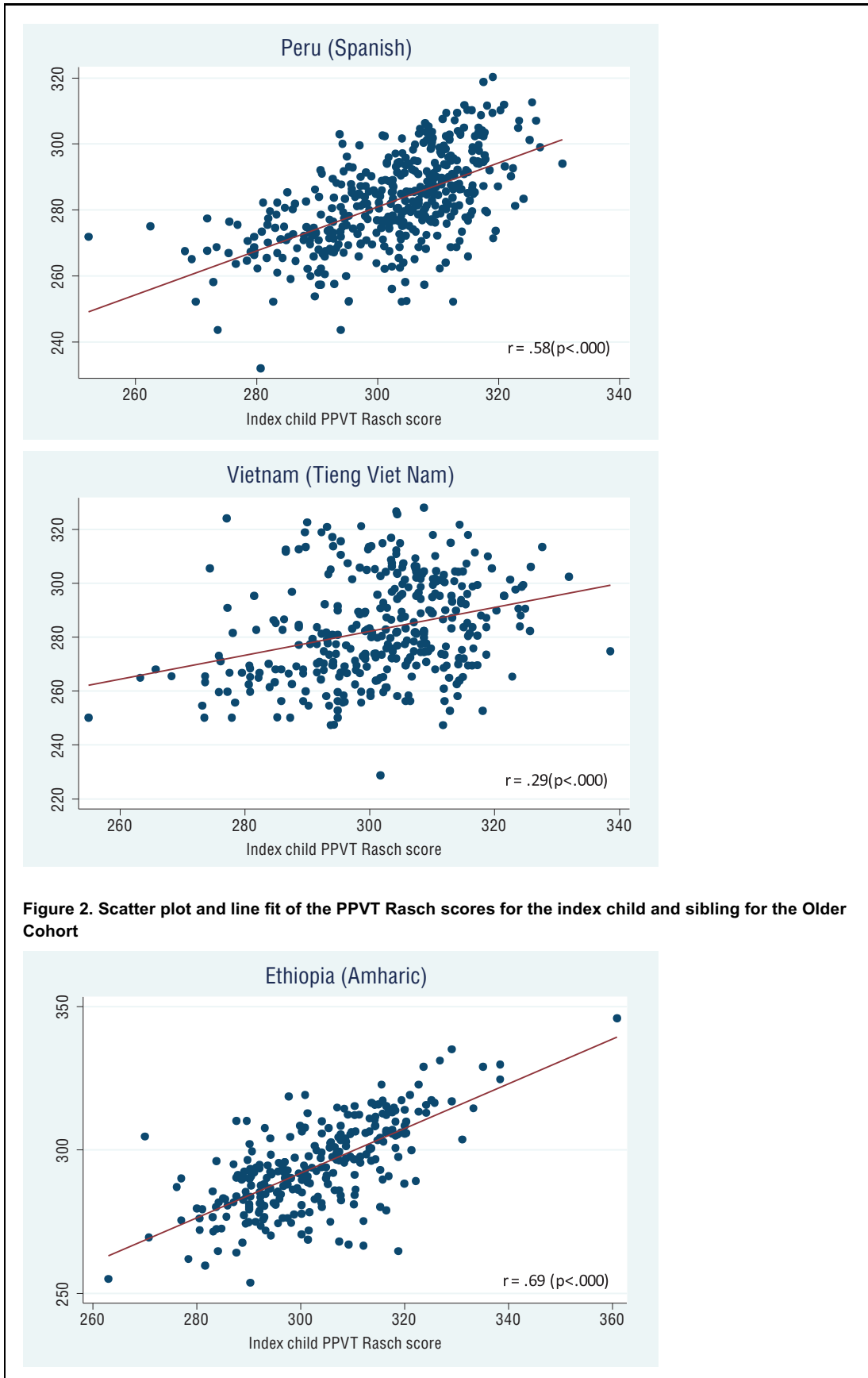


Figure 2. Scatter plot and line fit of the PPVT Rasch scores for the index child and sibling for the Older Cohort

4.6 Correlation among cognitive measures in Round 2 and Round 3

In general, the correlations between the tests administered in Rounds 2 and 3 were statistically significant and positive, as would be expected. As shown in Table 21, for the Younger Cohort the correlation between the PPVT scores and the CDA test administered in Round 2 is associated with the cognitive measures in Round 3, with the strongest correlations between the PPVT and the different cognitive achievements in Round 3. The only exception is for children who took the test in Oromiffa in Ethiopia, where the correlation between the CDA Rasch scores and cognitive tests in Round 3 was weak or non-significant, in comparison with the other languages in Ethiopia and the other countries. In terms of strength of the relationship, there are mixed results: in Ethiopia and Vietnam, the higher correlations occurred between the CDA Rasch scores and the cognitive tests in Round 3, while for Peru and Vietnam the correlation between the PPVT Rasch scores and the cognitive tests in Round 3 was the highest. The correlation coefficients for the Older Cohort are positive and statistically significant for all languages and countries.

Table 21. *Correlation among cognitive measures in Round 2 and Round 3 by language and cohort*

	Younger Cohort					
	PPVT - R2			CDA - R2		
	PPVT - R3	EGRA - R3	Maths - R3	PPVT - R3	EGRA - R3	Maths - R3
Ethiopia						
Amharic	0.52 ***	0.42 ***	0.49 ***	0.46 ***	0.37 ***	0.45 ***
Oromiffa	-	-	-	-	0.08	0.11 *
Tigrigna	-	-	-	-	0.24 ***	0.29 ***
India						
Telugu	0.34 ***	0.24 ***	0.22 ***	0.30 ***	0.24 ***	0.31 ***
Peru						
Spanish	0.69 ***	0.43 ***	0.49 ***	0.43 ***	0.28 ***	0.36 ***
Quechua	-	-	-	-	-	0.12
Vietnam						
Tieng Viet Nam	0.46 ***	0.36 ***	0.32 ***	0.20 ***	0.22 ***	0.24 ***
	Older Cohort					
	PPVT - R2			Maths - R2		
	PPVT - R3	Cloze - R3	Maths - R3	PPVT - R3	Cloze - R3	Maths - R3
Ethiopia						
Amharic	0.61 ***	0.53 ***	0.45 ***	0.38 ***	0.45 ***	0.56 ***
Oromiffa	-	-	-	-	0.46 ***	0.49 ***
Tigrigna	-	-	-	-	0.42 ***	0.55 ***
India						
Telugu	0.59 ***	0.49 ***	0.44 ***	0.53 ***	0.47 ***	0.54 ***
Peru						
Spanish	0.71 ***	0.64 ***	0.55 ***	0.48 ***	0.52 ***	0.43 ***
Vietnam						
Tieng Viet Nam	0.52 ***	0.39 ***	0.44 ***	0.42 ***	0.38 ***	0.51 ***

***p<.001, **p<.01, *p<.05

Finally, we tried to equate the scores between Round 2 and Round 3 for the PPVT, given that the same test was used. However, during this process we found that the levels of item difficulty between Round 2 and Round 3 were different, and the item difficulty was not invariant. The number of items with time-variant difficulty ranged from one-third to one-half of the entire test. Having items with difficulty levels that vary across populations breaks with the sample-independent assumption of the Rasch model (Reeve 2002). Thus we decided not to equate the scores. The reasons for this odd result might be changes in each country in the adaptation of the PPVT for Round 3, and/or in the administration. What we are saying is that in both rounds the PPVT measures verbal abilities (receptive vocabulary), but the scores are not directly comparable. In case researchers would like to see changes in vocabulary over time, we recommend using the relative position of a child in each round (a similar procedure could be used for the maths tests).

4.7 Comparison across cohorts over time (Younger Cohort 2009 and Older Cohort 2002)

One advantage of this study is that it follows two cohorts of children over time. As we stated in the first section, the children from the Older Cohort were on average 8 years old in Round 1. In Round 3, the Younger Cohort on average was also age 8. This allows us to compare these two cohorts over time. However, in Round 1 only the three achievement items mentioned above were administered (reading, writing and numeracy abilities) to the Older Cohort. These were administered to the Younger Cohort in Round 3.

Table 22 shows the percentage of correct responses in Round 1 (Older Cohort) and Round 3 (Younger Cohort); for this analysis only we have combined all languages for a single national result, assuming that the items that were adapted locally were of comparable difficulty. For the reading item, the highest improvement was shown by Ethiopia, with an increase of 6 per cent, followed by Vietnam and India with 2 per cent, and Peru with 1 per cent. However, Vietnam and Peru already had a high percentage of children able to read without problems (above 80 per cent in Round 1).

In writing, only Peru and Vietnam showed an increment over time, with Vietnam showing the highest rate (14 per cent). In Ethiopia and India there was a decrease in the percentage of children who are able to read a sentence.

Finally, for the multiplication test the results are again favorable for Vietnam; this country has 17 per cent more children who are able to answer correctly. This is followed by Ethiopia with 10 per cent and Peru with 7 per cent. In the case of India, again there is a decrease over time (7 per cent).

Table 22. *Percentage of correct answers for 8-year-olds across rounds by country*

	Older Cohort 2002	Younger Cohort 2009	Younger – Older
Read a sentence			
Ethiopia	21.0	26.6	5.6
India	50.8	52.8	2.0
Peru	80.5	81.1	0.6
Vietnam	87.2	89.3	2.1
Write without difficulty			
Ethiopia	22.9	17.4	-5.5
India	51.0	43.3	-7.7
Peru	53.9	61.2	7.3
Vietnam	74.1	87.7	13.6
Multiply 2 by 4			
Ethiopia	32.6	42.6	10.0
India	86.5	79.3	-7.2
Peru	55.3	62.7	7.4
Vietnam	66.3	83.3	17.0

In addition, we created a factor score combining the three measures – reading, writing and numeracy – which could be considered a rough estimate of basic abilities. We made this score comparable over time, pooling Round 1 and Round 3 data sets for each country.¹⁵ Table 23 shows the mean and standard deviation of the factor scores obtained for the 8-year-old children in each country over time. The results are pretty similar to those observed with the percentage of correct responses shown above. Vietnam is showing the highest increment over time, with a change of 0.32 SD ($p < .01$). Peru also showed an increment over time, but this was smaller than Vietnam's (0.13 SD; $p < .01$), while Ethiopia had an increment over time of .07 SD ($p < .06$), but it was not statistically significant. In the case of India, there is a decrease for 8-year-olds over time (-0.12 SD; $p < .01$).

Table 23. *Distribution of the basic cognitive abilities scores over time*

	Older Cohort 2002	Younger Cohort 2009	Younger – Older
Ethiopia	0.00 (1.00)	0.07 (1.03)	0.07
India	0.00 (1.00)	-0.12 (1.09)	-0.12*
Peru	0.00 (1.00)	0.13 (1.02)	0.13*
Vietnam	0.00 (1.00)	0.32 (0.79)	0.32*

* Difference statistically significant at 1% with robust standard errors.

Note: The factor scores were centred on the mean of the Older Cohort in R1 with a mean of 0 and a standard deviation of 1.

15 For this analysis we did not use sampling weights, since the attrition rate across rounds is low (Outes-Leon and Dercon 2008); therefore it does not bias our estimates and score comparison across rounds.

Finally, it might be of interest for some researchers to compare skills across cohorts in Round 3 (beyond the PPVT). This is why we repeated four maths items across cohorts (Appendix 1 has the details of those items). Table 24 shows the percentage of correct answers by country and cohort. As expected, the Older Cohort performs best in all cases. Another interesting result is that on average the difference between Older and Younger Cohorts is largest for Peru (55 per cent average of the four items), followed by Vietnam (45 per cent), India 40 per cent), and finally Ethiopia (35 per cent).

Table 24. *Repeated maths items across cohorts in Round 3 by country*

	Younger Cohort	Older Cohort	Older – Younger
Ethiopia	(n=1,883)	(n=933)	
Item 1905	2.1	29.5	27.4
Item 2209	2.7	47.6	44.9
Item 2310	0.7	30.0	29.3
Item 2807	0.3	39.4	39.2
India	(n=1,913)	(n=964)	
Item 1905	14.8	47.6	32.8
Item 2209	10.5	57.9	47.4
Item 2310	7.1	43.3	36.1
Item 2807	5.6	50.2	44.6
Peru	(n=1,921)	(n=673)	
Item 1905	19.2	68.6	49.4
Item 2209	14.4	71.3	57.0
Item 2310	9.5	62.0	52.5
Item 2807	3.3	65.8	62.5
Vietnam	(n=1,960)	(n=967)	
Item 1905	31.1	68.4	37.3
Item 2209	48.4	86.0	37.6
Item 2310	42.5	76.9	34.4
Item 2807	6.3	77.9	71.5

5. Conclusions and final considerations

This report described the methods used to develop and administer the cognitive and achievement tests used in Round 3. We also present psychometric analysis to show its reliability and validity. These include raw and converted scores (Rasch). The measures analysed here are Peabody Picture Vocabulary Test (PPVT), the Early Grade Reading Assessment (EGRA), the mathematics achievement test (versions for Older and for Younger Cohorts) and a Cloze reading-comprehension test. The PPVT, EGRA and Maths tests were administered to the Younger Cohort, and the PPVT, Cloze and Maths tests were administered to the older one.

In order to increase the reliability and validity of the tests, some items with poor indicators (such as those with poor in-fit or low item–test correlation) were excluded from the scores,¹⁶ generating a new score for each child (both raw and converted into a Rasch scale). A corrected data base was generated with the original raw scores and corrected raw scores and, based on these, a Rasch score was generated for all children. One of the main differences from the analysis in Round 2 is that we did not exclude items that presented Differential Item Functioning (DIF) by gender or language in the case of Maths test scores. We used a method that helped us to reduce the number of items deleted and increase the reliability and validity of our measures. Also, we correlated these scores with those obtained using the method employed in Round 2, and the correlations were above .90 in most of the tests, except maths achievement, where they were above .85.

While DIF procedures allowed us to correct for language bias in the maths scores, for verbal tests (EGRA and Cloze) we strongly suggest that future analysis is applied within each language. This is mostly because the process of adapting tests did not perform a sufficiently thorough process to ensure that the levels of difficulty are comparable across languages. Still, within a given language it is reasonable to assume that, given that administration was standardised (in most cases), results across children should be comparable.

Given that the tests were designed or selected on the basis of norms, we expected that test scores would be normally distributed; in many cases this was not achieved (see Tables 7 and 8 above). This was mostly due to attempting to develop a single set of tests, while the variation in abilities across languages within countries and across countries was very high. Thus for some children we may not have estimated their abilities well, be it that they scored at the lower or bottom end of the distributions. Thus we suggest that researchers using the scores check whether a normal distribution will affect their estimates.

The reliability indexes of each of the tests for each country and language were calculated using both Classical Test theory (CTT) and Item Response theory (IRT). The reliability indexes in most cases were acceptable (see Tables 9 to 12 above). This suggests that scores are stable, but again caution must be exercised when the reliability indexes are low.

Regarding the validity of the instruments used in Round 3, in most cases test scores are associated with several variables, as would be predicted given our aims. Thus children with better-educated parents, who are in school, in higher grades and from more wealthy families tend to score higher. There are some differences between boys and girls that are worth exploring in further analysis.

Round 3 included a measure of siblings using the PPVT. As expected, younger siblings had lower scores, but these were correlated positively. This data set is quite unique, as it will allow for intra-household estimates of variation, which are rare internationally.

In addition to the test mentioned above, reading, writing and numeracy assessment items administered to the children from the Older Cohort in Round 1 were also administered in Round 3, but this time to the children from the Younger Cohort. These are relatively weak measures of school achievement, but they allow for comparison over time and across cohorts, as presented above and in the technical report for the psychometric tests for Round 2 (Cueto et al. 2009).

¹⁶ See Appendix 3 for further details about the items deleted.

The psychometric analyses performed with each cognitive instrument administered to both age cohorts gives us an idea of what recommendations could be made in order to improve or discard the instruments administered in this round when planning Round 4 of the Young Lives Study.

- **PPVT.** On the one hand this test provides very interesting data and it has been used for several Young Lives papers, usually performing in ways that would be predicted by theory. It must be remembered, however, that it has limitations, arising mostly from the fact that it is an instrument developed in English in the USA. While local teams have worked to adapt the test to its local language and we have analysed DIF, some bias is likely remain. Again, we suggest not using the scores for comparisons across different language groups. For R4, the test perhaps could be used for the Younger Cohort but not (unless further adapted) for the Older Cohort, as many children were already reaching the higher items in Round 3.
- **Mathematics achievement tests.** As mentioned above, the maths test developed for the Younger Cohort is different from that for the Older Cohort, yet the results in general show good psychometric properties (i.e. high reliability, normal distribution and correlations with expected variables). This will be an interesting area to evaluate in both cohorts in R4. For comparison (equating) purposes across rounds, it would be good to retain some tests for comparison purposes but to introduce new ones of differing levels of difficulty.
- **EGRA.** This test also had good psychometric properties for many groups, but it seems unlikely to be usable in R4, as many children will require more difficult reading items in order to assess their range of skills. One option for Round 4 could be to construct a Cloze test, retaining some of the items used for the Older Cohort in Round 3.
- **Cloze.** The psychometric properties of this test are acceptable for most countries and languages analysed here. However, in most of the cases the distribution of test scores (both raw and Rasch scores) was found to be negatively skewed, showing a ceiling effect for many children. This indicates that in order to be used in Round 4 it would need to be modified by retaining some of the more difficult items and introducing new ones that could help to increase its reliability and validity to measure reading comprehension skills.

As indicated at the start of this report, the tests that we have selected to measure vocabulary, mathematics skills and reading comprehension are not supposed to capture abilities that are locally (culturally) relevant in specific contexts. The purpose of these tests in general is to capture a range of skills that are usually expected of children who attend formal schooling. For the most part, the results shown above suggest that indeed the tests are capturing this construct. With these caveats in mind, together with those mentioned throughout the report, we trust that the psychometric results will be helpful for researchers interested in using the Young Lives data base.

References

- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (1999) *Standards for Educational and Psychological Testing*, Washington DC: American Educational Research Association.
- Baker, F. (2004) *Item Response Theory: Parameter Estimation Techniques* (2nd ed.), New York: M. Dekker.
- Campbell, J.M. (1998) 'Review of the Peabody Picture Vocabulary Test – Third Edition', *Journal of Psychoeducational Assessment* 16.4: 334–8.
- Campbell, J.M., S.K. Bell and L.K. Keith (2001) 'Concurrent Validity of the Peabody Picture Vocabulary Test – Third Edition as an Intelligence and Achievement Screener for Low SES African American Children', *Assessment* 8.1: 85–94.
- Crocker, L. and J. Algina (2006) *Introduction to Classical and Modern Test Theory*, Pacific Grove, CA: Wadsworth.
- Cueto, S., J. Leon, G. Guerrero and I. Munoz (2009) *Psychometric Characteristics of Cognitive Development and Achievement Instruments in Round 2 of Young Lives*, Technical Note 15, Oxford: Young Lives.
- Dunn, L. and L. Dunn (1997) *Examiner's Manual for the PPVT-III. Form IIIA and IIIB*, Minnesota: AGS.
- Dunn, L., E. Padilla, D. Lugo and L. Dunn (1986) *Manual del Examinador para el Test de Vocabulario en Imágenes Peabody: Adaptación Hispanoamericana ((Peabody Picture Vocabulary Test: Hispanic-American Adaptation)*, Minnesota: AGS.
- Gray, S., E. Plante, R. Vance and M. Henrichsen (1999) 'The Diagnostic Accuracy of Four Vocabulary Tests Administered to Preschool-Age Children', *Language, Speech and Hearing Services in Schools* 30: 196–206.
- Gujarati, D. N. (2003) *Basic Econometrics* (4th ed.), Boston: McGraw Hill.
- Hambleton, Ronald K., H. Swaminathan and H. Jane Rogers (1991) *Fundamentals of Item Response Theory*, Newbury Park, CA: Sage.
- International Association for the Evaluation of Educational Achievement (IEA) (2007) *TIMSS 2003: Mathematics Items, Released Set for Fourth Grade* (see: http://timss.bc.edu/PDF/T03_RELEASED_M4.pdf).
- International Association for the Evaluation of Educational Achievement (IEA) (2007) *TIMSS 2003: Mathematics Items, Released Set for Eighth Grade* (see: http://timss.bc.edu/PDF/T03_RELEASED_M8.pdf).
- Linacre, M. (2008) *A User's Guide to WINSTEPS MINISTEPS Rasch Model Computer Programs*, Chicago: MESA Press.
- Matlock-Hetzel, S. (1997) 'Basic Concepts in Item and Test Analysis', available at <http://www.ericae.net/ft/tamu/Espy.htm> (accessed 1 December 2008).
- OECD (2009) *PISA Take the Test: Sample Questions from the OECD's PISA Assessments*, Paris: OECD Publications.

Outes-Leon, I. and S. Dercon (2008) *Survey Attrition and Attrition Bias in Young Lives*, Technical Note 5, Oxford: Young Lives.

Raven, J., J.C. Raven and J.H. Court (2003) *Manual for Raven's Progressive Matrices and Vocabulary Scales*, San Antonio, TX: Harcourt Assessment.

Reeve, B.B. (2002) 'An Introduction to Modern Measurement Theory.' National Cancer Institute.

Tabachnick, B.G. and L.S. Fidell (1996) *Using Multivariate Statistics* (3rd ed.), New York: Harper Collins.

Williams, K.T. and J. Wang (1997) *Technical References to the Peabody Picture Vocabulary Test – Third Edition*, Circle Pines, MN: American Guidance Service.

Appendix 1. Additional characteristics and sources of items used in the mathematics tests

Table 1. *Common items across rounds or cohorts*

Younger Cohort	Older Cohort	Comments
Item 1.68 in survey (item 9 in test)		The same item used with the Older Cohort in Rounds 1 and 2.
Item 1.82 in survey (item 19 in test)	Item 1.54 in survey (item 5 in test)	They are the same item.
Item 1.84 (item 21 in the test)		Same item was used with Older Cohort in Round 2 (item 2 in R2).
Item 1.85 (item 22 in the test)	Item 1.58 in survey (item 9 in test)	They are the same item.
Item 1.86 (item 23 in the test)	Item 1.59 in survey (item 10 in test)	They are the same item.
Item 1.91 (item 28 in the test)	Item 1.56 in survey (item 7 in test)	They are the same item.

Table 2. *Sources of mathematics items*

Younger Cohort	Older Cohort	Source
Item 1.92 (item 29 in the test)		TIMSS 4th grade-2003 (Item code M031305)
	Item 1.76 (item 21 in the test)	TIMSS 2003- 8th grade (Item M032699).
	Item 1.77 (item 22 in the test)	PISA 2003 (Item M515ShKid_ENG2)
	Item 1.78 (item 23 in the test)	TIMSS 2003- 8th grade (Item M032489)
	Item 1.79 (item 24 in the test)	TIMSS 2003- 8th grade (Item M012014)
	Item 1.80 (item 25 in the test)	TIMSS 2003- 8th grade (Item M012001)
	Item 1.81 (item 26 in the test)	TIMSS 2003- 8th grade (Item M012013)
	Item 1.82 (item 27 in the test)	Used in Round 2 for Older Cohort (item 10). TIMSS 2003-8th grade (Item M032671).
	Item 1.83 (item 28 in the test)	TIMSS 2003- 8th grade (Item M022148)
	Item 1.84 (item 29 in the test)	PISA 2003 (Item M547Stair-Eng3)
	Item 1.85 (item 30 in the test)	TIMSS 2003- 8th grade (Item M022156)

All other items were developed by the Young Lives team for this study.

Sources for documents:

IEA (2007) *TIMSS 2003: Mathematics Items, Released Set for Fourth Grade*. Retrieved from http://timss.bc.edu/PDF/T03_RELEASED_M4.pdf.

IEA (2007) *TIMSS 2003: Mathematics Items, Released Set for Eight Grade*. Retrieved from http://timss.bc.edu/PDF/T03_RELEASED_M8.pdf.

OECD (2009) *PISA Take the Test: Sample Questions from the OECD's PISA Assessments, Paris*: OECD Publications.

Appendix 2. Pearson correlation between the Rasch scores using the method in Round 2 and the method used in Round 3

	Ethiopia				India			Peru			Vietnam
	All	Amharic	Oromiffa	Tigrigna	All	Telugu	English	All	Spanish	Quechua	Tieng Viet Nam
Younger Cohort											
PPVT	-	0.95	-	-	-	0.93	-	-	0.99	-	0.93
EGRA	-	0.99	0.99	0.95	-	0.98	0.99	-	0.99	-	0.99
MATHS	0.86	-	-	-	0.95	-	-	0.97	-	-	0.99
Older Cohort											
PPVT	-	0.96	-	-	-	0.96	-	-	0.99	-	0.96
CLOZE	-	0.99	0.99	0.98	-	0.97	-	-	0.99	-	0.99
MATHS	0.92	-	-	-	0.94	-	-	0.99	-	-	0.99

Note: All the correlations are statistically significant at 1%.

Appendix 3. Item statistics for tests administered in Round 3

Table 1. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – Ethiopia (Amharic)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴	
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender			
											Male			Female
read1	769	36		57	7	0.47	0.65	0.99	1.05	0.48	*	__D	Yes	
read2	769	9		79	12	3.11	0.42	0.98	0.66	-0.51		__	Yes	
read3	769	37		51	13	0.40	0.73	0.78	0.65	0.68	*	__D	Yes	
read4	769	34		57	8	0.57	0.73	0.78	0.59	0.52	*	__D	Yes	
read5	769	22		62	16	1.57	0.65	0.82	0.51	-0.07		__	Yes	
read6	769	15		68	17	2.33	0.57	0.84	0.45	0.11		__C	Yes	
read7	769	36		53	11	0.45	0.78	0.65	0.47	0.12		__C	Yes	
read8	769	41		48	11	0.05	0.78	0.67	0.60	0.19		__	Yes	
oral1	777	85		14	1	-3.36	0.42	1.04	1.45	0.07		__	Yes	
oral2	777	91		7	1	-4.28	0.33	1.02	2.08	-0.17		__C	Yes	
oral3	777	21		70	10	1.73	0.44	1.29	1.41	-0.43		__	Yes	
oral4	777	24		68	8	1.41	0.49	1.22	1.21	-0.54	*	__D	Yes	
oral5	777	30	47	16	7	-2.54	0.35	1.26	2.82	-0.47	*	__CD	Yes	
oral6	777	35	34	21	10	-1.92	0.37	1.38	2.87	-0.51	*	__CD	Yes	

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 2. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – Ethiopia (Oromiffa)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴	
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender			
											Male			Female
read1	381	25		74	1	1.57	0.73	0.78	0.46	0.21		__C	Yes	
read2	381	10		88	3	3.60	0.46	1.23	0.55	0.00		__	Yes	
read3	381	28		68	4	1.33	0.76	0.68	0.44	-0.10		__C	Yes	
read4	381	24		75	1	1.76	0.72	0.79	0.44	0.11		__C	Yes	
read5	381	20		76	4	2.20	0.65	0.89	0.64	-0.33		__	Yes	
read6	381	17		77	6	2.51	0.63	0.90	0.44	0.08		__C	Yes	
read7	381	29		71	1	1.26	0.76	0.69	0.54	0.08		__	Yes	
read8	381	34		65	0	0.73	0.84	0.45	0.31	0.12		__BC	No	
oral1	367	85		14	1	-3.41	0.39	1.11	2.81	0.45		__C	Yes	
oral2	367	84		12	4	-3.30	0.46	0.93	1.54	0.20		__C	Yes	
oral3	367	65		29	7	-1.48	0.43	1.35	2.70	-0.30		__C	Yes	
oral4	367	43		50	7	0.13	0.50	1.54	1.66	0.02		__BC	No	
oral5	367	71	22	2	5	-4.61	0.29	1.17	9.90	-1.14	*	__CD	Yes	
oral6	367	56	9	8	8	-2.28	0.46	1.05	3.11	0.06		__C	Yes	

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 3. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – Ethiopia (Tigrigna)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
							Male				Female		
read1	291	8		92	0	1.25	0.50	0.90	0.55	-0.18		__C_	Yes
read2	291	1		97	1	3.82	0.42	0.82	0.05	2.27		__C_	Yes
read3	291	4		92	4	2.19	0.54	0.65	0.16	0.18		__C_	Yes
read4	291	4		92	3	2.07	0.53	0.73	0.17	-0.07		__C_	Yes
read5	291	2		93	5	3.23	0.43	0.82	0.17	-1.46		__C_	Yes
read6	291	1		93	6	4.23	0.31	1.14	0.14	0.39		__C_	Yes
read7	291	4		92	3	2.07	0.53	0.72	0.18	-0.54		__C_	Yes
read8	291	2		94	4	3.23	0.47	0.71	0.07	-0.45		__C_	Yes
oral1	345	90		10	0	-5.83	0.35	0.91	0.43	-1.10	*	__CD	Yes
oral2	345	82		15	3	-4.91	0.34	1.01	1.71	0.00		__C_	Yes
oral3	345	27		60	13	-0.91	0.42	1.10	1.80	-0.31		__C_	Yes
oral4	345	35		46	19	-1.52	0.37	1.15	1.80	0.05		__C_	Yes
oral5	345	37	42	2	18	-4.66	0.33	0.99	2.35	0.64		__C_	Yes
oral6	345	29	46	7	18	-4.28	0.33	1.08	1.92	0.46		__C_	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
 Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 4. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – India (Telugu)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
							Male				Female		
read1	1464	29		60	11	0.67	0.77	0.68	0.47	0.54	*	__CD	Yes
read2	1464	8		66	25	3.52	0.53	0.94	0.46	0.13		__C_	Yes
read3	1464	34		56	10	0.25	0.78	0.66	0.74	0.35	*	__D	Yes
read4	1464	21		69	10	1.50	0.68	0.92	1.27	0.52	*	__D	Yes
read5	1464	5		64	31	4.42	0.44	1.05	0.37	-0.25		__C_	Yes
read6	1464	9		66	25	3.37	0.56	0.88	0.39	0.13		__C_	Yes
read7	1464	27		63	10	0.82	0.75	0.74	0.59	0.34	*	__D	Yes
read8	1464	31		57	12	0.46	0.77	0.68	0.79	0.68	*	__D	Yes
oral1	1431	95		4	1	-5.78	0.26	1.08	3.51	0.00		__C_	Yes
oral2	1431	90		6	4	-4.49	0.37	0.92	2.46	-0.44	*	__CD	Yes
oral3	1431	70		12	18	-2.40	0.44	1.15	2.67	-0.26		__C_	Yes
oral4	1431	32		47	21	0.48	0.58	1.31	1.27	-0.74	*	__D	Yes
oral5	1431	40	21	14	26	-1.68	0.53	1.11	1.50	-0.29	*	__CD	Yes
oral6	1431	34	19	13	33	-1.14	0.47	1.35	2.45	-0.57	*	__CD	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
 Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 5. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – India (English)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
							Male				Female		
read1	407	48		29	23	-1.15	0.69	0.94	0.76	0.08		___	Yes
read2	407	5		44	51	3.45	0.29	1.01	6.54	0.16		__C_	Yes
read3	407	52		29	19	-1.43	0.66	1.04	0.97	0.39		___	Yes
read4	407	24		46	30	0.77	0.60	0.98	0.98	0.28		___	Yes
read5	407	7		41	52	2.92	0.42	1.00	0.50	-0.70		__C_	Yes
read6	407	3		41	56	4.01	0.30	0.92	0.50	0.97		__C_	Yes
read7	407	38		39	24	-0.40	0.72	0.82	0.76	0.49		___	Yes
read8	407	20		39	42	1.16	0.66	0.74	0.48	0.12		__C_	Yes
oral1	360	86		10	4	-3.89	0.43	1.05	0.99	0.13		___	Yes
oral2	360	77		11	12	-2.87	0.56	0.89	0.72	-0.17		___	Yes
oral3	360	55		17	28	-1.24	0.61	0.98	1.18	0.00		___	Yes
oral4	360	39		22	39	-0.18	0.58	1.13	1.07	-0.35		___	Yes
oral5	360	34	15	13	39	-0.86	0.59	1.06	1.48	-0.20		___	Yes
oral6	360	28	3	3	48	-0.28	0.50	1.28	1.79	-0.67	*	__CD	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
- Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 6. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – Peru (Spanish)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
							Male				Female		
read1	1622	62		37	1	0.30	0.53	1.01	1.03	0.34		___	Yes
read2	1622	42		53	4	1.39	0.54	0.96	0.89	1.39		___	Yes
read3	1622	84		13	3	-1.37	0.59	0.80	0.71	-1.17		___	Yes
read4	1622	38		61	1	1.65	0.51	0.98	0.91	1.65		___	Yes
read5	1622	59		35	6	0.48	0.64	0.84	0.75	0.50		___	Yes
read6	1622	35		59	6	1.80	0.54	0.91	0.80	1.77		___	Yes
read7	1622	73		26	1	-0.42	0.63	0.82	0.74	-0.33		___	Yes
read8	1622	74		25	1	-0.49	0.64	0.80	0.74	-0.40		___	Yes
oral1	1743	93		7	0	-2.69	0.28	1.16	1.55	-2.69		__C_	Yes
oral2	1743	95		4	1	-3.13	0.31	1.03	0.98	-3.16		___	Yes
oral3	1743	66		30	3	-0.05	0.41	1.24	1.23	-0.32	*	__D	Yes
oral4	1743	38		58	4	1.58	0.41	1.12	1.37	1.58		___	Yes
oral5	1743	77	2	18	3	-0.98	0.46	1.08	1.10	-1.15		___	Yes
oral6	1743	24	8	64	3	1.94	0.36	1.15	1.51	1.99		__C_	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
- Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 7. Item statistics in the Early Grade Reading Assessment (EGRA) – Reading Comprehension items – Vietnam (Tieng Viet Nam)

	N	Response options				Difficulty ² (IRT)	CTT item-fit indicator		IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Partially correct (%)	Incorrect (%)	NR ¹ (%)		Item-test correlation	In-fit (IRT)	Outfit (IRT)	Gender				
										Male	Female			
read1	1883	61		32	8	1.08	0.55	0.94	0.91	0.19		___	Yes	
read2	1883	27		54	19	3.05	0.42	1.05	1.14	-0.08		___	Yes	
read3	1883	75		12	14	0.21	0.54	0.93	0.94	0.37	*	___D	Yes	
read4	1883	55		39	5	1.40	0.43	1.13	1.22	0.00		___	Yes	
read5	1883	41		39	20	2.17	0.54	0.91	0.87	0.25	*	___D	Yes	
read6	1883	55		30	15	1.38	0.49	1.03	1.02	0.00		___	Yes	
read7	1883	87		12	1	-0.80	0.54	0.86	0.66	0.27		___	Yes	
read8	1883	82		14	4	-0.36	0.53	0.90	0.83	0.17		___	Yes	
oral1	1912	97		3	0	-2.81	0.22	1.05	1.02	-0.22		___	Yes	
oral2	1912	97		3	1	-2.71	0.25	1.04	0.90	-1.07	*	___D	Yes	
oral3	1912	84		13	4	-0.55	0.38	1.10	1.27	-0.59	*	___D	Yes	
oral4	1912	78		17	5	-0.08	0.46	1.03	1.07	-0.38	*	___D	Yes	
oral5	1912	2	88	5	4	-1.37	0.38	1.01	0.87	-0.02		___	Yes	
oral6	1912	6	78	12	3	-0.61	0.44	1.03	0.93	-0.14		___	Yes	

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 8. Item statistics in the Maths achievement test for Younger Cohort – Ethiopia

	N	Response options			Difficulty (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis				Warnings	Item kept for the analysis	
		Correct (%)	Incorrect (%)	NR (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender	Language				
											Male-Female	First-Second			First-Third
math01	1498	60	40	0	-4.39	0.68	0.89	0.73	0.27 *	0.26	-0.42 *	-0.67 *	___DE	Yes	
math02	1498	52	47	1	-3.71	0.69	0.97	1.05	0.11	0.03	-0.17	-0.20	_____	Yes	
math03	1498	56	43	1	-4.02	0.63	1.16	1.71	0.06	0.66 *	0.43 *	-0.23	__C_E	Yes	
math04	1498	90	9	1	-7.62	0.35	1.01	1.83	-0.22	-0.29	0.56	0.85 *	__C_E	Yes	
math05	1498	22	72	6	-0.85	0.59	1.32	1.82	0.30	0.14	0.66 *	0.52	__C_E	Yes	
math06	1498	14	80	6	0.10	0.51	1.34	2.83	0.45 *	0.31	1.48 *	1.17 *	__CDE	Yes	
math07	1498	77	21	1	-6.20	0.50	1.02	1.30	0.15	0.10	0.26	0.16	_____	Yes	
math08	1498	39	58	3	-2.91	0.70	1.28	1.35	-0.27	-0.29	0.64 *	0.92 *	___E	Yes	
math09	1498	46	50	4	-3.46	0.65	1.33	1.52	0.00	0.32	0.47 *	0.15	__C_E	Yes	
math10	1498	54	16	31	-3.78	0.75	0.68	0.54	-0.05	-0.10	-0.38 *	-0.29	___E	Yes	
math11	1498	37	33	31	-2.35	0.74	0.94	0.86	-0.14	0.02	-0.70 *	-0.72 *	___E	Yes	
math12	1498	29	31	39	-1.54	0.76	0.82	0.62	-0.05	-0.23	-0.20	0.03	_____	Yes	
math13	1498	30	36	34	-1.63	0.74	0.89	0.65	-0.13	-0.37	-0.05	0.32	_____	Yes	
math14	1498	32	32	36	-1.82	0.77	0.78	0.56	0.00	-0.53 *	-0.75 *	-0.22	___E	Yes	
math15	1498	9	38	53	1.08	0.58	0.83	0.27	0.00	0.02	-0.90	-0.92	__C__	Yes	
math16	1498	21	37	42	-0.74	0.72	0.84	0.55	-0.13	-0.03	-0.44	-0.41	_____	Yes	
math17	1498	17	34	48	-0.22	0.67	0.88	0.53	0.00	-0.19	-0.44	-0.25	_____	Yes	
math18	1498	7	32	61	1.59	0.53	0.84	0.22	0.02	-0.65	-1.33	-0.67	__C__	Yes	
math19	1498	3	29	68	3.13	0.37	0.90	0.23	0.00	-0.29	-1.55	-1.26	__C__	Yes	
math20	1498	5	22	73	2.24	0.46	0.83	0.21	0.13	-0.35	-2.49	-2.14	__C__	Yes	
math21	1498	8	23	69	1.32	0.51	1.02	0.49	-0.43	0.57 *	-0.13	-0.71	__C_E	Yes	
math22	1498	3	19	77	2.77	0.38	0.95	1.26	-0.73 *	-0.30	-1.92	-1.61	__D	Yes	
math23	1498	1	14	85	4.55	0.24	0.91	1.39	-0.26	0.46	-0.05	-0.51	_____	Yes	
math24	1498	2	12	86	3.37	0.33	0.91	1.85	0.07	-0.20	0.67	0.87	__C__	Yes	
math25	1498	2	10	88	3.81	0.26	1.02	7.63	-0.45	0.42	1.29	0.87	__C__	Yes	
math26	1498	1	8	91	5.33	0.21	0.83	0.14	-0.84	0.91	0.17	-0.74	__C__	Yes	
math27	1498	2	7	91	3.21	0.26	1.24	4.99	-0.20	-0.03	0.62	0.65	__C__	Yes	
math28	1498	0	7	93	6.00	0.07	1.44	9.90	0.26	1.81	3.73 *	1.92	A_C_E	No	
math29	1498	0	6	93	6.73	0.13	1.21	0.07	0.35	-0.17	-0.17	0.00	__C__	Yes	

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic
- Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%. E: The difference between language groups is significant at 5%.

Table 9. Item statistics in the Maths achievement test for Younger Cohort – India

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator Item-test correlation	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴	
		Correct (%)	Incorrect (%)	NR ¹ (%)			In-fit (IRT)	Out-fit (IRT)	Gender	Language			
						Male - Female			First - Second				
math01	1861	93	6	0	-6.04	0.39	1.00	0.51	-0.19		2.62 *	___E	Yes
math02	1861	84	15	1	-4.26	0.53	0.95	0.72	-0.15		0.58 *	___E	Yes
math03	1861	83	15	2	-4.08	0.50	1.06	6.23	0.23		0.34	__C__	Yes
math04	1861	96	4	1	-6.78	0.28	1.08	9.90	0.06		0.22	__C__	Yes
math05	1861	31	61	8	0.81	0.57	1.34	1.80	-0.51 *		0.04	__CD__	Yes
math06	1861	18	72	10	2.15	0.53	1.23	2.31	-0.44 *		-0.12	__CD__	Yes
math07	1861	83	15	2	-4.07	0.46	1.30	9.90	0.14		-1.42 *	__C_E	Yes
math08	1861	68	25	8	-2.40	0.63	1.06	1.87	-0.16		-0.50 *	__C_E	Yes
math09	1861	80	10	10	-3.69	0.55	1.00	1.31	0.51 *		-0.28	__D__	Yes
math10	1861	79	11	10	-3.64	0.58	0.91	1.47	0.23		-0.47 *	___E	Yes
math11	1861	64	24	13	-2.01	0.66	1.00	1.12	0.15		-0.07	___	Yes
math12	1861	50	28	23	-0.82	0.71	0.85	1.13	0.02		0.16	___	Yes
math13	1861	63	20	18	-1.93	0.69	0.87	1.23	0.10		0.30	___	Yes
math14	1861	55	24	20	-1.27	0.72	0.84	1.23	-0.25		0.63 *	___E	Yes
math15	1861	30	37	32	0.88	0.69	0.89	0.69	0.17		-0.22	___	Yes
math16	1861	37	35	28	0.29	0.65	1.08	1.17	-0.39 *		0.35 *	__DE	Yes
math17	1861	39	26	36	0.15	0.74	0.79	0.62	-0.05		0.04	___	Yes
math18	1861	28	30	42	1.08	0.69	0.87	0.57	0.24		-0.24	___	Yes
math19	1861	15	28	57	2.57	0.57	1.02	0.68	0.00		0.00	___	Yes
math20	1861	10	32	58	3.41	0.49	1.12	0.62	0.23		0.18	___	Yes
math21	1861	19	20	61	2.10	0.63	0.88	0.61	0.11		0.52 *	___E	Yes
math22	1861	11	19	70	3.23	0.56	0.86	0.39	0.00		0.27	__C__	Yes
math23	1861	7	15	78	3.93	0.49	0.91	2.40	0.00		-0.09	__C__	Yes
math24	1861	20	11	69	1.89	0.58	1.08	1.61	0.27		0.26	__C__	Yes
math25	1861	11	9	80	3.18	0.56	0.79	2.05	-0.03		-0.16	__C__	Yes
math26	1861	8	9	83	3.82	0.51	0.84	0.50	0.20		-0.13	__C__	Yes
math27	1861	9	9	82	3.50	0.54	0.76	2.67	0.08		-0.17	__C__	Yes
math28	1861	6	7	87	4.32	0.49	0.75	0.19	0.07		-0.66 *	__C_E	Yes
math29	1861	8	8	84	3.67	0.49	0.94	0.93	-0.44 *		-0.43	__D__	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%. E: The difference between language groups is significant at 5%.

Table 10. *Item statistics in the Maths achievement test for Younger Cohort – Peru*

	N	Response options			Difficulty ² (IRT)	CTT item- fit indicator	IRT item fit Indicator		Bias Analysis ³		Warnings	Item kept for the analysis ⁴	
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender			Language
						Male- Female				First- Second			
math01	1827	95	5	0	-4.89	0.37	1.05	1.88	0.22		0.40	__C__	Yes
math02	1827	91	9	0	-3.87	0.50	0.87	0.81	0.31		-0.67	_____	Yes
math03	1827	93	7	0	-4.31	0.46	0.88	1.03	-0.24		-0.28	_____	Yes
math04	1827	96	3	0	-5.54	0.24	1.13	9.90	0.31		0.19	__C__	Yes
math05	1827	49	49	2	0.11	0.59	1.06	1.38	-0.09		-0.25	_____	Yes
math06	1827	28	69	3	1.57	0.55	1.02	1.75	-0.31	*	0.02	__CD__	Yes
math07	1827	85	15	0	-2.90	0.46	1.20	2.09	-0.28		0.14	__C__	Yes
math08	1827	49	47	4	0.09	0.57	1.09	1.37	-0.31	*	0.53	__D__	Yes
math09	1827	64	23	12	-0.97	0.62	0.98	1.14	0.11		0.51	_____	Yes
math10	1827	93	5	2	-4.33	0.36	1.03	4.21	0.21		0.07	__C__	Yes
math11	1827	79	19	2	-2.23	0.44	1.34	2.80	0.31	*	0.88	* __CDE	Yes
math12	1827	77	17	5	-2.06	0.50	1.14	3.27	0.30	*	0.49	__CD__	Yes
math13	1827	79	17	4	-2.21	0.55	0.99	3.16	0.12		0.16	__C__	Yes
math14	1827	72	21	7	-1.59	0.60	0.97	1.52	0.13		-0.50	__C__	Yes
math15	1827	57	29	14	-0.41	0.62	0.97	1.39	0.21		0.03	_____	Yes
math16	1827	56	30	13	-0.40	0.61	1.01	1.33	0.05		0.69	_____	Yes
math17	1827	55	26	19	-0.27	0.67	0.85	1.15	0.09		-0.36	_____	Yes
math18	1827	53	20	27	-0.16	0.65	0.90	0.96	0.10		-1.27	_____	Yes
math19	1827	20	47	33	2.31	0.49	1.12	0.86	-0.06		0.13	_____	Yes
math20	1827	44	18	37	0.41	0.65	0.89	0.94	-0.07		-0.67	_____	Yes
math21	1827	30	26	44	1.43	0.62	0.88	0.70	-0.21		-1.70	_____	Yes
math22	1827	15	30	56	2.85	0.51	0.94	0.60	-0.18		-1.67	_____	Yes
math23	1827	10	22	68	3.54	0.45	0.98	0.54	-0.06		0.41	_____	Yes
math24	1827	12	11	77	3.26	0.50	0.86	0.50	-0.19		0.13	__C__	Yes
math25	1827	9	7	84	3.72	0.48	0.75	4.70	-0.06		0.60	__C__	Yes
math26	1827	9	7	85	3.75	0.48	0.76	0.35	-0.17		-0.77	__C__	Yes
math27	1827	9	4	87	3.73	0.48	0.75	0.38	0.00		-0.79	__C__	Yes
math28	1827	3	6	91	5.09	0.35	0.81	0.19	0.00		0.00	__C__	Yes
math29	1827	6	5	88	4.27	0.42	0.78	0.39	0.09		-0.26	__C__	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%. E: The difference between language groups is significant at 5%.

Table 11. Item statistics in the Maths achievement test for Younger Cohort – Vietnam

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
						Male				Female		
math01	1909	99	1	0	-4.91	0.30	0.68	0.65	0.20		___	Yes
math02	1909	96	4	0	-2.98	0.30	1.14	2.63	-0.24		___C_	Yes
math03	1909	98	2	1	-4.13	0.34	0.83	0.25	0.53		___C_	Yes
math04	1909	97	3	0	-3.57	0.10	1.70	8.26	0.43		ABC_	No
math05	1909	56	40	3	1.32	0.57	1.14	1.18	0.05		___	Yes
math06	1909	45	51	4	2.08	0.58	1.11	1.32	0.00		___	Yes
math07	1909	92	7	1	-2.03	0.34	1.19	2.62	0.23		___C_	Yes
math08	1909	79	17	4	-0.46	0.53	1.04	1.05	-0.17		___	Yes
math09	1909	84	8	8	-0.96	0.53	0.96	0.86	-0.30	*	___D	Yes
math10	1909	97	2	1	-3.50	0.23	1.08	7.36	-0.15		___C_	Yes
math11	1909	96	3	1	-3.29	0.34	0.94	1.66	0.25		___C_	Yes
math12	1909	93	5	2	-2.31	0.36	1.07	2.46	0.19		___C_	Yes
math13	1909	89	8	2	-1.64	0.42	1.06	1.94	-0.19		___C_	Yes
math14	1909	88	9	2	-1.47	0.44	1.01	1.84	0.00		___C_	Yes
math15	1909	56	34	10	1.34	0.56	1.16	1.26	0.10		___	Yes
math16	1909	89	7	4	-1.57	0.46	0.97	1.34	0.20		___	Yes
math17	1909	88	8	4	-1.42	0.48	0.93	1.20	-0.05		___	Yes
math18	1909	52	37	11	1.61	0.61	1.05	1.06	0.00		___	Yes
math19	1909	32	51	18	3.05	0.55	1.12	1.00	0.24		___	Yes
math20	1909	42	38	20	2.27	0.60	1.09	0.96	-0.06		___	Yes
math21	1909	63	20	17	0.82	0.64	0.92	0.91	0.21		___	Yes
math22	1909	49	29	22	1.78	0.65	0.95	0.87	0.21		___	Yes
math23	1909	43	23	33	2.19	0.65	0.94	0.80	-0.16		___	Yes
math24	1909	49	8	42	1.79	0.70	0.81	0.71	-0.16		___	Yes
math25	1909	53	5	42	1.52	0.73	0.73	0.62	-0.10		___	Yes
math26	1909	33	16	51	2.97	0.68	0.76	0.55	-0.12		___	Yes
math27	1909	48	5	47	1.88	0.72	0.76	0.59	-0.13		___	Yes
math28	1909	6	28	66	6.24	0.34	1.08	0.75	-0.06		___	Yes
math29	1909	28	12	60	3.39	0.64	0.81	0.54	0.05		___	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
 Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 12. Item statistics by item in the Peabody Picture Vocabulary Test for the Younger Cohort, Ethiopia – Amharic

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
		Male-Female												
ppvt001	390	2	2	5	91	0	-3.70	0.22	0.91	0.92	1.02		Yes	
ppvt002	390	3	2	92	3	0	-3.77	0.21	0.93	0.90	2.50	*	Yes	
ppvt003	390	94	3	1	2	0	-4.10	0.20	0.91	0.69	-1.38		Yes	
ppvt004	390	84	6	4	6	0	-2.98	0.30	0.94	0.80	0.38		Yes	
ppvt005	390	4	9	40	47	0	-0.90	0.50	0.99	1.01	1.30		Yes	
ppvt006	390	86	6	5	3	0	-3.12	0.28	0.95	0.79	-0.52		Yes	
ppvt007	390	18	43	15	24	0	-0.69	0.53	0.98	0.95	-2.31	*	Yes	
ppvt008	390	29	14	50	8	0	-1.06	0.51	0.95	0.92	0.72		Yes	
ppvt009	390	24	7	9	60	0	-1.53	0.46	0.93	0.87	1.02		Yes	
ppvt010	390	31	45	11	13	0	-0.83	0.48	1.10	1.17	0.43		Yes	
ppvt011	390	15	67	12	6	0	-1.90	0.42	0.88	0.81	0.33		Yes	
ppvt012	390	6	22	67	5	0	-1.88	0.39	1.04	1.06	-0.25		Yes	
ppvt013	477	8	84	5	3	0	-2.81	0.34	0.83	0.64	0.28		Yes	
ppvt014	477	80	9	6	5	0	-2.48	0.31	1.09	1.25	0.49		Yes	
ppvt015	477	11	17	66	5	0	-1.66	0.46	0.89	0.85	0.90		Yes	
ppvt016	477	85	8	4	3	0	-2.88	0.30	0.95	1.02	-0.26		Yes	
ppvt017	477	13	14	61	11	0	-1.39	0.50	0.86	0.79	-0.67		Yes	
ppvt018	477	18	40	26	16	0	-0.37	0.53	1.17	1.26	0.00		Yes	
ppvt019	477	74	5	12	9	0	-2.12	0.34	1.14	1.16	-1.16		Yes	
ppvt020	477	19	24	19	38	0	-0.27	0.55	1.04	1.04	0.54		Yes	
ppvt021	477	8	15	72	4	0	-2.00	0.38	1.04	1.08	-0.25		Yes	
ppvt022	477	23	12	11	55	0	-1.09	0.46	1.04	1.08	-0.37		Yes	
ppvt023	477	16	40	32	13	0	-0.34	0.57	0.99	0.99	-0.79		Yes	
ppvt024	477	24	8	15	53	0	-1.00	0.49	1.00	1.00	-1.48		Yes	
ppvt025	616	76	11	6	7	0	-1.80	0.50	0.84	0.69	-0.52		Yes	
ppvt026	616	11	11	70	8	0	-1.43	0.53	0.85	0.75	0.13		Yes	
ppvt027	616	15	14	24	46	0	-0.24	0.58	0.96	0.99	0.00		Yes	
ppvt028	616	7	6	12	75	0	-1.73	0.50	0.83	0.73	-0.83		Yes	
ppvt029	616	70	12	9	8	0	-1.46	0.46	1.01	1.02	0.13		Yes	
ppvt030	616	16	63	10	11	0	-1.03	0.56	0.88	0.83	1.19		Yes	
ppvt031	616	12	19	58	11	0	-0.79	0.58	0.85	0.82	-0.48		Yes	
ppvt032	616	9	75	10	6	0	-1.76	0.48	0.88	0.76	-1.34		Yes	
ppvt033	616	16	64	7	13	0	-1.09	0.55	0.88	0.82	-1.10		Yes	
ppvt034	616	57	10	12	20	0	-0.78	0.46	1.13	1.21	-0.57		Yes	
ppvt035	616	17	17	11	56	0	-0.68	0.50	1.09	1.10	-1.64		Yes	
ppvt036	616	9	30	58	3	0	-0.83	0.49	1.06	1.10	1.60		Yes	
ppvt037	612	20	49	12	19	0	-0.08	0.57	1.08	1.10	0.63		Yes	
ppvt038	612	3	4	2	91	0	-2.73	0.48	0.95	0.79	1.51		Yes	
ppvt039	612	11	18	64	7	0	-0.77	0.54	1.01	1.04	1.04		Yes	
ppvt040	612	27	21	13	39	0	1.06	0.60	1.08	1.25	0.75		Yes	
ppvt041	612	30	53	8	9	0	-0.26	0.54	1.08	1.11	0.52		Yes	
ppvt042	612	43	21	17	18	0	0.20	0.59	1.04	1.09	1.81		Yes	
ppvt043	612	26	19	35	18	1	0.60	0.62	1.00	1.06	0.55		Yes	
ppvt044	612	29	29	11	31	0	0.84	0.58	1.10	1.19	0.98		Yes	
ppvt045	612	18	22	3	57	0	-0.45	0.54	1.05	1.06	-0.82		Yes	
ppvt046	612	14	52	20	14	0	-0.22	0.64	0.85	0.83	-0.62		Yes	
ppvt047	612	11	5	76	8	0	-1.43	0.57	0.88	0.76	-0.71		Yes	
ppvt048	612	89	7	3	2	0	-2.46	0.45	1.04	1.17	-0.46		Yes	
ppvt049	666	14	22	56	8	1	-0.16	0.61	0.98	0.97	0.24		Yes	
ppvt050	666	80	9	6	6	0	-1.49	0.55	0.94	0.96	0.46		Yes	
ppvt051	666	51	16	21	11	1	0.06	0.55	1.10	1.11	0.52		Yes	
ppvt052	666	35	43	9	13	0	0.45	0.64	0.93	0.94	0.12		Yes	
ppvt053	666	34	14	18	34	0	0.92	0.58	1.05	1.19	3.12	*	Yes	
ppvt054	666	15	52	23	11	0	0.05	0.55	1.13	1.17	0.00		Yes	

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt055	666	45	41	10	5	0	0.38	0.62	0.98	0.99	1.22		___	Yes
ppvt056	666	8	5	44	43	0	0.47	0.64	0.91	0.88	-0.62		___	Yes
ppvt057	666	16	58	16	10	0	-0.25	0.60	0.99	0.99	1.84		___	Yes
ppvt058	666	28	20	30	22	0	1.13	0.58	1.05	1.12	1.14		___	Yes
ppvt059	666	7	10	80	4	0	-1.52	0.54	0.97	0.96	-0.99		___	Yes
ppvt060	666	14	25	8	53	0	-0.03	0.67	0.85	0.81	0.91		___	Yes
ppvt061	631	1	1	1	97	0	-3.32	0.62	1.06	1.94	1.25		__C_	Yes
ppvt062	631	91	5	1	2	0	-2.21	0.64	0.99	1.33	1.73		___	Yes
ppvt063	631	7	85	4	3	0	-1.53	0.67	0.94	0.97	-1.37		___	Yes
ppvt064	631	6	14	76	4	0	-0.89	0.70	0.85	0.82	-1.52		___	Yes
ppvt065	631	17	8	26	49	0	0.55	0.53	1.20	1.29	-1.38		___	Yes
ppvt066	631	17	19	40	24	0	0.95	0.54	1.14	1.20	-1.04		___	Yes
ppvt067	631	21	54	13	11	0	0.29	0.56	1.16	1.24	2.31	*	__D	Yes
ppvt068	631	70	14	9	6	0	-0.52	0.66	0.96	0.99	-0.71		___	Yes
ppvt069	631	9	11	7	73	0	-0.66	0.61	1.08	1.16	0.91		___	Yes
ppvt070	631	22	58	12	8	0	0.13	0.62	1.05	1.18	1.23		___	Yes
ppvt071	631	13	4	59	23	0	0.04	0.61	1.02	1.01	-1.31		___	Yes
ppvt072	631	71	9	8	12	0	-0.55	0.62	1.06	1.16	-2.14	*	__D	Yes
ppvt073	742	9	84	2	6	0	-1.19	0.59	1.15	1.09	-0.45		___	Yes
ppvt074	742	17	7	6	70	0	-0.27	0.70	0.86	0.80	0.33		___	Yes
ppvt075	742	14	13	67	6	0	-0.12	0.66	0.91	0.88	0.00		___	Yes
ppvt076	742	55	24	13	8	0	0.47	0.60	0.98	1.01	2.11	*	__D	Yes
ppvt077	742	77	5	10	8	0	-0.74	0.65	0.96	1.01	1.41		___	Yes
ppvt078	742	17	6	70	7	0	-0.30	0.63	0.99	1.07	0.00		___	Yes
ppvt079	742	2	87	8	2	0	-1.55	0.68	0.92	0.86	-1.56		___	Yes
ppvt080	742	50	15	25	10	0	0.71	0.60	0.95	0.97	-1.64		___	Yes
ppvt081	742	35	33	11	20	0	1.56	0.38	1.16	1.38	-0.78		___	Yes
ppvt082	742	17	9	7	67	0	-0.12	0.69	0.87	0.81	-0.67		___	Yes
ppvt083	742	15	24	45	15	0	0.98	0.51	1.04	1.14	0.00		___	Yes
ppvt084	742	28	6	11	55	0	0.48	0.57	1.01	1.01	2.69	*	__D	Yes
ppvt085	620	3	84	13	1	0	-0.92	0.71	1.02	1.05	-0.48		___	Yes
ppvt086	620	37	25	21	17	0	2.79	0.34	1.04	1.03	-1.63		___	Yes
ppvt087	620	60	12	19	9	0	0.48	0.66	1.02	1.08	0.28		___	Yes
ppvt088	620	23	29	17	31	0	1.87	0.48	0.96	0.95	0.00		___	Yes
ppvt089	620	20	29	33	18	0	1.75	0.41	1.13	1.19	-2.25	*	__D	Yes
ppvt090	620	55	15	15	14	0	0.71	0.66	0.95	0.93	-1.37		___	Yes
ppvt091	620	18	15	12	55	0	0.72	0.57	1.07	1.17	0.47		___	Yes
ppvt092	620	22	34	40	4	0	1.72	0.46	1.04	1.10	-1.91		___	Yes
ppvt093	620	7	7	84	2	0	-0.90	0.75	0.96	0.85	-0.23		___	Yes
ppvt094	620	10	60	23	7	0	0.46	0.59	1.08	1.09	-1.32		___	Yes
ppvt095	620	6	9	57	28	0	0.63	0.62	1.00	0.97	0.40		___	Yes
ppvt096	620	58	15	9	18	0	0.58	0.59	1.07	1.06	-0.96		___	Yes
ppvt097	468	16	33	29	22	0	2.00	0.49	1.06	1.07	-0.51		___	Yes
ppvt098	468	28	44	12	16	0	1.49	0.57	1.02	1.05	0.84		___	Yes
ppvt099	468	13	16	16	54	0	1.01	0.62	1.04	1.02	-0.81		___	Yes
ppvt100	468	0	3	96	1	0	-2.21	0.82	1.00	1.03	-1.30		___	Yes
ppvt101	468	6	13	13	67	0	0.40	0.76	0.86	0.79	-1.52		___	Yes
ppvt102	468	19	18	14	49	0	2.84	0.35	1.12	1.13	-2.19	*	__D	Yes
ppvt103	468	2	16	69	12	0	0.29	0.72	0.97	0.94	0.71		___	Yes
ppvt104	468	1	2	8	88	0	-1.04	0.82	0.96	0.85	1.80		___	Yes
ppvt105	468	88	7	2	2	0	-1.02	0.81	0.97	0.93	2.23	*	__D	Yes
ppvt106	468	15	38	20	27	0	1.76	0.56	0.93	0.90	-0.46		___	Yes
ppvt107	468	58	17	15	10	0	0.86	0.63	1.04	1.02	-3.12	*	__D	Yes
ppvt108	468	9	20	68	4	0	0.36	0.69	1.04	1.03	0.00		___	Yes
ppvt109	414	2	0	2	96	0	-2.18	0.85	1.02	1.15	2.13	*	__D	Yes
ppvt110	414	14	59	21	6	0	0.92	0.70	0.93	0.93	0.96		___	Yes
ppvt111	414	8	30	58	3	0	0.94	0.72	0.89	0.85	0.55		___	Yes
ppvt112	414	4	16	67	14	0	0.55	0.73	0.94	0.88	0.00		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item- fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male-Female						
ppvt113	414	34	17	37	11	0	2.06	0.46	1.19	1.25	0.40		___	Yes
ppvt114	414	74	3	3	20	0	0.17	0.73	1.07	1.10	-1.09		___	Yes
ppvt115	414	5	2	4	89	0	-0.97	0.85	0.93	0.74	1.13		___	Yes
ppvt116	414	3	85	2	9	0	-0.60	0.80	1.00	1.02	-0.82		___	Yes
ppvt117	414	33	44	17	6	0	2.13	0.50	1.02	1.01	0.56		___	Yes
ppvt118	414	2	90	4	3	0	-1.11	0.84	0.99	0.91	2.60	*	__D	Yes
ppvt119	414	0	2	2	96	0	-1.99	0.84	1.03	1.72	0.00		__C	Yes
ppvt120	414	20	19	41	21	0	1.75	0.57	0.94	0.93	-1.00		___	Yes
ppvt121	394	10	79	7	5	0	-0.08	0.78	1.01	1.01	0.64		___	Yes
ppvt122	394	7	10	12	72	0	0.35	0.77	0.95	0.93	0.93		___	Yes
ppvt123	394	15	12	52	20	0	1.28	0.65	0.97	0.93	0.42		___	Yes
ppvt124	394	17	43	27	12	0	1.66	0.58	0.98	0.98	1.32		___	Yes
ppvt125	394	14	28	39	18	0	1.85	0.54	1.02	1.03	0.00		___	Yes
ppvt126	394	50	30	10	9	0	1.37	0.63	0.97	0.95	-2.29	*	__D	Yes
ppvt127	394	2	89	9	1	0	-0.91	0.81	1.10	1.24	-0.61		___	Yes
ppvt128	394	84	4	4	7	0	-0.44	0.83	0.94	0.83	-0.33		___	Yes
ppvt129	394	22	15	24	40	0	1.84	0.58	0.91	0.88	0.25		___	Yes
ppvt130	394	82	4	3	10	0	-0.33	0.84	0.89	0.76	-1.08		___	Yes
ppvt131	394	22	15	27	36	0	2.02	0.52	1.03	1.04	0.26		___	Yes
ppvt132	394	2	7	88	3	0	-0.78	0.83	1.03	1.23	0.58		___	Yes
ppvt133	349	13	28	46	13	0	1.62	0.59	0.98	0.98	0.00		___	Yes
ppvt134	349	7	76	4	13	0	0.18	0.78	0.93	0.84	-0.28		___	Yes
ppvt135	349	53	17	22	7	0	1.31	0.66	0.96	0.94	0.47		___	Yes
ppvt136	349	3	3	3	91	0	-1.07	0.84	0.98	0.85	0.52		___	Yes
ppvt137	349	54	5	32	9	0	1.27	0.65	0.95	0.93	1.91		___	Yes
ppvt138	349	47	30	16	7	0	1.59	0.60	0.97	0.96	0.67		___	Yes
ppvt139	349	1	2	1	96	0	-1.90	0.86	1.02	0.94	-0.20		___	Yes
ppvt140	349	4	16	34	46	0	1.64	0.57	1.07	1.13	1.17		___	Yes
ppvt141	349	11	69	11	8	0	0.57	0.74	0.92	0.84	-0.63		___	Yes
ppvt142	349	30	12	52	5	0	1.35	0.66	0.91	0.90	0.38		___	Yes
ppvt143	349	6	91	2	1	0	-1.00	0.85	0.97	0.86	0.15		___	Yes
ppvt144	349	4	7	71	17	0	0.47	0.73	1.03	1.06	0.10		___	Yes
ppvt145	323	0	3	1	95	0	-1.69	0.85	1.02	1.01	-0.96		___	Yes
ppvt146	323	28	21	44	7	0	1.77	0.55	1.12	1.17	1.25		___	Yes
ppvt147	323	2	94	2	2	0	-1.38	0.85	0.98	0.80	0.00		___	Yes
ppvt148	323	47	6	34	13	0	2.26	0.52	0.92	0.91	-1.07		___	Yes
ppvt149	323	76	9	5	10	0	0.22	0.78	0.96	0.94	0.85		___	Yes
ppvt150	323	13	28	24	34	1	2.79	0.43	0.99	1.05	-0.69		___	Yes
ppvt151	323	11	9	9	70	0	0.58	0.74	0.97	0.93	-0.90		___	Yes
ppvt152	323	34	47	6	14	0	2.26	0.49	1.03	1.06	0.67		___	Yes
ppvt153	323	18	41	25	16	0	1.92	0.56	1.01	1.01	-0.42		___	Yes
ppvt154	323	15	38	24	23	0	2.07	0.53	0.96	0.96	0.65		___	Yes
ppvt155	323	89	6	2	4	0	-0.69	0.83	1.00	1.06	1.81		___	Yes
ppvt156	323	11	4	9	76	0	0.24	0.78	0.96	0.92	-0.63		___	Yes
ppvt157	290	2	91	6	1	0	-0.94	0.82	0.98	1.03	-0.88		___	Yes
ppvt158	290	16	27	15	42	0	3.41	0.35	0.96	0.94	0.00		___	Yes
ppvt159	290	16	36	18	31	0	2.49	0.48	0.92	0.97	0.00		___	Yes
ppvt160	290	6	7	80	7	0	0.05	0.76	1.05	1.15	-0.14		___	Yes
ppvt161	290	16	3	4	77	0	0.25	0.75	1.01	1.11	0.08		___	Yes
ppvt162	290	79	2	2	18	0	0.16	0.76	1.02	1.06	-2.56	*	__D	Yes
ppvt163	290	9	55	30	6	0	1.36	0.64	0.96	0.96	-0.24		___	Yes
ppvt164	290	70	11	4	16	0	0.67	0.70	1.06	1.07	-1.99	*	__D	Yes
ppvt165	290	9	9	9	74	0	0.45	0.72	1.05	1.03	-0.30		___	Yes
ppvt166	290	4	92	2	2	0	-1.04	0.82	1.03	1.23	0.00		___	Yes
ppvt167	290	6	4	86	4	0	-0.36	0.80	0.97	0.85	1.29		___	Yes
ppvt168	290	6	6	64	23	1	0.96	0.69	0.99	1.03	-0.56		___	Yes
ppvt169	277	2	2	3	94	0	-1.29	0.82	1.00	1.46	-0.93		___	Yes
ppvt170	277	5	84	5	5	0	-0.19	0.78	1.01	1.25	-0.47		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
											Male-Female			
ppvt171	277	68	12	11	9	0	0.79	0.69	1.06	1.10	1.78		___	Yes
ppvt172	277	52	10	29	10	0	2.63	0.43	1.08	1.11	2.04	*	___D	Yes
ppvt173	277	41	14	14	30	1	2.53	0.45	1.01	1.03	1.27		___	Yes
ppvt174	277	4	89	3	4	0	-0.65	0.81	0.98	0.84	-1.13		___	Yes
ppvt175	277	2	3	92	3	0	-1.06	0.81	1.04	1.13	-1.67		___	Yes
ppvt176	277	13	55	9	23	0	1.39	0.61	1.07	1.15	-2.03	*	___D	Yes
ppvt177	277	39	11	23	27	0	2.10	0.52	1.01	1.02	-1.66		___	Yes
ppvt178	277	2	8	88	3	0	-0.54	0.79	1.05	1.09	-0.06		___	Yes
ppvt179	277	4	3	8	85	0	-0.25	0.78	1.01	1.00	-1.86		___	Yes
ppvt180	277	28	13	53	6	0	2.65	0.45	0.94	0.93	0.00		___	Yes
ppvt181	265	10	17	7	66	0	0.92	0.67	1.01	1.02	1.15		___	Yes
ppvt182	265	57	15	14	14	0	1.34	0.62	1.00	0.99	1.55		___	Yes
ppvt183	265	2	5	89	4	0	-0.56	0.78	1.05	1.16	0.56		___	Yes
ppvt184	265	5	3	6	85	0	-0.25	0.77	1.02	0.94	1.08		___	Yes
ppvt185	265	26	29	11	34	0	2.77	0.41	1.12	1.11	0.09		___	Yes
ppvt186	265	3	4	91	2	0	-0.77	0.80	0.99	0.95	0.35		___	Yes
ppvt187	265	5	92	1	2	0	-1.02	0.80	1.02	1.32	-0.35		___	Yes
ppvt188	265	14	28	25	34	0	2.38	0.47	1.00	1.02	-0.73		___	Yes
ppvt189	265	13	55	23	9	0	1.40	0.61	1.00	0.97	0.28		___	Yes
ppvt190	265	5	2	91	2	0	-0.82	0.80	0.99	0.99	0.14		___	Yes
ppvt191	265	70	22	2	7	0	0.72	0.70	0.97	0.91	-0.42		___	Yes
ppvt192	265	30	55	10	4	0	1.39	0.62	0.95	0.90	-1.15		___	Yes
ppvt193	254	41	16	7	36	0	2.27	0.49	0.97	0.98	-0.97		___	Yes
ppvt194	254	9	79	7	5	0	0.20	0.72	1.03	1.04	-1.11		___	Yes
ppvt195	254	9	4	83	5	0	-0.04	0.75	0.95	0.88	2.59	*	___D	Yes
ppvt196	254	76	2	5	17	0	0.39	0.71	1.04	1.06	1.41		___	Yes
ppvt197	254	69	7	7	16	0	0.75	0.68	0.96	0.98	1.31		___	Yes
ppvt198	254	0	6	2	93	0	-1.02	0.79	0.98	0.83	-0.11		___	Yes
ppvt199	254	0	10	79	11	0	0.23	0.73	0.99	0.90	0.32		___	Yes
ppvt200	254	70	3	2	26	0	0.73	0.68	0.99	0.95	1.17		___	Yes
ppvt201	254	19	26	3	52	0	1.54	0.57	1.08	1.06	-2.49	*	___D	Yes
ppvt202	254	3	96	1	0	0	-1.71	0.80	1.01	0.97	1.34		___	Yes
ppvt203	254	3	41	52	4	0	1.54	0.60	0.93	0.90	-0.39		___	Yes
ppvt204	254	26	51	3	19	0	1.59	0.60	0.88	0.86	-0.49		___	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 13. Item statistics by item in the Peabody Picture Vocabulary Test for the Younger Cohort, India – Telugu

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt001	1002	1	0	1	98	0	-4.89	0.11	0.95	0.78	0.90		___	Yes
ppvt002	1002	1	0	97	1	0	-4.77	0.12	0.94	0.64	0.00		___	Yes
ppvt003	1002	98	0	1	0	0	-5.04	0.12	0.93	0.48	0.95		__C_	Yes
ppvt004	1002	96	1	1	1	0	-4.39	0.13	0.96	0.70	0.11		___	Yes
ppvt005	1002	2	1	32	66	0	-1.62	0.31	1.01	1.04	0.19		___	Yes
ppvt006	1002	98	1	0	1	0	-4.81	0.12	0.92	0.61	-0.27		___	Yes
ppvt007	1002	5	92	1	3	0	-3.51	0.16	0.95	0.92	-0.25		___	Yes
ppvt008	1002	25	9	59	7	0	-1.30	0.37	0.98	0.98	-0.22		___	Yes
ppvt009	1002	3	2	5	89	0	-3.20	0.21	0.92	0.83	-0.23		___	Yes
ppvt010	1002	2	75	13	10	0	-2.13	0.28	1.00	0.97	0.39	*	__D	Yes
ppvt011	1002	2	93	3	2	0	-3.67	0.19	0.89	0.67	0.00		___	Yes
ppvt012	1002	2	6	90	2	0	-3.28	0.22	0.89	0.71	-0.10		___	Yes
ppvt013	1180	15	74	5	6	0	-1.98	0.32	0.91	0.88	-0.13		___	Yes
ppvt014	1180	88	3	5	5	0	-2.98	0.23	0.93	0.82	0.48	*	__D	Yes
ppvt015	1180	4	4	87	4	0	-2.90	0.22	0.94	0.89	0.15		___	Yes
ppvt016	1180	86	7	4	3	0	-2.79	0.25	0.94	0.83	0.12		___	Yes
ppvt017	1180	5	5	87	2	0	-2.90	0.25	0.91	0.77	0.07		___	Yes
ppvt018	1180	23	54	12	11	0	-1.00	0.40	0.96	0.95	0.19		___	Yes
ppvt019	1180	65	4	21	10	0	-1.47	0.34	1.00	1.00	0.00		___	Yes
ppvt020	1180	27	17	14	43	0	-0.51	0.47	0.96	0.97	-0.04		___	Yes
ppvt021	1180	21	20	43	16	0	-0.48	0.40	1.05	1.06	0.08		___	Yes
ppvt022	1180	24	2	10	64	0	-1.44	0.35	0.96	0.96	-0.19		___	Yes
ppvt023	1180	22	59	9	10	0	-1.19	0.38	0.97	0.96	-0.09		___	Yes
ppvt024	1180	23	10	15	52	0	-0.88	0.38	1.02	1.03	0.31	*	__D	Yes
ppvt025	1417	93	3	1	3	0	-3.53	0.22	0.88	0.56	-0.27		___	Yes
ppvt026	1417	3	2	89	5	0	-2.99	0.26	0.86	0.65	-0.15		___	Yes
ppvt027	1417	12	18	47	23	0	0.68	0.53	0.96	1.02	-0.11		___	Yes
ppvt028	1417	7	4	6	83	0	-2.39	0.28	0.91	0.81	-0.10		___	Yes
ppvt029	1417	68	5	5	22	0	-1.51	0.28	1.03	1.05	-0.10		___	Yes
ppvt030	1417	2	64	33	1	0	-1.30	0.22	1.13	1.19	-0.05		___	Yes
ppvt031	1417	5	20	62	13	0	-1.18	0.39	0.94	0.92	-0.05		___	Yes
ppvt032	1417	14	64	8	14	0	-1.31	0.31	1.02	1.02	-0.31	*	__D	Yes
ppvt033	1417	6	76	7	10	0	-1.96	0.35	0.87	0.77	0.23		___	Yes
ppvt034	1417	55	14	20	12	0	-0.87	0.39	0.99	0.99	0.21		___	Yes
ppvt035	1417	14	18	16	52	0	-0.72	0.37	1.02	1.04	-0.44	*	__D	Yes
ppvt036	1417	3	15	79	2	0	-2.14	0.29	0.94	0.90	0.00		___	Yes
ppvt037	1434	20	48	18	14	0	-0.53	0.33	1.08	1.11	0.00		___	Yes
ppvt038	1434	3	4	2	91	0	-3.12	0.22	0.93	0.79	0.61	*	__D	Yes
ppvt039	1434	9	41	44	6	0	-0.34	0.39	1.02	1.04	-0.14		___	Yes
ppvt040	1434	29	13	13	45	0	0.35	0.46	1.01	1.17	0.11		___	Yes
ppvt041	1434	18	62	6	14	0	-1.20	0.36	0.97	0.96	-0.09		___	Yes
ppvt042	1434	13	29	16	41	0	1.42	0.39	1.13	1.65	-0.10		__C_	Yes
ppvt043	1434	26	21	33	20	0	0.18	0.40	1.05	1.14	0.08		___	Yes
ppvt044	1434	18	37	11	34	0	0.11	0.33	1.12	1.21	0.16		___	Yes
ppvt045	1434	25	41	10	23	0	0.71	0.40	1.09	1.30	-0.16		___	Yes
ppvt046	1434	10	30	51	9	0	0.33	0.45	1.00	1.09	0.00		___	Yes
ppvt047	1434	7	4	86	3	0	-2.64	0.30	0.86	0.68	-0.19		___	Yes
ppvt048	1434	86	6	5	4	0	-2.59	0.24	0.97	0.93	0.28		___	Yes
ppvt049	1388	34	29	25	11	0	0.62	0.45	1.05	1.11	-0.16		___	Yes
ppvt050	1388	27	57	10	6	0	0.52	0.46	1.05	1.22	0.00		___	Yes
ppvt051	1388	23	33	30	13	0	0.76	0.49	1.01	1.24	-0.12		___	Yes
ppvt052	1388	27	55	7	11	0	-0.82	0.47	0.93	0.93	0.03		___	Yes
ppvt053	1388	42	15	24	19	0	1.01	0.46	1.05	1.25	0.22		___	Yes
ppvt054	1388	13	47	21	19	0	-0.47	0.36	1.10	1.12	0.12		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt055	1388	55	32	8	5	0	-0.83	0.43	1.01	1.01	0.00		___	Yes
ppvt056	1388	27	25	11	36	0	0.04	0.51	0.96	0.98	0.30	*	___D	Yes
ppvt057	1388	28	39	18	15	0	-0.10	0.35	1.13	1.20	0.26	*	___D	Yes
ppvt058	1388	25	19	29	27	0	0.40	0.44	1.05	1.10	0.06		___	Yes
ppvt059	1388	9	13	71	7	0	-1.59	0.40	0.95	0.90	-0.14		___	Yes
ppvt060	1388	17	12	18	53	0	-0.74	0.43	0.99	0.98	-0.24	*	___D	Yes
ppvt061	1070	25	26	14	35	0	0.25	0.60	0.94	1.13	0.03		___	Yes
ppvt062	1070	30	44	14	12	0	0.54	0.52	1.03	1.14	0.00		___	Yes
ppvt063	1070	11	64	11	14	0	-1.10	0.51	1.07	1.08	-0.24		___	Yes
ppvt064	1070	7	13	73	6	0	-1.61	0.59	0.87	0.78	0.00		___	Yes
ppvt065	1070	12	18	40	30	0	0.53	0.47	1.10	1.22	0.12		___	Yes
ppvt066	1070	19	24	26	32	0	0.78	0.38	1.19	1.47	-0.30	*	___D	Yes
ppvt067	1070	19	38	20	23	0	0.12	0.46	1.14	1.26	-0.05		___	Yes
ppvt068	1070	64	19	9	8	0	-1.12	0.61	0.88	0.83	0.50	*	___D	Yes
ppvt069	1070	12	5	12	70	0	-1.45	0.54	0.96	0.94	-0.20		___	Yes
ppvt070	1070	17	61	11	10	0	-0.96	0.57	0.99	0.97	0.00		___	Yes
ppvt071	1070	9	11	71	9	0	-1.49	0.56	0.92	0.91	0.15		___	Yes
ppvt072	1070	29	27	18	26	0	0.60	0.51	1.04	1.32	-0.32	*	___D	Yes
ppvt073	910	13	44	30	13	0	0.15	0.51	1.04	1.10	0.16		___	Yes
ppvt074	910	6	5	4	85	0	-2.12	0.67	0.84	0.67	0.00		___	Yes
ppvt075	910	30	8	56	6	0	-0.41	0.59	0.99	1.01	-0.37	*	___D	Yes
ppvt076	910	38	18	31	13	0	0.44	0.53	0.99	1.04	0.30	*	___D	Yes
ppvt077	910	64	15	13	7	0	-0.83	0.64	0.92	0.91	0.40	*	___D	Yes
ppvt078	910	8	9	77	6	0	-1.59	0.67	0.88	0.82	-0.06		___	Yes
ppvt079	910	10	56	25	9	0	-0.42	0.58	1.01	1.03	-0.13		___	Yes
ppvt080	910	42	15	29	13	0	0.22	0.59	0.93	0.92	0.06		___	Yes
ppvt081	910	27	42	15	17	0	0.26	0.52	1.04	1.14	-0.13		___	Yes
ppvt082	910	23	12	13	52	0	-0.23	0.63	0.93	0.93	0.00		___	Yes
ppvt083	910	10	20	49	21	0	-0.11	0.59	0.97	0.99	0.00		___	Yes
ppvt084	910	16	29	15	40	0	0.32	0.47	1.10	1.25	0.64	*	___D	Yes
ppvt085	599	23	40	28	10	0	0.75	0.50	1.08	1.09	0.00		___	Yes
ppvt086	599	39	12	31	19	0	1.88	0.43	0.97	0.96	0.30		___	Yes
ppvt087	599	65	8	16	10	0	-0.35	0.66	1.06	1.06	0.03		___	Yes
ppvt088	599	14	19	15	51	0	0.26	0.60	1.03	1.04	0.00		___	Yes
ppvt089	599	23	33	25	19	0	1.49	0.40	1.05	1.08	0.04		___	Yes
ppvt090	599	45	22	12	21	0	0.54	0.59	0.99	1.00	-0.28		___	Yes
ppvt091	599	13	13	18	56	0	0.08	0.60	1.05	1.08	-0.13		___	Yes
ppvt092	599	7	42	46	6	0	0.67	0.51	1.06	1.07	-0.35	*	___D	Yes
ppvt093	599	2	3	94	1	0	-2.50	0.79	0.97	1.01	-1.01	*	___D	Yes
ppvt094	599	24	55	16	5	0	0.09	0.58	1.10	1.11	-0.18		___	Yes
ppvt095	599	6	16	47	32	0	0.47	0.60	1.00	1.01	-0.12		___	Yes
ppvt096	599	61	16	6	17	0	-0.16	0.69	0.97	0.96	-0.05		___	Yes
ppvt097	470	33	28	16	23	0	1.47	0.47	1.01	1.02	-0.75	*	___D	Yes
ppvt098	470	18	47	8	27	0	0.59	0.57	1.07	1.07	0.02		___	Yes
ppvt099	470	2	12	16	70	0	-0.45	0.70	1.01	1.04	-0.29		___	Yes
ppvt100	470	36	8	25	30	0	1.61	0.54	0.89	0.86	0.29		___	Yes
ppvt101	470	6	17	7	70	0	-0.46	0.75	0.96	0.92	-0.44	*	___D	Yes
ppvt102	470	19	14	11	56	0	1.96	0.34	1.08	1.15	-0.18		___	Yes
ppvt103	470	20	27	25	27	0	1.62	0.48	0.95	0.96	0.65	*	___D	Yes
ppvt104	470	2	4	2	92	0	-2.08	0.79	1.02	1.05	0.37		___	Yes
ppvt105	470	63	18	8	11	0	-0.14	0.69	0.99	1.00	-0.02		___	Yes
ppvt106	470	14	46	29	11	0	0.60	0.61	0.98	0.98	0.12		___	Yes
ppvt107	470	37	14	38	12	0	1.02	0.52	1.04	1.05	0.00		___	Yes
ppvt108	470	5	28	65	2	0	-0.22	0.70	1.01	1.01	0.36		___	Yes
ppvt109	359	1	2	4	94	0	-2.24	0.80	0.98	0.82	0.00		___	Yes
ppvt110	359	14	83	1	2	0	-1.09	0.77	0.99	1.01	0.25		___	Yes
ppvt111	359	7	40	40	13	0	1.01	0.58	0.95	0.95	0.00		___	Yes
ppvt112	359	9	24	38	28	0	1.07	0.54	0.99	1.00	0.13		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt113	359	29	16	39	16	0	1.53	0.44	1.07	1.12	0.02		___	Yes
ppvt114	359	39	11	12	38	0	1.02	0.58	0.95	0.95	0.26		___	Yes
ppvt115	359	22	1	1	77	0	-0.70	0.73	1.01	1.02	-0.19		___	Yes
ppvt116	359	8	54	14	24	0	0.40	0.61	1.05	1.06	0.29		___	Yes
ppvt117	359	20	39	29	12	0	2.02	0.39	1.03	1.03	0.47		___	Yes
ppvt118	359	8	51	18	22	0	0.50	0.68	0.92	0.91	0.60	*	___D	Yes
ppvt119	359	10	32	12	46	0	0.73	0.63	0.93	0.93	-0.19		___	Yes
ppvt120	359	15	28	27	30	0	1.60	0.49	0.97	0.98	-0.10		___	Yes
ppvt121	259	14	55	22	9	0	0.49	0.63	0.97	0.97	-0.20		___	Yes
ppvt122	259	11	23	16	50	0	0.70	0.61	0.98	0.98	0.09		___	Yes
ppvt123	259	18	14	54	14	0	0.54	0.61	1.02	1.02	0.05		___	Yes
ppvt124	259	33	17	31	19	0	2.38	0.38	0.96	0.93	0.08		___	Yes
ppvt125	259	8	24	46	22	0	0.87	0.60	0.96	0.96	0.32		___	Yes
ppvt126	259	93	3	3	2	0	-1.93	0.79	0.97	1.05	-0.65		___	Yes
ppvt127	259	30	29	14	27	0	1.68	0.50	0.90	0.88	0.58		___	Yes
ppvt128	259	97	1	1	1	0	-2.99	0.80	0.99	0.94	0.14		___	Yes
ppvt129	259	15	3	31	51	0	0.67	0.59	1.03	1.03	-0.85	*	___D	Yes
ppvt130	259	32	31	12	24	0	1.49	0.46	1.07	1.09	-0.14		___	Yes
ppvt131	259	14	8	15	63	0	0.13	0.64	1.05	1.07	-0.07		___	Yes
ppvt132	259	3	6	88	3	0	-1.38	0.78	0.97	0.95	0.00		___	Yes
ppvt133	225	0	6	93	1	0	-1.89	0.77	1.00	1.03	-0.13		___	Yes
ppvt134	225	46	17	8	28	0	2.44	0.40	0.89	0.78	-0.47		___	Yes
ppvt135	225	25	11	60	4	0	1.91	0.42	1.03	1.01	-0.05		___	Yes
ppvt136	225	3	15	5	77	0	-0.48	0.71	1.00	0.99	0.36		___	Yes
ppvt137	225	89	4	5	2	0	-1.43	0.75	1.01	1.03	0.07		___	Yes
ppvt138	225	17	47	18	19	0	2.44	0.34	1.01	1.07	-0.20		___	Yes
ppvt139	225	32	24	19	24	0	1.96	0.42	1.00	1.00	-0.55		___	Yes
ppvt140	225	0	12	35	52	0	0.67	0.59	1.00	1.00	0.18		___	Yes
ppvt141	225	15	52	14	19	0	0.69	0.63	0.92	0.91	-0.09		___	Yes
ppvt142	225	21	21	49	9	0	0.82	0.58	0.99	0.99	0.26		___	Yes
ppvt143	225	49	40	6	5	0	1.20	0.53	0.96	0.97	-0.27		___	Yes
ppvt144	225	11	18	32	40	0	1.59	0.52	0.88	0.85	0.15		___	Yes
ppvt145	164	3	14	4	79	0	-0.43	0.68	1.00	0.98	-0.24		___	Yes
ppvt146	164	16	37	35	12	0	1.55	0.45	1.04	1.05	0.04		___	Yes
ppvt147	164	5	92	1	1	0	-1.60	0.73	1.01	0.98	-3.48		___	Yes
ppvt148	164	31	22	21	26	0	2.32	0.37	0.97	0.97	-0.83	*	___D	Yes
ppvt149	164	57	36	5	2	0	0.62	0.56	1.08	1.09	0.04		___	Yes
ppvt150	164	15	33	34	18	0	1.61	0.45	1.03	1.04	-0.88	*	___D	Yes
ppvt151	164	7	34	9	50	0	0.92	0.54	1.03	1.03	-0.10		___	Yes
ppvt152	164	27	29	15	29	0	1.97	0.39	1.05	1.07	0.34		___	Yes
ppvt153	164	5	85	7	2	0	-0.90	0.71	0.97	0.92	0.55		___	Yes
ppvt154	164	3	63	4	30	0	0.38	0.62	1.00	1.00	0.00		___	Yes
ppvt155	164	41	41	5	12	0	1.28	0.53	0.91	0.91	0.58		___	Yes
ppvt156	164	31	16	11	42	0	1.26	0.51	0.98	0.98	-0.19		___	Yes
ppvt157	131	6	77	10	7	0	-0.28	0.61	1.10	1.24	-0.40		___	Yes
ppvt158	131	18	23	20	39	0	2.53	0.31	1.02	1.04	-0.08		___	Yes
ppvt159	131	17	31	24	28	0	1.95	0.37	1.06	1.14	-0.09		___	Yes
ppvt160	131	31	21	24	24	0	2.20	0.35	1.01	1.07	-0.33		___	Yes
ppvt161	131	30	5	5	60	0	0.55	0.56	0.99	0.98	0.18		___	Yes
ppvt162	131	50	8	8	34	0	0.97	0.50	1.05	1.06	-1.05	*	___D	Yes
ppvt163	131	36	34	11	18	0	1.66	0.44	0.94	0.93	0.53		___	Yes
ppvt164	131	69	11	3	17	0	0.13	0.57	1.11	1.17	-0.16		___	Yes
ppvt165	131	15	16	16	53	0	0.84	0.53	0.99	0.99	-0.56		___	Yes
ppvt166	131	9	73	8	10	0	-0.03	0.62	1.00	0.98	0.26		___	Yes
ppvt167	131	2	1	97	0	0	-2.55	0.70	1.01	0.99	-0.76		___	Yes
ppvt168	131	24	21	27	29	0	2.03	0.39	1.00	0.99	0.30		___	Yes
ppvt169	104	16	26	23	35	0	1.67	0.39	1.02	1.03	-0.17		___	Yes
ppvt170	104	13	39	38	10	0	1.45	0.44	0.91	0.89	0.19		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
		Male-Female												
ppvt171	104	47	32	4	17	0	1.13	0.44	1.07	1.07	-0.11		_____	Yes
ppvt172	104	29	35	18	18	0	2.56	0.26	1.05	1.24	0.37		_____	Yes
ppvt173	104	15	34	38	13	0	3.02	0.26	0.95	0.90	0.83		_____	Yes
ppvt174	104	22	23	45	10	0	2.26	0.35	0.90	0.83	0.30		_____	Yes
ppvt175	104	14	21	41	23	0	1.37	0.42	1.01	1.01	0.00		_____	Yes
ppvt176	104	37	31	13	19	0	1.85	0.36	1.03	1.04	-0.30		_____	Yes
ppvt177	104	36	7	28	30	0	1.63	0.37	1.12	1.16	-0.46		_____	Yes
ppvt178	104	20	25	36	19	0	1.63	0.43	0.91	0.90	-0.28		_____	Yes
ppvt179	104	11	10	12	68	0	0.21	0.54	0.99	1.00	-0.47		_____	Yes
ppvt180	104	33	23	34	11	0	1.76	0.40	0.97	0.94	0.04		_____	Yes
ppvt181	39	5	3	38	54	0	1.06	0.34	0.98	0.97	-0.13		_____	Yes
ppvt182	39	13	77	8	3	0	3.21	0.18	0.98	0.86	0.00		_____	Yes
ppvt183	39	10	15	62	13	0	0.73	0.35	1.02	1.01	0.95		_____	Yes
ppvt184	39	3	0	5	92	0	-1.34	0.42	1.06	1.25	0.84		_____	Yes
ppvt185	39	36	10	41	13	0	1.83	0.26	1.11	1.11	-0.62		_____	Yes
ppvt186	39	3	3	92	3	0	-1.34	0.43	0.99	0.83	-0.67		_____	Yes
ppvt187	39	54	21	10	15	0	2.63	0.22	0.96	0.91	0.87		_____	Yes
ppvt188	39	21	46	13	21	0	2.63	0.20	1.00	1.15	1.86		_____	Yes
ppvt189	39	13	44	10	33	0	1.49	0.31	0.97	0.98	0.00		_____	Yes
ppvt190	39	3	49	31	18	0	2.07	0.27	0.90	0.88	1.02		_____	Yes
ppvt191	39	49	21	10	21	0	1.28	0.34	0.87	0.86	0.37		_____	Yes
ppvt192	39	23	49	10	18	0	1.28	0.33	0.95	0.94	0.82		_____	Yes
ppvt193	25	48	16	20	16	0	3.09	0.15	1.06	1.20	-0.88		_____	Yes
ppvt194	25	8	28	12	52	0	2.36	0.21	1.02	1.05	-0.59		_____	Yes
ppvt195	25	8	12	80	0	0	-0.06	0.34	1.06	1.22	1.14		_____	Yes
ppvt196	25	56	12	4	28	0	1.12	0.29	1.04	1.04	-0.13		_____	Yes
ppvt197	25	56	4	24	16	0	1.12	0.29	1.06	1.08	-1.71		_____	Yes
ppvt198	25	16	20	12	52	0	1.29	0.30	0.80	0.79	0.33		_____	Yes
ppvt199	25	12	16	48	24	0	1.46	0.29	0.85	0.84	0.05		_____	Yes
ppvt200	25	36	12	28	24	0	1.97	0.22	1.09	1.11	1.97		_____	Yes
ppvt201	25	28	32	0	40	0	1.80	0.25	0.99	1.00	-0.46		_____	Yes
ppvt202	25	12	80	8	0	0	-0.06	0.34	1.06	1.08	-1.15		_____	Yes
ppvt203	25	8	32	44	16	0	1.63	0.27	0.89	0.87	-0.19		_____	Yes
ppvt204	25	20	32	32	16	0	2.16	0.23	0.96	0.95	-0.27		_____	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 14. Item statistics by item in the Peabody Picture Vocabulary Test for the Younger Cohort, Peru – Spanish

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt001	275	7	81	7	5	0	-5.64	0.34	1.00	0.95	0.54		___	Yes
ppvt002	277	16	40	6	37	0	-3.13	0.53	1.41	1.62	0.23		__C__	Yes
ppvt003	278	95	3	2	0	0	-7.34	0.18	1.08	1.48	-0.70		___	Yes
ppvt004	316	3	92	0	4	0	-6.78	0.24	0.97	0.80	-0.43		___	Yes
ppvt005	321	46	28	17	9	0	-3.37	0.54	1.19	1.20	0.55	*	___D	Yes
ppvt006	324	8	8	5	79	0	-5.27	0.36	1.12	1.19	0.40		___	Yes
ppvt007	329	55	15	15	15	1	-3.81	0.51	1.12	1.14	0.00		___	Yes
ppvt008	342	50	19	12	19	0	-3.54	0.55	1.11	1.10	0.34		___	Yes
ppvt009	331	10	6	8	76	0	-4.81	0.45	0.85	0.79	0.00		___	Yes
ppvt010	346	10	79	5	5	0	-4.89	0.44	0.93	0.87	0.00		___	Yes
ppvt011	348	3	3	86	7	1	-5.49	0.40	0.89	0.71	-0.96	*	___D	Yes
ppvt012	375	15	13	36	35	1	-2.45	0.64	1.14	1.21	-0.07		___	Yes
ppvt013	387	2	9	82	7	0	-4.82	0.45	0.99	0.94	0.03		___	Yes
ppvt014	390	15	5	75	5	0	-4.30	0.52	0.88	0.82	0.02		___	Yes
ppvt015	393	11	74	9	6	0	-4.23	0.54	0.90	0.76	-0.20		___	Yes
ppvt016	398	12	62	18	7	1	-3.49	0.58	1.05	1.17	0.06		___	Yes
ppvt017	410	76	9	7	9	0	-4.17	0.55	0.89	0.78	0.00		___	Yes
ppvt018	419	54	20	9	17	0	-2.92	0.64	0.96	0.93	-0.17		___	Yes
ppvt019	412	8	75	4	13	0	-3.99	0.57	0.93	0.98	0.00		___	Yes
ppvt020	425	22	46	25	7	0	-2.36	0.65	1.17	1.20	0.39		___	Yes
ppvt021	429	10	11	16	62	0	-3.15	0.61	1.09	1.16	0.00		___	Yes
ppvt022	491	6	9	78	7	0	-3.76	0.60	0.95	1.09	-0.30		___	Yes
ppvt023	497	12	77	5	6	0	-3.63	0.63	0.90	1.04	0.39		___	Yes
ppvt024	519	13	8	8	72	0	-3.19	0.66	0.89	0.85	-0.23		___	Yes
ppvt025	545	8	12	67	13	0	-2.82	0.68	0.92	0.89	0.00		___	Yes
ppvt026	532	2	7	76	16	0	-3.30	0.66	0.93	0.97	-0.21		___	Yes
ppvt027	605	7	6	81	6	0	-3.48	0.67	0.87	0.71	-0.25		___	Yes
ppvt028	632	5	6	9	80	0	-3.30	0.66	0.98	1.09	-0.33		___	Yes
ppvt029	675	6	89	1	4	0	-3.98	0.64	1.00	1.25	0.00		___	Yes
ppvt030	730	9	12	19	59	1	-1.86	0.72	0.97	0.97	-0.05		___	Yes
ppvt031	741	7	9	77	6	0	-2.82	0.70	0.90	0.84	-0.40	*	___D	Yes
ppvt032	746	15	79	4	2	0	-2.89	0.66	1.14	1.30	0.53	*	___D	Yes
ppvt033	758	19	6	7	67	0	-2.15	0.72	0.90	0.93	-0.63	*	___D	Yes
ppvt034	803	0	1	0	99	0	-5.88	0.66	0.97	0.46	0.99		__C__	Yes
ppvt035	812	57	13	13	16	1	-1.46	0.73	0.98	0.96	0.00		___	Yes
ppvt036	840	6	9	18	66	0	-1.91	0.70	1.05	1.11	0.13		___	Yes
ppvt037	878	62	9	9	19	1	-1.56	0.70	1.08	1.11	0.00		___	Yes
ppvt038	886	6	56	24	13	1	-1.20	0.70	1.10	1.15	0.29		___	Yes
ppvt039	898	9	16	68	6	1	-1.83	0.75	0.88	0.78	-0.04		___	Yes
ppvt040	901	13	10	60	16	1	-1.31	0.74	0.96	0.93	-0.18		___	Yes
ppvt041	910	1	4	81	14	0	-2.44	0.73	1.00	0.99	0.00		___	Yes
ppvt042	943	36	52	5	7	0	-0.66	0.72	1.03	1.02	-0.18		___	Yes
ppvt043	1054	77	5	7	10	0	-1.81	0.76	0.90	0.79	-0.23		___	Yes
ppvt044	1067	1	62	14	23	0	-0.84	0.73	0.93	0.90	0.06		___	Yes
ppvt045	1063	58	11	19	10	0	-0.60	0.73	0.95	0.94	0.09		___	Yes
ppvt046	1108	1	11	4	83	0	-2.00	0.74	1.02	1.14	0.47	*	___D	Yes
ppvt047	1138	5	82	4	9	0	-1.82	0.77	0.88	0.75	0.62	*	___D	Yes
ppvt048	1172	9	75	14	2	0	-1.26	0.75	0.93	0.90	0.29	*	___D	Yes
ppvt049	1192	81	7	3	10	0	-1.55	0.73	1.08	1.06	-0.18		___	Yes
ppvt050	1307	11	7	76	6	0	-1.16	0.73	0.96	0.92	0.38	*	___D	Yes
ppvt051	1304	6	23	17	53	1	0.12	0.62	1.09	1.13	0.05		___	Yes
ppvt052	1294	4	83	11	2	0	-1.57	0.73	1.05	1.12	0.34	*	___D	Yes
ppvt053	1381	3	10	80	6	0	-1.26	0.76	0.89	0.78	-0.07		___	Yes
ppvt054	1406	3	20	68	8	0	-0.47	0.66	1.05	1.08	-0.23		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender Male-Female		
ppvt055	1399	6	6	17	72	0	-0.66	0.71	0.92	0.89	-0.43	*	___D	Yes
ppvt056	1382	12	16	10	61	0	-0.05	0.67	0.92	0.89	-0.25	*	___D	Yes
ppvt057	1368	68	21	8	2	1	-0.41	0.63	1.14	1.22	0.00		___	Yes
ppvt058	1341	26	27	36	11	0	1.19	0.56	1.00	1.01	0.00		___	Yes
ppvt059	1329	85	6	3	7	0	-1.41	0.75	1.02	1.00	-0.13		___	Yes
ppvt060	1341	78	10	6	6	0	-0.85	0.75	0.86	0.73	-0.13		___	Yes
ppvt061	1322	12	36	26	25	1	1.88	0.49	1.01	1.04	0.18		___	Yes
ppvt062	1285	29	30	16	24	1	1.69	0.46	1.14	1.30	0.16		___	Yes
ppvt063	1251	17	16	11	56	0	0.44	0.62	0.97	0.97	0.00		___	Yes
ppvt064	1206	18	73	4	4	0	-0.36	0.70	0.94	0.90	0.05		___	Yes
ppvt065	1192	7	10	82	2	0	-0.84	0.70	1.06	1.16	-0.13		___	Yes
ppvt066	1171	9	11	71	9	0	-0.17	0.66	1.04	1.03	0.22		___	Yes
ppvt067	1176	55	9	30	5	0	0.65	0.61	0.94	0.91	0.37	*	___D	Yes
ppvt068	1145	19	51	16	13	1	0.88	0.53	1.17	1.21	-0.20		___	Yes
ppvt069	1125	14	13	14	58	1	0.61	0.61	0.97	0.95	-0.37	*	___D	Yes
ppvt070	1118	13	60	18	8	0	0.55	0.58	1.08	1.10	-0.07		___	Yes
ppvt071	1109	8	32	43	17	0	1.37	0.55	0.88	0.85	-0.07		___	Yes
ppvt072	1092	42	24	17	16	1	1.41	0.52	0.98	0.97	0.00		___	Yes
ppvt073	1056	17	12	11	59	1	0.71	0.58	1.00	1.00	0.14		___	Yes
ppvt074	1005	51	19	15	14	1	1.10	0.55	0.99	0.99	-0.09		___	Yes
ppvt075	945	21	21	41	16	1	1.60	0.50	0.98	0.98	-0.06		___	Yes
ppvt076	882	21	59	8	12	0	0.87	0.55	1.05	1.10	-0.13		___	Yes
ppvt077	834	12	18	36	34	0	1.93	0.44	1.05	1.09	0.07		___	Yes
ppvt078	775	13	35	36	16	1	2.08	0.44	1.00	0.99	0.29		___	Yes
ppvt079	703	13	17	61	8	0	0.96	0.54	1.00	0.99	-0.06		___	Yes
ppvt080	676	18	9	20	52	1	1.38	0.50	1.01	1.02	-0.13		___	Yes
ppvt081	647	17	32	34	17	1	2.33	0.38	1.11	1.16	-0.40	*	___D	Yes
ppvt082	596	2	8	60	29	0	1.11	0.52	0.94	0.93	0.71	*	___D	Yes
ppvt083	555	20	24	17	38	1	2.90	0.35	0.97	0.95	0.00		___	Yes
ppvt084	517	39	19	14	28	1	2.19	0.41	1.01	1.03	-0.46	*	___D	Yes
ppvt085	468	14	26	45	14	2	1.99	0.41	1.08	1.09	0.30		___	Yes
ppvt086	436	16	49	25	10	0	1.85	0.45	0.90	0.89	0.27		___	Yes
ppvt087	413	15	39	13	32	0	2.32	0.40	0.94	0.94	0.20		___	Yes
ppvt088	362	7	55	12	25	1	1.68	0.44	0.91	0.90	0.10		___	Yes
ppvt089	327	12	25	28	34	2	2.71	0.34	1.04	1.06	-0.72	*	___D	Yes
ppvt090	298	9	5	62	24	0	3.29	0.28	1.04	1.06	-0.54		___	Yes
ppvt091	256	17	20	26	35	2	2.80	0.31	1.06	1.06	0.22		___	Yes
ppvt092	240	15	21	31	33	0	3.04	0.30	1.05	1.08	-0.47		___	Yes
ppvt093	215	13	21	38	28	0	3.24	0.27	1.05	1.05	0.14		___	Yes
ppvt094	195	45	5	22	29	0	3.28	0.28	0.96	0.92	0.63		___	Yes
ppvt095	161	8	17	25	50	0	2.38	0.33	0.93	0.91	0.30		___	Yes
ppvt096	140	24	31	30	13	1	4.52	0.19	0.96	0.95	0.08		___	Yes
ppvt097	105	30	43	23	3	1	2.86	0.27	1.02	1.01	0.08		___	Yes
ppvt098	90	29	27	10	32	2	3.71	0.22	0.91	0.91	-0.14		___	Yes
ppvt099	81	0	25	64	11	0	2.00	0.29	0.97	1.14	0.68		___	Yes
ppvt100	74	24	11	49	15	1	2.76	0.27	0.85	0.83	-0.48		___	Yes
ppvt101	66	26	29	36	8	2	3.95	0.19	1.10	1.16	1.28		___	Yes
ppvt102	59	51	17	14	19	0	4.46	0.17	1.08	1.15	0.23		___	Yes
ppvt103	55	27	33	13	27	0	3.97	0.21	0.83	0.81	0.00		___	Yes
ppvt104	39	54	26	10	10	0	2.90	0.23	1.00	1.07	-0.22		___	Yes
ppvt105	34	35	12	18	32	3	3.81	0.19	1.01	0.96	-0.35		___	Yes
ppvt106	31	26	39	10	23	3	3.73	0.20	0.89	0.86	0.14		___	Yes
ppvt107	28	14	64	18	4	0	2.63	0.23	0.95	0.90	1.57		___	Yes
ppvt108	26	19	19	46	12	4	3.50	0.20	0.84	0.81	0.94		___	Yes
ppvt109	21	5	14	67	14	0	2.75	0.21	1.20	1.27	-0.44		___	Yes
ppvt110	21	5	33	19	43	0	5.03	0.15	0.89	0.77	-1.05		___	Yes
ppvt111	19	5	0	5	89	0	1.29	0.23	0.91	0.62	0.91		___	Yes
ppvt112	20	40	35	5	20	0	4.29	0.17	0.92	0.85	-1.15		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male-Female						
ppvt113	20	45	30	10	15	0	3.83	0.18	0.86	0.84	-0.45		___	Yes
ppvt114	18	6	50	11	33	0	3.70	0.18	0.92	0.89	-0.64		___	Yes
ppvt115	17	12	35	41	12	0	4.12	0.17	0.88	0.83	-1.30		___	Yes
ppvt116	15	13	20	40	27	0	4.28	0.15	1.14	1.14	-1.63		___	Yes
ppvt117	14	0	43	36	14	7	4.11	0.15	1.17	1.19	-0.05		___	Yes
ppvt118	15	13	20	47	20	0	4.11	0.16	0.82	0.80	-0.05		___	Yes
ppvt119	15	80	7	0	13	0	2.44	0.20	0.84	0.71	-0.33		___	Yes
ppvt120	12	33	8	42	17	0	4.55	0.13	1.39	1.42	-1.21		___	Yes
ppvt121	11	18	64	18	0	0	3.54	0.16	0.85	0.80	-3.53		___	Yes
ppvt122	11	18	0	73	9	0	3.07	0.16	1.16	1.51	-2.76		__C_	Yes
ppvt123	11	91	0	0	9	0	1.69	0.18	1.05	1.11	1.72		___	Yes
ppvt124	11	9	18	27	45	0	4.40	0.13	1.10	1.10	0.85		___	Yes
ppvt125	11	27	0	18	55	0	3.97	0.15	0.85	0.83	1.61		___	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 15. Item statistics by item in the Peabody Picture Vocabulary Test for the Younger Cohort, Vietnam – Tieng Viet Nam

	N	Response options					Difficulty ² (IRT)	CTT item - fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt001	252	3	0	2	95	0	-5.27	0.13	1.06	0.95	-0.19		___	Yes
ppvt002	252	1	3	96	1	0	-5.27	0.14	1.03	0.68	-0.54		___	Yes
ppvt003	252	98	0	1	0	0	-6.53	0.10	0.89	0.33	1.06		__C_	Yes
ppvt004	252	90	6	1	3	0	-4.47	0.22	0.89	0.67	0.40		___	Yes
ppvt005	252	5	5	26	64	0	-2.63	0.37	0.96	1.06	0.17		___	Yes
ppvt006	252	91	4	2	4	0	-4.57	0.20	0.99	0.82	0.00		___	Yes
ppvt007	252	6	89	3	2	0	-4.34	0.22	0.95	0.75	0.29		___	Yes
ppvt008	252	18	15	64	3	0	-2.58	0.37	1.04	1.03	-0.09		___	Yes
ppvt009	252	4	3	8	85	0	-4.00	0.22	1.06	1.48	-0.21		___	Yes
ppvt010	252	17	48	18	17	0	-1.89	0.42	1.11	1.13	0.27		___	Yes
ppvt011	252	9	75	8	8	0	-3.28	0.31	1.06	1.05	0.02		___	Yes
ppvt012	252	12	14	62	13	0	-2.52	0.38	1.01	0.96	0.11		___	Yes
ppvt013	307	7	85	4	4	0	-3.77	0.28	0.98	0.83	0.00		___	Yes
ppvt014	307	70	11	7	11	0	-2.80	0.36	1.05	1.27	-0.13		___	Yes
ppvt015	307	3	4	92	2	0	-4.48	0.21	0.96	0.73	-0.19		___	Yes
ppvt016	307	57	24	9	9	0	-2.06	0.42	1.06	1.09	0.34		___	Yes
ppvt017	307	10	13	63	13	0	-2.40	0.43	0.88	0.88	-0.13		___	Yes
ppvt018	307	22	45	21	12	0	-1.55	0.44	1.28	1.43	-0.69	*	__D	Yes
ppvt019	307	69	13	8	10	0	-2.66	0.37	0.95	1.15	0.37		___	Yes
ppvt020	307	8	22	18	52	0	-1.84	0.44	1.07	1.13	-0.14		___	Yes
ppvt021	307	7	13	70	9	0	-2.77	0.36	1.04	1.11	-0.47		___	Yes
ppvt022	307	27	3	10	59	0	-2.15	0.40	1.07	1.15	0.09		___	Yes
ppvt023	307	19	60	15	6	0	-2.20	0.39	1.10	1.31	-0.32		___	Yes
ppvt024	307	17	4	9	71	0	-2.85	0.34	1.11	1.28	0.00		___	Yes
ppvt025	745	47	26	13	14	0	-0.49	0.50	1.18	1.28	0.00		___	Yes
ppvt026	745	3	3	90	4	0	-3.28	0.37	0.88	0.71	-0.24		___	Yes
ppvt027	745	4	6	3	87	0	-2.97	0.42	0.82	0.57	-0.35		___	Yes
ppvt028	745	4	2	3	91	0	-3.45	0.37	0.84	0.53	-0.08		___	Yes
ppvt029	745	86	9	4	2	0	-2.87	0.41	0.89	0.78	-0.06		___	Yes
ppvt030	745	4	90	5	1	0	-3.31	0.37	0.96	0.82	0.60	*	__D	Yes
ppvt031	745	3	8	82	7	0	-2.50	0.44	0.84	0.69	-0.09		___	Yes
ppvt032	745	3	93	2	3	0	-3.66	0.35	0.91	0.60	0.15		___	Yes
ppvt033	745	7	82	3	8	0	-2.53	0.44	0.88	0.81	-0.20		___	Yes
ppvt034	745	84	9	4	3	0	-2.64	0.42	0.87	0.82	0.23		___	Yes
ppvt035	745	4	6	11	79	0	-2.21	0.46	0.87	0.78	-0.41	*	__D	Yes
ppvt036	745	2	20	77	2	0	-2.08	0.47	0.90	0.87	0.02		___	Yes
ppvt037	910	18	50	28	4	0	-0.28	0.49	1.16	1.22	0.00		___	Yes
ppvt038	910	1	3	3	93	0	-3.34	0.43	0.93	0.56	-0.02		___	Yes
ppvt039	910	7	13	69	11	0	-1.28	0.53	0.89	0.85	0.38	*	__D	Yes
ppvt040	910	46	21	24	10	0	-0.07	0.50	1.14	1.26	-0.19		___	Yes
ppvt041	910	36	56	4	5	0	-0.57	0.49	1.12	1.21	-0.35	*	__D	Yes
ppvt042	910	23	28	23	25	1	1.15	0.55	1.13	1.33	0.00		___	Yes
ppvt043	910	7	6	79	8	0	-1.96	0.48	0.98	0.84	0.58	*	__D	Yes
ppvt044	910	11	11	8	70	0	-1.31	0.51	0.95	0.97	0.00		___	Yes
ppvt045	910	13	10	7	70	0	-1.32	0.49	1.05	1.01	-0.13		___	Yes
ppvt046	910	10	79	8	3	0	-1.88	0.51	0.87	0.76	-0.34		___	Yes
ppvt047	910	3	3	92	2	0	-3.20	0.42	0.94	0.80	0.11		___	Yes
ppvt048	910	84	7	2	7	0	-2.28	0.42	1.06	1.10	0.36		___	Yes
ppvt049	1148	1	5	93	1	0	-2.93	0.49	0.96	0.90	0.17		___	Yes
ppvt050	1148	67	14	6	12	0	-0.77	0.53	0.99	1.01	0.10		___	Yes
ppvt051	1148	85	7	2	6	0	-1.95	0.54	0.94	0.82	0.06		___	Yes
ppvt052	1148	22	64	10	4	0	-0.58	0.54	0.98	0.98	-0.08		___	Yes
ppvt053	1148	28	15	32	25	0	1.36	0.55	1.07	1.12	0.29	*	__D	Yes
ppvt054	1148	8	60	19	12	1	-0.40	0.47	1.15	1.24	0.18		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item - fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt055	1148	72	26	1	1	0	-1.04	0.50	1.05	1.07	-0.11		___	Yes
ppvt056	1148	13	18	28	41	0	0.52	0.52	1.08	1.15	-0.07		___	Yes
ppvt057	1148	7	79	8	6	0	-1.51	0.54	0.96	0.88	0.29		___	Yes
ppvt058	1148	4	10	80	5	0	-1.59	0.50	0.97	1.03	0.32	*	___D	Yes
ppvt059	1148	7	9	75	9	0	-1.25	0.52	0.96	1.00	-0.07		___	Yes
ppvt060	1148	11	10	7	72	0	-1.05	0.56	0.90	0.83	-0.11		___	Yes
ppvt061	1431	15	1	1	82	0	-1.37	0.54	1.03	0.98	0.00		___	Yes
ppvt062	1431	19	48	21	12	0	2.03	0.48	1.08	1.59	-0.15		__C_	Yes
ppvt063	1431	1	95	1	2	0	-3.02	0.58	0.92	0.67	0.00		___	Yes
ppvt064	1431	3	5	90	2	0	-2.09	0.59	0.89	0.72	-0.28		___	Yes
ppvt065	1431	15	4	6	74	0	-0.81	0.55	0.99	0.98	-0.10		___	Yes
ppvt066	1431	12	16	57	14	0	0.09	0.54	0.96	0.96	0.00		___	Yes
ppvt067	1431	22	60	8	9	0	-0.05	0.46	1.15	1.23	-0.07		___	Yes
ppvt068	1431	89	3	6	2	0	-1.93	0.56	1.00	0.92	0.10		___	Yes
ppvt069	1431	6	3	6	85	0	-1.56	0.53	0.97	0.98	-0.20		___	Yes
ppvt070	1431	11	82	5	2	0	-1.35	0.57	0.95	0.87	-0.08		___	Yes
ppvt071	1431	7	7	82	4	0	-1.30	0.52	1.04	1.04	0.00		___	Yes
ppvt072	1431	85	8	3	4	0	-1.63	0.56	0.98	0.88	-0.35	*	___D	Yes
ppvt073	1935	46	18	10	25	1	2.46	0.23	1.05	1.39	0.00		___	Yes
ppvt074	1935	5	2	1	92	0	-2.02	0.58	0.90	0.77	-0.16		___	Yes
ppvt075	1935	5	2	92	1	0	-2.03	0.56	0.95	0.85	-0.13		___	Yes
ppvt076	1935	66	12	13	10	0	-0.01	0.46	0.97	1.00	0.49	*	___D	Yes
ppvt077	1935	87	3	3	6	0	-1.48	0.54	0.96	0.89	0.43	*	___D	Yes
ppvt078	1935	4	4	85	6	0	-1.24	0.57	0.88	0.78	-0.23		___	Yes
ppvt079	1935	1	91	6	2	0	-1.89	0.56	1.00	0.83	0.00		___	Yes
ppvt080	1935	71	6	16	6	0	-0.30	0.51	0.94	0.93	-0.12		___	Yes
ppvt081	1935	10	75	8	7	0	-0.51	0.51	0.94	0.92	0.25	*	___D	Yes
ppvt082	1935	10	4	6	79	0	-0.82	0.49	0.97	1.01	0.34	*	___D	Yes
ppvt083	1935	3	15	73	9	0	-0.42	0.52	0.91	0.87	0.26	*	___D	Yes
ppvt084	1935	17	15	10	58	0	0.39	0.46	0.98	0.98	0.57	*	___D	Yes
ppvt085	1827	32	28	19	21	0	1.88	0.35	1.04	1.13	0.25	*	___D	Yes
ppvt086	1827	18	19	41	22	0	2.25	0.35	1.02	1.09	-0.17		___	Yes
ppvt087	1827	21	21	40	19	0	2.31	0.30	1.01	1.21	0.13		___	Yes
ppvt088	1827	7	21	9	63	0	0.24	0.53	1.01	0.99	-0.11		___	Yes
ppvt089	1827	18	24	43	16	0	1.15	0.33	1.13	1.19	-0.21	*	___D	Yes
ppvt090	1827	74	6	6	14	0	-0.30	0.56	0.95	0.94	-0.39	*	___D	Yes
ppvt091	1827	24	26	32	17	0	2.64	0.28	1.04	1.12	0.00		___	Yes
ppvt092	1827	19	38	40	4	0	1.41	0.42	1.00	1.02	-0.05		___	Yes
ppvt093	1827	32	3	62	2	0	0.28	0.44	1.11	1.13	-0.36	*	___D	Yes
ppvt094	1827	34	26	34	6	0	2.01	0.36	1.01	1.10	0.10		___	Yes
ppvt095	1827	8	16	44	32	0	1.12	0.48	0.97	1.02	0.25	*	___D	Yes
ppvt096	1827	50	35	6	8	1	0.84	0.50	0.96	0.99	-0.16		___	Yes
ppvt097	905	2	62	23	12	0	0.72	0.60	0.98	0.97	0.36	*	___D	Yes
ppvt098	905	27	45	14	15	0	1.51	0.51	1.04	1.06	-0.16		___	Yes
ppvt099	905	2	4	12	82	0	-0.36	0.69	1.01	1.06	-0.54	*	___D	Yes
ppvt100	905	14	14	54	17	0	1.08	0.61	0.90	0.88	-0.08		___	Yes
ppvt101	905	1	9	4	85	0	-0.61	0.68	1.05	1.28	-0.13		___	Yes
ppvt102	905	42	10	8	39	0	1.63	0.50	0.98	0.97	-0.26		___	Yes
ppvt103	905	1	3	90	5	0	-1.12	0.72	1.00	1.07	-0.17		___	Yes
ppvt104	905	3	6	3	88	0	-0.86	0.70	1.06	1.38	0.08		___	Yes
ppvt105	905	76	12	5	7	0	0.05	0.70	0.91	0.88	0.15		___	Yes
ppvt106	905	9	50	30	11	0	1.28	0.56	1.02	1.03	0.15		___	Yes
ppvt107	905	44	15	32	8	0	1.52	0.51	1.03	1.04	-0.10		___	Yes
ppvt108	905	2	21	74	1	2	0.13	0.68	0.96	0.92	0.38	*	___D	Yes
ppvt109	847	14	18	29	39	0	1.85	0.49	1.03	1.05	-0.14		___	Yes
ppvt110	847	18	71	8	2	0	0.38	0.66	1.02	0.99	0.15		___	Yes
ppvt111	847	7	38	47	8	0	1.48	0.57	0.97	0.96	0.08		___	Yes
ppvt112	847	9	16	68	8	0	0.55	0.64	1.02	1.08	-0.39	*	___D	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item - fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt113	847	17	23	21	38	0	3.05	0.28	1.10	1.30	-0.30		___	Yes
ppvt114	847	75	9	3	12	0	0.15	0.68	1.02	1.04	-0.07		___	Yes
ppvt115	847	21	18	23	37	1	1.93	0.51	0.97	0.98	0.20		___	Yes
ppvt116	847	18	63	8	11	0	0.77	0.63	0.99	0.99	0.17		___	Yes
ppvt117	847	31	24	34	10	0	2.17	0.46	0.99	1.01	0.00		___	Yes
ppvt118	847	1	92	3	3	0	-1.23	0.74	1.03	1.27	0.19		___	Yes
ppvt119	847	2	1	2	94	0	-1.56	0.75	1.02	1.09	-0.52		___	Yes
ppvt120	847	28	15	39	13	5	1.82	0.53	0.94	0.95	-0.32	*	___D	Yes
ppvt121	716	24	65	9	2	0	0.79	0.67	0.92	0.88	0.14		___	Yes
ppvt122	716	9	17	19	55	0	1.22	0.55	1.11	1.13	0.19		___	Yes
ppvt123	716	9	21	64	7	0	0.83	0.66	0.97	0.96	0.00		___	Yes
ppvt124	716	25	12	36	27	0	3.52	0.31	0.96	1.00	-0.14		___	Yes
ppvt125	716	14	25	44	16	0	1.67	0.53	0.99	0.99	-0.07		___	Yes
ppvt126	716	36	34	16	14	0	2.06	0.47	1.00	1.03	-0.15		___	Yes
ppvt127	716	1	90	6	2	0	-0.94	0.72	1.07	1.32	-0.45		___	Yes
ppvt128	716	42	12	27	19	0	1.78	0.53	0.97	0.97	-0.31	*	___D	Yes
ppvt129	716	3	2	8	87	0	-0.53	0.71	1.07	1.26	-0.07		___	Yes
ppvt130	716	50	14	17	18	0	1.43	0.60	0.95	0.95	0.32	*	___D	Yes
ppvt131	716	16	19	15	50	0	1.47	0.57	1.01	1.01	0.16		___	Yes
ppvt132	716	2	5	83	4	6	-0.22	0.73	0.99	0.91	0.35		___	Yes
ppvt133	592	1	8	88	3	0	-0.56	0.72	0.99	1.01	0.52	*	___D	Yes
ppvt134	592	5	87	5	3	0	-0.51	0.73	0.93	0.82	-0.16		___	Yes
ppvt135	592	54	9	28	9	0	1.35	0.57	1.01	1.01	-0.22		___	Yes
ppvt136	592	7	44	34	16	0	3.31	0.35	0.94	0.84	0.00		___	Yes
ppvt137	592	24	39	16	20	0	2.76	0.38	1.03	1.10	-0.39	*	___D	Yes
ppvt138	592	22	34	18	26	0	2.88	0.37	1.03	1.05	0.17		___	Yes
ppvt139	592	24	9	12	54	0	1.33	0.58	0.99	1.00	0.00		___	Yes
ppvt140	592	2	25	26	47	0	1.65	0.54	0.99	0.99	0.20		___	Yes
ppvt141	592	12	63	12	13	0	0.94	0.63	0.99	0.98	-0.11		___	Yes
ppvt142	592	48	10	38	5	0	2.06	0.50	0.97	0.98	0.16		___	Yes
ppvt143	592	12	74	10	4	0	0.40	0.66	1.00	0.97	-0.17		___	Yes
ppvt144	592	6	7	57	22	8	1.23	0.62	0.93	0.93	0.13		___	Yes
ppvt145	449	6	22	14	57	0	1.37	0.57	0.99	0.98	0.06		___	Yes
ppvt146	449	16	38	35	12	0	2.35	0.45	1.02	1.03	0.00		___	Yes
ppvt147	449	13	76	5	6	0	0.45	0.64	1.03	1.04	-0.57	*	___D	Yes
ppvt148	449	10	14	67	8	0	0.89	0.60	1.04	1.07	-0.27		___	Yes
ppvt149	449	47	28	9	16	0	1.80	0.50	1.06	1.07	0.16		___	Yes
ppvt150	449	9	32	26	33	0	2.76	0.41	0.92	0.90	-0.31		___	Yes
ppvt151	449	6	6	4	84	0	-0.04	0.67	0.99	1.08	-0.23		___	Yes
ppvt152	449	52	16	10	22	0	1.58	0.55	1.01	1.01	0.19		___	Yes
ppvt153	449	5	69	19	6	0	0.82	0.61	1.06	1.08	-0.15		___	Yes
ppvt154	449	8	38	12	41	0	2.18	0.49	0.96	0.96	0.12		___	Yes
ppvt155	449	34	31	25	10	0	2.38	0.45	0.96	0.96	0.15		___	Yes
ppvt156	449	33	16	8	35	8	2.34	0.47	0.95	0.95	-0.23		___	Yes
ppvt157	343	24	41	27	8	0	2.19	0.47	0.98	0.98	0.17		___	Yes
ppvt158	343	21	30	14	36	0	3.20	0.34	0.99	0.96	0.18		___	Yes
ppvt159	343	6	27	32	34	0	2.47	0.44	0.99	1.00	0.26		___	Yes
ppvt160	343	21	35	17	26	1	3.44	0.33	0.96	0.98	0.41		___	Yes
ppvt161	343	6	7	3	84	0	0.09	0.63	1.06	1.22	0.00		___	Yes
ppvt162	343	35	20	24	21	0	2.46	0.44	0.98	0.97	-0.47	*	___D	Yes
ppvt163	343	20	25	37	17	0	2.93	0.38	1.00	1.04	-0.45		___	Yes
ppvt164	343	23	33	20	24	0	3.04	0.37	0.99	1.01	0.00		___	Yes
ppvt165	343	29	9	23	39	0	2.24	0.45	1.01	1.02	0.07		___	Yes
ppvt166	343	4	81	8	7	0	0.26	0.63	1.04	1.10	-0.36		___	Yes
ppvt167	343	11	6	75	8	0	0.61	0.62	1.00	0.99	0.84	*	___D	Yes
ppvt168	343	17	17	26	27	13	2.90	0.37	1.04	1.10	0.19		___	Yes
ppvt169	179	2	23	7	68	0	1.15	0.51	0.91	0.87	0.00		___	Yes
ppvt170	179	6	62	26	6	0	1.43	0.48	0.95	0.94	-0.39		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item - fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
		Male-Female												
ppvt171	179	53	20	17	11	0	1.84	0.44	1.02	1.02	0.34		___	Yes
ppvt172	179	13	27	27	32	0	3.01	0.34	0.94	0.92	0.26		___	Yes
ppvt173	179	27	28	11	34	0	2.67	0.36	0.99	0.98	0.37		___	Yes
ppvt174	179	2	56	41	1	0	1.67	0.45	1.04	1.04	0.07		___	Yes
ppvt175	179	12	17	61	11	0	1.48	0.48	0.95	0.93	-0.48		___	Yes
ppvt176	179	6	74	6	14	0	0.86	0.52	0.98	0.97	0.07		___	Yes
ppvt177	179	46	4	8	41	0	2.10	0.40	1.05	1.06	-0.42		___	Yes
ppvt178	179	11	36	40	13	0	2.39	0.40	0.95	0.93	0.54		___	Yes
ppvt179	179	6	16	11	68	0	1.15	0.50	0.99	0.98	0.12		___	Yes
ppvt180	179	47	6	40	2	4	2.05	0.44	0.91	0.90	-0.24		___	Yes
ppvt181	139	6	24	23	47	0	2.20	0.39	0.98	0.98	-0.40		___	Yes
ppvt182	139	71	20	7	1	0	1.12	0.46	1.14	1.25	-0.08		___	Yes
ppvt183	139	4	7	87	2	0	0.09	0.51	1.09	1.47	0.00		___	Yes
ppvt184	139	2	13	5	80	0	0.63	0.50	1.01	1.04	0.13		___	Yes
ppvt185	139	41	29	13	17	0	2.45	0.35	1.13	1.15	-0.21		___	Yes
ppvt186	139	7	4	86	3	0	0.21	0.52	1.00	1.01	0.05		___	Yes
ppvt187	139	0	99	1	1	0	-2.25	0.55	1.01	1.19	-1.67		___	Yes
ppvt188	139	12	33	22	32	0	2.84	0.33	0.98	0.99	-0.53		___	Yes
ppvt189	139	14	69	14	4	0	1.23	0.47	0.98	0.97	-0.26		___	Yes
ppvt190	139	2	7	78	13	0	0.77	0.50	0.98	0.94	-0.10		___	Yes
ppvt191	139	83	9	1	7	0	0.38	0.51	1.05	1.22	0.00		___	Yes
ppvt192	139	27	59	7	6	1	1.69	0.44	0.99	0.98	-0.24		___	Yes
ppvt193	136	43	13	23	21	0	3.49	0.25	1.04	1.12	-0.40		___	Yes
ppvt194	136	7	54	18	21	0	1.92	0.41	0.98	1.00	-0.38		___	Yes
ppvt195	136	20	10	54	17	0	1.92	0.41	1.04	1.05	-0.26		___	Yes
ppvt196	136	85	6	3	6	0	0.25	0.52	0.95	0.90	-0.68		___	Yes
ppvt197	136	37	18	26	20	0	2.65	0.34	1.04	1.04	0.16		___	Yes
ppvt198	136	16	13	14	57	0	1.77	0.42	1.01	1.01	-0.64		___	Yes
ppvt199	136	1	7	87	5	0	0.12	0.52	0.97	0.96	0.59		___	Yes
ppvt200	136	65	20	6	10	0	1.44	0.44	1.06	1.09	0.05		___	Yes
ppvt201	136	5	41	10	44	0	2.32	0.34	1.21	1.23	0.06		___	Yes
ppvt202	136	33	51	10	6	0	2.05	0.41	0.90	0.89	-0.86	*	___D	Yes
ppvt203	136	3	5	87	5	0	0.12	0.52	0.98	0.93	0.06		___	Yes
ppvt204	136	17	70	9	4	0	1.20	0.48	0.87	0.82	-1.04	*	___D	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 16. Item statistics by item in the Cloze reading comprehension test for the Older Cohort, Ethiopia – Amharic

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Outfit (IRT)	Gender		
						Male				Female		
cloze01	452	34	53	14	-0.93	0.47	1.18	1.41	-0.09		___	Yes
cloze02	452	18	72	10	0.27	0.44	1.12	1.28	0.54		___	Yes
cloze03	452	30	67	3	-0.65	0.46	1.21	1.23	0.45		___	Yes
cloze04	452	59	37	3	-2.60	0.35	1.42	2.77	0.20		__C_	Yes
cloze05	452	29	63	8	-0.61	0.53	1.05	1.22	0.00		___	Yes
cloze06	452	35	57	8	-1.00	0.56	1.01	1.01	-0.17		___	Yes
cloze07	452	64	30	6	-2.91	0.55	1.00	0.92	0.00		___	Yes
cloze08	452	30	51	19	-0.68	0.57	1.00	0.89	-0.12		___	Yes
cloze09	452	9	74	17	1.40	0.40	1.00	0.95	0.55		___	Yes
cloze10	452	50	35	15	-1.96	0.67	0.80	0.70	-0.16		___	Yes
cloze11	452	26	48	26	-0.38	0.56	0.99	0.88	0.91	*	__D	Yes
cloze12	452	34	46	20	-0.96	0.57	1.00	0.99	-0.51	*	__D	Yes
cloze13	452	41	27	32	-1.43	0.57	1.01	1.10	-0.07		___	Yes
cloze14	452	21	36	42	-0.03	0.57	0.94	0.81	-0.29		___	Yes
cloze15	452	38	19	43	-1.23	0.68	0.81	0.68	-0.21		___	Yes
cloze16	452	14	34	53	0.72	0.45	1.09	0.88	-0.02		___	Yes
cloze17	452	10	33	57	1.14	0.50	0.92	0.59	0.17		___	Yes
cloze18	452	9	36	55	1.40	0.49	0.88	0.60	-0.05		___	Yes
cloze19	452	4	33	63	2.48	0.31	1.01	0.84	-0.52		___	Yes
cloze20	452	6	31	64	1.95	0.48	0.78	0.44	1.13	*	__CD	Yes
cloze21	452	15	18	66	0.53	0.64	0.75	0.53	-0.41		___	Yes
cloze22	452	9	20	71	1.40	0.47	0.94	0.63	-0.47		___	Yes
cloze23	452	12	15	73	0.96	0.58	0.80	0.54	-0.20		___	Yes
cloze24	452	2	17	80	3.10	0.32	0.97	0.26	-0.26		__C_	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 17. Item statistics by item in the Cloze reading comprehension test for the Older Cohort, Ethiopia – Oromiffa

	N	Response options			Difficulty ² (IRT)	CTT item- fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
						Male				Female		
cloze01	195	32	67	2	-1.17	0.37	1.28	1.77	0.26		__C_	Yes
cloze02	195	14	83	4	0.26	0.47	0.96	1.28	-0.45		___	Yes
cloze03	195	33	62	6	-1.24	0.38	1.27	1.44	0.13		___	Yes
cloze04	195	58	37	5	-2.89	0.40	1.24	1.71	0.53		__C_	Yes
cloze05	195	15	77	8	0.10	0.42	1.06	1.04	-0.27		___	Yes
cloze06	195	31	63	7	-1.10	0.56	0.98	0.96	-0.72		___	Yes
cloze07	195	16	74	10	0.05	0.52	0.92	0.77	-0.15		___	Yes
cloze08	195	24	64	12	-0.63	0.43	1.12	1.08	0.61		___	Yes
cloze09	195	55	34	11	-2.65	0.61	0.92	0.86	-0.22		___	Yes
cloze10	195	41	43	16	-1.76	0.58	0.96	0.89	0.15		___	Yes
cloze11	195	28	47	25	-0.89	0.57	0.94	0.81	-0.19		___	Yes
cloze12	195	20	51	29	-0.31	0.47	1.04	0.99	-0.24		___	Yes
cloze13	195	35	34	30	-1.41	0.61	0.91	0.84	0.21		___	Yes
cloze14	195	29	31	39	-1.00	0.58	0.93	0.91	0.60		___	Yes
cloze15	195	7	43	50	1.15	0.41	0.92	0.95	-1.52	*	__D	Yes
cloze16	195	11	39	50	0.61	0.54	0.82	0.57	-0.98		___	Yes
cloze17	195	6	35	58	1.34	0.36	0.95	0.83	-0.25		___	Yes
cloze18	195	13	28	58	0.31	0.51	0.91	0.67	0.32		___	Yes
cloze19	195	17	19	63	-0.09	0.55	0.89	0.69	0.16		___	Yes
cloze20	195	8	25	67	0.97	0.43	0.92	0.74	-0.08		___	Yes
cloze21	195	8	22	70	1.06	0.45	0.92	0.56	0.06		___	Yes
cloze22	195	1	26	73	3.37	0.28	0.82	0.18	0.69		__C_	Yes
cloze23	195	4	19	77	1.83	0.38	0.91	0.40	0.62		__C_	Yes
cloze24	195	1	19	81	4.10	0.02	1.07	1.81	-0.47		A_C_	No

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 18. Item statistics by item in the Cloze reading comprehension test for the Older Cohort, Ethiopia – Tigrigna

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
						Male				Female		
cloze01	160	12	59	29	-1.04	0.49	1.05	1.13	-1.09		___	Yes
cloze02	160	3	81	17	1.03	0.37	0.85	1.05	-0.34		___	Yes
cloze03	160	39	45	16	-3.70	0.62	1.08	1.26	0.93		___	Yes
cloze04	160	35	46	19	-3.31	0.54	1.28	1.42	0.05		___	Yes
cloze05	160	8	67	26	-0.36	0.45	1.01	1.14	-0.59		___	Yes
cloze06	160	11	56	33	-0.95	0.50	1.03	1.01	1.23	*	___D	Yes
cloze07	160	21	54	26	-2.01	0.65	0.90	0.78	-0.16		___	Yes
cloze08	160	21	42	37	-2.08	0.58	1.01	1.11	-0.27		___	Yes
cloze09	160	15	45	40	-1.42	0.70	0.75	0.54	-0.27		___	Yes
cloze10	160	8	35	57	-0.47	0.56	0.87	0.49	-0.76		__C_	Yes
cloze11	160	5	30	65	0.18	0.32	1.14	0.86	0.88		___	Yes
cloze12	160	17	16	67	-1.63	0.54	1.07	1.10	1.05	*	___D	Yes
cloze13	160	2	23	76	1.36	0.01	1.23	4.86	-1.26		A_C_	No
cloze14	160	1	19	80	1.81	0.28	0.95	0.34	-0.85		__C_	Yes
cloze15	160	1	14	86	2.55	0.28	0.91	0.11	-0.35		__C_	Yes
cloze16	160	0	18	82	3.79		c	c	c		__BC_	No
cloze17	160	3	10	88	1.03	0.47	0.75	0.24	-0.34		__C_	Yes
cloze18	160	3	12	85	0.76	0.35	1.02	0.48	-0.70		__C_	Yes
cloze19	160	0	10	90	3.79		c	c	c		__BC_	No
cloze20	160	2	4	94	1.36	0.30	0.90	0.82	-1.26		___	Yes
cloze21	160	1	4	94	1.81	0.35	0.81	0.23	-0.85		__C_	Yes
cloze22	160	0	5	95	3.79		c	c	c		__BC_	No
cloze23	160	1	3	96	2.55	0.34	0.76	0.06	-0.35		__C_	Yes
cloze24	160	1	4	95	2.55	0.19	1.02	0.28	-0.35		__C_	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 19. Item statistics by item in the Cloze reading comprehension test for the Older Cohort, India – Telugu

	N	Response options			Difficulty ² (IRT)	CTT item- fit indicator	IRT item-fit indicator		Bias analysis ³	Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)			
						Male-Female					
cloze01	791	55	17	28	-0.76	0.65	0.94	0.89	-0.44 *	___D	Yes
cloze02	791	47	24	28	-0.27	0.59	1.04	1.19	0.00	___	Yes
cloze03	791	63	30	7	-1.33	0.48	1.32	1.55	0.82 *	___CD	Yes
cloze04	791	76	18	6	-2.41	0.50	1.24	1.38	0.67 *	___D	Yes
cloze05	791	45	33	22	-0.15	0.60	1.03	1.03	0.30	___	Yes
cloze06	791	59	22	19	-1.01	0.64	0.97	0.97	0.08	___	Yes
cloze07	791	66	23	11	-1.57	0.60	1.03	1.15	0.16	___	Yes
cloze08	791	78	12	11	-2.58	0.58	0.96	1.85	0.38	___C_	Yes
cloze09	791	34	34	33	0.64	0.52	1.10	1.23	0.00	___	Yes
cloze10	791	46	26	28	-0.20	0.66	0.90	0.77	-0.21	___	Yes
cloze11	791	25	43	32	1.27	0.54	0.99	0.84	-0.21	___	Yes
cloze12	791	63	19	18	-1.32	0.63	0.97	0.93	-0.76 *	___D	Yes
cloze13	791	15	61	24	2.14	0.40	1.11	1.24	0.15	___	Yes
cloze14	791	52	24	24	-0.57	0.67	0.89	0.80	0.51 *	___D	Yes
cloze15	791	9	37	54	2.95	0.36	1.01	0.87	-0.71 *	___D	Yes
cloze16	791	25	30	45	1.27	0.53	1.00	1.10	0.31	___	Yes
cloze17	791	53	15	32	-0.65	0.62	1.00	0.99	0.06	___	Yes
cloze18	791	52	24	25	-0.56	0.64	0.95	0.96	0.18	___	Yes
cloze19	791	26	29	45	1.17	0.57	0.94	0.78	-0.27	___	Yes
cloze20	791	34	34	32	0.63	0.56	1.03	0.99	0.00	___	Yes
cloze21	791	47	20	33	-0.28	0.65	0.92	0.82	-0.36 *	___D	Yes
cloze22	791	18	36	46	1.86	0.49	0.97	0.82	-0.71 *	___D	Yes
cloze23	791	40	19	41	0.21	0.69	0.82	0.68	-0.29	___	Yes
cloze24	791	22	36	42	1.51	0.53	0.95	0.76		___	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 20. Item statistics by item in the Cloze reading comprehension test for the Older Cohort, Peru – Spanish

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
						Male				Female		
cloze01	658	85	8	7	-1.77	0.59	0.91	0.78	-0.34		___	Yes
cloze02	658	81	11	7	-1.38	0.58	0.93	0.89	0.25		___	Yes
cloze03	658	75	17	8	-0.79	0.38	1.35	1.62	0.22		__C_	Yes
cloze04	658	81	13	7	-1.32	0.47	1.16	1.28	-0.19		___	Yes
cloze05	658	85	14	1	-1.75	0.49	1.08	1.53	0.27		__C_	Yes
cloze06	658	72	24	3	-0.58	0.53	1.08	1.14	-0.46	*	___D	Yes
cloze07	658	77	20	3	-0.98	0.57	0.98	0.98	0.45		___	Yes
cloze08	658	86	7	7	-1.90	0.47	1.09	1.21	0.47		___	Yes
cloze09	658	70	21	10	-0.40	0.47	1.18	1.43	-0.07		___	Yes
cloze10	658	78	17	5	-1.03	0.50	1.08	1.28	-0.28		___	Yes
cloze11	658	79	12	10	-1.11	0.64	0.84	0.69	0.55	*	___D	Yes
cloze12	658	73	19	8	-0.64	0.56	0.99	1.21	0.07		___	Yes
cloze13	658	72	15	13	-0.54	0.58	0.96	1.11	-0.44	*	___D	Yes
cloze14	658	74	12	14	-0.74	0.64	0.86	0.73	0.00		___	Yes
cloze15	658	44	32	24	1.14	0.53	0.98	1.03	0.00		___	Yes
cloze16	658	50	30	21	0.84	0.57	0.95	0.87	0.30		___	Yes
cloze17	658	47	30	24	1.01	0.58	0.91	0.82	-0.32		___	Yes
cloze18	658	47	35	19	1.01	0.50	1.06	1.24	0.07		___	Yes
cloze19	658	58	19	22	0.34	0.64	0.84	0.76	0.23		___	Yes
cloze20	658	31	36	33	1.95	0.49	0.99	0.85	0.37		___	Yes
cloze21	658	24	39	37	2.44	0.41	1.05	1.01	0.30		___	Yes
cloze22	658	34	25	40	1.73	0.55	0.88	0.80	-0.24		___	Yes
cloze23	658	31	26	43	1.95	0.53	0.86	1.23	-0.66	*	___D	Yes
cloze24	658	23	34	43	2.52	0.48	0.90	0.69	-0.15		___	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
 Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 21. Item statistics by item in the Cloze reading comprehension test for the Older Cohort, Vietnam – Tieng Viet Nam

	N	Response options			Difficulty ² (IRT)	CTT item- fit indicator	IRT item-fit indicator		Bias analysis ³	Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)			
						Gender Male-Female					
cloze01	941	88	6	6	-0.63	0.46	1.11	1.27	0.50 *	___D	Yes
cloze02	941	96	3	1	-2.25	0.38	1.10	1.13	0.32	___	Yes
cloze03	941	77	18	5	0.47	0.40	1.31	1.48	-0.44 *	___D	Yes
cloze04	941	95	5	1	-1.90	0.33	1.11	1.93	-0.30	___C_	Yes
cloze05	941	91	8	1	-1.07	0.45	1.07	1.07	-0.33	___	Yes
cloze06	941	83	13	3	-0.14	0.41	1.23	1.46	0.21	___	Yes
cloze07	941	96	3	1	-2.25	0.37	1.02	0.93	0.32	___	Yes
cloze08	941	94	3	2	-1.79	0.38	1.02	1.42	0.23	___	Yes
cloze09	941	89	7	4	-0.80	0.41	1.13	1.53	-0.80 *	___CD	Yes
cloze10	941	89	7	4	-0.77	0.51	0.97	1.21	-0.21	___	Yes
cloze11	941	59	33	8	1.72	0.54	1.03	1.02	0.00	___	Yes
cloze12	941	94	2	4	-1.73	0.48	0.86	1.09	-0.33	___	Yes
cloze13	941	78	13	9	0.39	0.50	1.13	1.23	-0.16	___	Yes
cloze14	941	57	26	17	1.85	0.50	1.13	1.22	0.00	___	Yes
cloze15	941	62	28	10	1.53	0.50	1.13	1.17	-0.18	___	Yes
cloze16	941	67	19	14	1.18	0.61	0.93	0.89	-0.12	___	Yes
cloze17	941	70	17	13	0.98	0.59	0.97	0.91	0.18	___	Yes
cloze18	941	74	13	13	0.70	0.63	0.89	0.83	-0.10	___	Yes
cloze19	941	81	6	13	0.10	0.70	0.74	0.64	0.30	___	Yes
cloze20	941	80	6	14	0.18	0.69	0.77	0.62	0.33	___	Yes
cloze21	941	65	20	15	1.36	0.65	0.86	0.81	0.00	___	Yes
cloze22	941	71	11	18	0.94	0.65	0.86	0.81	0.26	___	Yes
cloze23	941	72	9	19	0.85	0.65	0.87	0.86	0.21	___	Yes
cloze24	941	69	13	18	1.07	0.69	0.79	0.68	0.09	___	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 22. Item statistics by item in the Maths achievement test for the Older Cohort, Ethiopia

	N	Response options			Difficulty (IRT) ²	CTT item-fit indicator	IRT item-fit indicator			Bias analysis ³				Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		Language			
										Male-Female	First-Second	First-Third	Second-Third		
math01	816	58	38	5	-2.94	0.53	0.96	0.99	0.41 *	0.69 *	0.13	-0.55 *	___DE	Yes	
math02	816	63	33	4	-3.27	0.54	0.89	0.87	0.29	-0.05	-0.31	-0.26	_____	Yes	
math03	816	24	48	28	-0.75	0.66	0.78	0.71	-0.17	0.12	-0.86 *	-0.98 *	___E	Yes	
math04	816	21	35	45	-0.54	0.55	0.95	1.19	-0.20	0.44	0.89 *	0.45	___E	Yes	
math05	816	31	56	13	-1.31	0.59	0.93	0.82	-0.14	0.20	-0.30	-0.50	_____	Yes	
math06	816	44	39	18	-2.15	0.60	0.91	0.91	-0.31	-0.15	-0.08	0.07	_____	Yes	
math07	816	41	37	22	-1.96	0.67	0.81	0.75	-0.29	-0.13	-0.31	-0.18	_____	Yes	
math08	816	24	50	26	-0.73	0.66	0.80	0.61	-0.17	0.57 *	0.09	-0.49	___E	Yes	
math09	816	50	26	24	-2.47	0.65	0.79	0.71	-0.42 *	-0.43 *	-0.53 *	-0.10	___DE	Yes	
math10	816	32	34	34	-1.35	0.71	0.70	0.59	-0.30	-0.01	-0.39	-0.38	_____	Yes	
math11	816	26	34	40	-0.93	0.60	0.88	0.78	0.19	0.13	-0.06	-0.20	_____	Yes	
math12	816	21	37	42	-0.48	0.69	0.73	0.53	0.25	0.09	0.04	-0.05	_____	Yes	
math13	816	8	39	53	0.92	0.50	0.87	0.52	0.38	-0.86 *	-0.30	0.57	___E	Yes	
math14	816	21	23	56	-0.56	0.59	0.89	0.70	0.37	0.05	-0.31	-0.35	_____	Yes	
math15	816	3	36	60	2.09	0.35	0.90	1.06	-0.14	-0.19	0.01	0.20	_____	Yes	
math16	816	3	16	81	2.19	0.43	0.77	0.26	-0.64	0.53	-0.21	-0.74	_C_	Yes	
math17	816	2	12	86	2.50	0.41	0.75	0.29	-0.77	-0.08	-0.84	-0.76	_C_	Yes	
math18	816	3	18	80	2.33	0.42	0.77	0.20	-0.70	0.51	0.86	0.35	_C_	Yes	
math19	816	4	15	81	1.78	0.40	0.87	0.82	0.57	-0.22	-0.37	-0.15	_____	Yes	
math20	816	3	11	86	2.19	0.44	0.76	0.23	0.20	0.07	-0.37	-0.43	_C_	Yes	
math21	803	33	59	8	-1.46	0.33	1.40	1.72	0.12	-0.79 *	0.12	0.91 *	_C_E	Yes	
math22	803	11	44	45	0.44	0.37	1.16	1.15	0.35	0.82 *	-0.35	-1.17 *	___E	Yes	
math23	803	10	84	5	0.61	0.12	1.42	2.83	0.17	0.34	1.14 *	0.80 *	_C_E	Yes	
math24	803	36	56	8	-1.53	0.22	1.51	1.98	0.14	0.25	0.68 *	0.43	_BC_E	No	
math25	803	23	67	10	-0.76	0.23	1.46	2.05	0.13	-0.31	0.08	0.39	_C_	Yes	
math26	803	13	74	13	0.07	0.09	1.54	3.61	0.10	-0.31	1.62 *	1.93 *	ABC_E	No	
math27	803	6	69	26	1.47	0.39	1.01	0.53	-0.36	-2.58 *	-0.81	1.77 *	___E	Yes	
math28	803	3	64	33	2.33	0.22	0.99	2.34	-0.18	-1.43 *	-0.51	0.92	_C_E	Yes	
math29	803	3	55	42	2.18	0.36	0.94	0.47	0.00	0.72	0.75	0.04	_C_	Yes	
math30	803	3	31	67	2.09	0.24	1.11	1.20	0.03	-1.45 *	-0.57	0.88	___E	Yes	

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
 Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%. E: The difference between language groups is significant at 5%.

Table 23. Item statistics by item in the Maths achievement test for the Older Cohort, India

	N	Response options			Difficulty ² (IRT)	CTT item - fit indicator	IRT item-fit indicator		Bias analysis ³				Warnings	Item kept for the analysis ⁴	
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		Language			
										Male- Female	First- Second				
math01	893	77	20	3	-3.27	0.49	0.96	1.11	0.44	*	-0.14		__D_	Yes	
math02	893	81	16	3	-3.60	0.44	1.00	1.36	0.14		-0.35		_____	Yes	
math03	893	42	17	41	-0.95	0.73	0.72	0.57	0.00		1.00	*	___E	Yes	
math04	893	40	7	53	-0.73	0.69	0.77	0.80	0.11		0.16		_____	Yes	
math05	893	51	37	13	-1.43	0.54	1.05	1.11	0.49	*	-0.60	*	___DE	Yes	
math06	893	53	17	30	-1.58	0.70	0.75	0.63	-0.17		0.35		_____	Yes	
math07	893	53	13	33	-1.59	0.71	0.69	0.61	-0.39	*	-0.13		__D_	Yes	
math08	893	46	31	23	-1.14	0.62	0.90	0.83	0.19		0.24		_____	Yes	
math09	893	61	29	10	-2.11	0.55	0.99	1.12	-0.29		-0.08		_____	Yes	
math10	893	46	30	24	-1.14	0.64	0.89	0.87	-0.08		0.27		_____	Yes	
math11	893	54	22	23	-1.65	0.62	0.92	0.85	0.29		0.38		_____	Yes	
math12	893	9	46	45	1.87	0.45	0.95	0.68	-0.85	*	-0.03		__D_	Yes	
math13	893	21	39	40	0.60	0.54	0.97	0.80	0.73	*	-0.20		__D_	Yes	
math14	893	13	35	52	1.40	0.60	0.75	0.39	0.10		0.33		__C_	Yes	
math15	893	19	42	39	0.75	0.52	1.01	0.88	0.85	*	0.30		__D_	Yes	
math16	893	8	15	77	2.04	0.50	0.81	0.40	-0.16		1.27	*	__C_E	Yes	
math17	893	6	14	80	2.40	0.47	0.80	0.36	0.02		0.75	*	__C_E	Yes	
math18	893	11	38	51	1.60	0.53	0.85	0.54	0.15		0.26		_____	Yes	
math19	893	5	31	64	2.63	0.36	0.99	0.77	-0.99	*	-0.70		__D_	Yes	
math20	893	3	18	79	3.19	0.36	0.83	0.75	-0.44		0.85		_____	Yes	
math21	888	37	58	5	-0.56	0.33	1.44	1.93	0.00		-0.32		__C_	Yes	
math22	888	45	26	29	-1.08	0.53	1.08	1.24	-0.26		0.53		_____	Yes	
math23	888	14	81	6	1.25	0.25	1.33	2.75	-0.22		-1.08	*	__C_E	Yes	
math24	888	37	56	7	-0.61	0.44	1.26	1.48	0.47	*	-0.34		__D_	Yes	
math25	888	32	57	11	-0.23	0.32	1.40	2.01	-0.15		-0.47		__C_	Yes	
math26	888	21	72	7	0.59	0.03	1.76	4.60	0.86	*	-1.44	*	ABCDE	No	
math27	888	25	50	25	0.24	0.64	0.82	0.72	-0.95	*	0.05		__D_	Yes	
math28	888	22	50	28	0.49	0.60	0.86	0.72	-0.64	*	-0.02		__D_	Yes	
math29	888	19	34	47	0.74	0.61	0.82	0.58	-0.58	*	-0.21		__D_	Yes	
math30	888	9	28	63	1.90	0.41	0.98	0.89	-0.63	*	-0.43		__D_	Yes	

1. NR: No Response.
2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%. E: The difference between language groups is significant at 5%.

Table 24. *Item statistics by item in the Maths achievement test for the Older Cohort, Peru*

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Outfit (IRT)	Gender		
						Male				Female		
math01	663	83	16	1	-2.48	0.16	1.37	2.47	0.35		__C_	Yes
math02	663	92	7	1	-3.78	0.22	1.15	8.79	0.46		__C_	Yes
math03	663	73	20	7	-1.74	0.62	0.80	0.63	0.18		___	Yes
math04	663	75	5	20	-1.89	0.61	0.79	0.66	0.00		___	Yes
math05	663	69	27	5	-1.45	0.47	1.01	1.29	0.28		___	Yes
math06	663	73	15	12	-1.69	0.57	0.85	0.83	-0.25		___	Yes
math07	663	66	17	17	-1.27	0.63	0.79	0.69	-0.38	*	__D	Yes
math08	663	58	33	9	-0.81	0.45	1.06	1.13	0.32		___	Yes
math09	663	71	22	6	-1.61	0.55	0.87	1.08	-0.29		___	Yes
math10	663	62	26	12	-1.05	0.56	0.90	0.88	0.33		___	Yes
math11	663	66	21	13	-1.26	0.40	1.10	1.81	0.27		__C_	Yes
math12	663	46	38	16	-0.16	0.59	0.85	0.81	0.12		___	Yes
math13	663	38	37	25	0.24	0.52	0.93	0.87	0.38	*	__D	Yes
math14	663	42	33	24	0.03	0.46	1.04	0.98	0.16		___	Yes
math15	663	21	61	18	1.33	0.35	1.05	1.30	-0.16		___	Yes
math16	663	17	22	61	1.67	0.50	0.85	0.61	-0.31		___	Yes
math17	663	15	16	69	1.86	0.49	0.84	0.54	-0.14		___	Yes
math18	663	6	51	43	3.12	0.32	0.91	0.67	1.22	*	__D	Yes
math19	663	13	32	55	2.06	0.42	0.92	0.65	0.23		___	Yes
math20	663	10	20	70	2.39	0.45	0.83	0.50	0.00		__C_	Yes
math21	662	76	24	0	-1.93	0.38	1.14	1.37	-0.38		___	Yes
math22	662	52	33	15	-0.47	0.46	1.05	1.06	0.33		___	Yes
math23	662	36	64	1	0.40	0.43	1.05	1.08	-0.37	*	__D	Yes
math24	662	46	54	1	-0.15	0.35	1.19	1.44	-0.21		___	Yes
math25	662	28	71	1	0.83	0.21	1.29	1.86	-0.51	*	__CD	Yes
math26	662	23	77	1	1.21	0.21	1.25	2.24	0.00		__C_	Yes
math27	662	26	66	8	1.00	0.51	0.90	0.77	-0.20		___	Yes
math28	662	21	79	1	1.35	0.44	0.95	0.86	-0.31		___	Yes
math29	662	16	43	41	1.72	0.49	0.87	0.62	-0.27		___	Yes
math30	662	9	35	56	2.53	0.34	0.95	0.92	0.08		___	Yes

1. NR: No Response.
 2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.
 3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).
 4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.
 Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 25. Item statistics by item in the Maths achievement test for the Older Cohort, Vietnam

	N	Response options			Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		Correct (%)	Incorrect (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
						Male				Female		
math01	944	90	9	0	-2.46	0.33	1.09	1.81	0.08		__C_	Yes
math02	944	93	7	0	-2.85	0.28	1.11	3.97	-0.17		__C_	Yes
math03	944	76	16	8	-0.96	0.64	0.79	0.57	0.24		___	Yes
math04	944	78	6	17	-1.09	0.61	0.82	0.67	0.23		___	Yes
math05	944	70	27	3	-0.52	0.48	1.12	1.73	-0.23		__C_	Yes
math06	944	82	8	9	-1.53	0.53	0.89	0.91	0.12		___	Yes
math07	944	80	10	11	-1.26	0.63	0.77	0.52	-0.17		___	Yes
math08	944	64	29	7	-0.12	0.48	1.18	1.45	0.00		___	Yes
math09	944	88	9	3	-2.13	0.46	0.92	1.15	-0.31		___	Yes
math10	944	79	16	5	-1.18	0.46	1.08	1.16	-0.48	*	__D	No
math11	944	71	19	9	-0.62	0.49	1.09	1.36	0.18		___	Yes
math12	944	74	17	9	-0.85	0.62	0.84	0.80	0.44	*	__D	Yes
math13	944	59	25	16	0.20	0.60	0.98	0.94	0.54	*	__D	Yes
math14	944	62	20	19	0.04	0.63	0.91	0.79	0.17		___	Yes
math15	944	47	39	14	0.94	0.64	0.90	0.83	0.00		___	Yes
math16	944	36	19	45	1.67	0.66	0.79	0.67	-0.60	*	__D	Yes
math17	944	31	18	51	1.98	0.64	0.81	0.67	-0.48	*	__D	Yes
math18	944	25	41	34	2.49	0.57	0.91	0.69	0.30		___	Yes
math19	944	32	30	38	1.96	0.54	1.05	0.97	0.18		___	Yes
math20	944	40	16	43	1.38	0.69	0.77	0.64	0.10		___	Yes
math21	942	85	14	1	-1.76	0.37	1.23	1.62	-0.84	*	__CD	Yes
math22	942	71	21	8	-0.59	0.49	1.12	1.64	0.20		__C_	Yes
math23	942	52	45	3	0.61	0.43	1.30	1.79	0.00		__C_	Yes
math24	942	66	31	3	-0.26	0.47	1.19	1.35	-0.02		___	Yes
math25	942	50	40	11	0.79	0.45	1.26	1.50	-0.12		__C_	Yes
math26	942	47	48	6	0.96	0.50	1.16	1.40	-0.11		___	Yes
math27	942	63	25	12	-0.05	0.60	0.96	1.11	-0.08		___	Yes
math28	942	37	30	34	1.62	0.62	0.89	0.83	-0.07		___	Yes
math29	942	34	28	38	1.82	0.65	0.80	0.67	0.31		___	Yes
math30	942	35	26	39	1.75	0.63	0.84	0.74	0.25		___	Yes

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 26. Item statistics by item in the Peabody Picture Vocabulary Test for the Older Cohort, Ethiopia – Amharic

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
		Male-Female												
ppvt001	102	2	1	4	93	0	-3.09	0.21	0.97	0.68	1.29		Yes	
ppvt002	102	1	1	95	3	0	-3.46	0.19	0.90	0.53	-0.18		Yes	
ppvt003	102	98	0	0	2	0	-4.43	0.11	0.99	0.71	0.25		Yes	
ppvt004	102	96	1	1	2	0	-3.70	0.16	0.94	0.74	-0.91		Yes	
ppvt005	102	2	7	34	57	0	-0.59	0.44	1.06	1.05	-0.49		Yes	
ppvt006	102	93	4	2	1	0	-3.09	0.21	0.89	0.54	-0.03		Yes	
ppvt007	102	25	50	6	19	0	-0.20	0.43	1.08	1.08	0.34		Yes	
ppvt008	102	15	3	78	4	0	-1.64	0.34	0.90	0.82	0.45		Yes	
ppvt009	102	16	5	9	71	0	-1.24	0.38	0.94	0.85	-0.49		Yes	
ppvt010	102	21	40	25	15	0	0.15	0.47	1.17	1.17	0.73		Yes	
ppvt011	102	12	78	5	5	0	-1.58	0.34	0.85	0.78	0.33		Yes	
ppvt012	102	3	20	74	4	0	-1.40	0.38	0.95	0.88	0.23		Yes	
ppvt013	120	3	95	1	1	0	-3.18	0.21	0.88	0.49	0.48	_C_	Yes	
ppvt014	120	88	5	6	2	0	-2.30	0.28	1.00	0.93	0.67		Yes	
ppvt015	120	6	5	84	5	0	-2.01	0.34	0.84	0.66	-0.81		Yes	
ppvt016	120	67	22	11	1	0	-0.94	0.38	1.08	1.06	0.83	* _D	Yes	
ppvt017	120	9	6	76	9	0	-1.38	0.39	0.90	0.82	0.00		Yes	
ppvt018	120	14	41	33	13	0	0.28	0.50	1.14	1.26	-0.40		Yes	
ppvt019	120	77	4	13	6	0	-1.43	0.36	0.98	0.90	-0.13		Yes	
ppvt020	120	10	32	24	34	0	0.63	0.53	1.10	1.12	-0.78		Yes	
ppvt021	120	3	17	78	3	0	-1.48	0.32	1.03	1.06	-0.45		Yes	
ppvt022	120	3	24	17	56	0	-0.39	0.44	1.08	1.08	0.43		Yes	
ppvt023	120	17	43	30	11	0	0.16	0.51	1.01	1.02	-0.46		Yes	
ppvt024	120	23	5	4	68	0	-0.94	0.36	1.05	1.11	0.16		Yes	
ppvt025	202	93	1	1	4	0	-2.64	0.32	0.83	0.43	-0.16	_C_	Yes	
ppvt026	202	6	6	81	6	0	-1.39	0.44	0.81	0.66	-0.36		Yes	
ppvt027	202	7	7	12	73	0	-0.86	0.50	0.88	0.81	-0.35		Yes	
ppvt028	202	1	4	7	88	0	-1.87	0.35	0.86	0.70	0.24		Yes	
ppvt029	202	84	5	5	5	0	-1.57	0.36	0.89	0.88	0.14		Yes	
ppvt030	202	9	70	13	8	0	-0.70	0.47	0.90	0.85	0.09		Yes	
ppvt031	202	4	16	70	9	0	-0.70	0.50	0.84	0.78	-0.77	* _D	Yes	
ppvt032	202	3	88	4	5	0	-1.91	0.39	0.85	0.60	-0.84		Yes	
ppvt033	202	11	66	4	18	0	-0.55	0.48	0.97	0.96	0.37		Yes	
ppvt034	202	58	8	6	27	0	-0.12	0.45	1.12	1.16	0.30		Yes	
ppvt035	202	9	10	6	75	0	-0.95	0.46	0.93	0.83	-0.18		Yes	
ppvt036	202	4	28	64	4	0	-0.43	0.43	1.08	1.14	-0.09		Yes	
ppvt037	216	13	54	10	23	0	0.13	0.55	0.95	0.93	0.00		Yes	
ppvt038	216	1	2	2	95	0	-2.93	0.26	0.93	0.60	-0.03		Yes	
ppvt039	216	6	19	70	5	0	-0.65	0.45	1.01	1.03	0.45		Yes	
ppvt040	216	23	18	13	47	0	1.66	0.61	1.03	1.24	0.42		Yes	
ppvt041	216	31	51	5	13	0	0.26	0.51	1.05	1.06	-0.02		Yes	
ppvt042	216	32	23	19	26	0	1.17	0.59	1.05	1.03	0.37		Yes	
ppvt043	216	31	23	34	12	0	1.10	0.59	1.01	1.14	-0.23		Yes	
ppvt044	216	39	36	9	15	0	2.21	0.62	1.10	1.31	0.22		Yes	
ppvt045	216	14	10	3	73	0	-0.78	0.49	0.93	0.85	0.00		Yes	
ppvt046	216	8	72	9	11	0	-0.75	0.53	0.87	0.80	-0.45		Yes	
ppvt047	216	6	4	86	4	0	-1.72	0.40	0.82	0.56	-0.29		Yes	
ppvt048	216	87	5	1	6	0	-1.85	0.34	0.95	0.88	-0.60		Yes	
ppvt049	248	9	17	63	11	0	-0.17	0.49	1.05	1.17	0.30		Yes	
ppvt050	248	89	4	4	3	0	-1.92	0.38	0.89	0.72	-0.29		Yes	
ppvt051	248	63	15	10	12	0	-0.22	0.52	1.01	1.10	0.35		Yes	
ppvt052	248	40	42	6	12	0	0.81	0.61	0.96	1.04	-0.11		Yes	
ppvt053	248	30	17	17	37	0	1.08	0.54	1.21	1.36	0.76	* _D	Yes	
ppvt054	248	10	58	26	6	0	0.06	0.42	1.25	1.29	0.17		Yes	

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item- fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender Male-Female		
ppvt055	248	41	48	7	3	0	0.86	0.59	1.00	1.02	-0.40		___	Yes
ppvt056	248	5	2	50	44	0	0.75	0.60	0.95	0.98	-0.28		___	Yes
ppvt057	248	13	63	15	9	0	-0.19	0.53	0.98	0.94	0.42		___	Yes
ppvt058	248	23	23	39	15	0	0.96	0.58	1.07	1.17	0.55		___	Yes
ppvt059	248	2	6	91	2	0	-2.05	0.35	0.90	0.87	0.00		___	Yes
ppvt060	248	7	19	4	69	0	-0.49	0.55	0.86	0.80	0.02		___	Yes
ppvt061	302	0	0	0	99	0	-4.77	0.33	1.01	1.38	-1.84		___	Yes
ppvt062	302	96	2	0	1	0	-2.91	0.42	0.86	0.39	0.02		__C_	Yes
ppvt063	302	4	92	2	2	0	-2.14	0.42	0.97	0.90	-0.18		___	Yes
ppvt064	302	5	9	82	4	0	-1.09	0.56	0.82	0.71	-0.74	*	__D	Yes
ppvt065	302	13	4	48	35	0	1.32	0.55	1.25	1.51	0.26		__C_	Yes
ppvt066	302	13	19	37	31	0	1.19	0.54	1.22	1.33	0.10		___	Yes
ppvt067	302	12	68	11	10	0	-0.26	0.48	1.11	1.21	0.63	*	__D	Yes
ppvt068	302	72	13	10	6	0	-0.49	0.56	0.95	0.85	0.08		___	Yes
ppvt069	302	4	7	6	83	0	-1.17	0.44	1.02	1.17	-0.29		___	Yes
ppvt070	302	19	70	5	7	0	-0.34	0.55	0.97	0.97	0.03		___	Yes
ppvt071	302	14	2	55	30	0	0.39	0.60	0.97	0.98	0.14		___	Yes
ppvt072	302	79	4	5	12	0	-0.93	0.48	1.04	1.03	0.08		___	Yes
ppvt073	386	6	89	2	4	0	-1.48	0.42	1.03	1.05	0.58		___	Yes
ppvt074	386	13	5	4	78	0	-0.63	0.57	0.86	0.78	0.00		___	Yes
ppvt075	386	8	9	76	8	0	-0.48	0.55	0.97	0.87	-0.32		___	Yes
ppvt076	386	60	24	10	6	0	0.36	0.60	0.96	0.97	0.46	*	__D	Yes
ppvt077	386	82	4	6	8	0	-0.90	0.55	0.92	0.81	0.00		___	Yes
ppvt078	386	23	4	66	8	0	0.05	0.49	1.09	1.11	0.54	*	__D	Yes
ppvt079	386	1	92	4	2	0	-1.96	0.45	0.97	0.83	0.45		___	Yes
ppvt080	386	51	10	30	9	0	0.81	0.62	0.95	0.96	-0.11		___	Yes
ppvt081	386	31	40	12	17	0	1.29	0.56	1.11	1.21	0.29		___	Yes
ppvt082	386	21	9	8	63	0	0.18	0.62	0.95	0.91	0.19		___	Yes
ppvt083	386	9	28	53	9	0	0.67	0.52	1.11	1.14	-0.15		___	Yes
ppvt084	386	36	4	5	54	0	0.64	0.58	1.02	1.03	1.01	*	__D	Yes
ppvt085	444	2	89	8	1	0	-1.32	0.56	0.92	0.78	-0.17		___	Yes
ppvt086	444	30	26	19	25	0	2.28	0.58	1.07	1.15	0.11		___	Yes
ppvt087	444	62	14	11	13	0	0.44	0.62	0.92	0.88	0.00		___	Yes
ppvt088	444	17	26	20	37	0	1.63	0.61	0.98	0.99	-0.29		___	Yes
ppvt089	444	18	23	32	27	0	1.91	0.53	1.14	1.29	0.32		___	Yes
ppvt090	444	63	12	14	11	0	0.41	0.67	0.87	0.81	-0.38		___	Yes
ppvt091	444	9	11	7	73	0	-0.14	0.61	0.91	0.81	0.58	*	__D	Yes
ppvt092	444	22	32	42	4	0	1.91	0.65	0.89	0.88	0.13		___	Yes
ppvt093	444	9	4	86	1	0	-1.07	0.55	0.96	0.86	-0.79	*	__D	Yes
ppvt094	444	5	66	25	5	0	0.24	0.58	1.01	0.98	0.16		___	Yes
ppvt095	444	6	7	71	16	0	-0.06	0.56	1.01	1.07	-0.37		___	Yes
ppvt096	444	73	13	5	9	0	-0.14	0.51	1.06	1.11	0.63	*	__D	Yes
ppvt097	495	28	40	14	18	0	1.74	0.57	1.02	1.10	0.59	*	__D	Yes
ppvt098	495	22	48	9	21	0	1.37	0.59	0.97	1.04	0.23		___	Yes
ppvt099	495	6	7	12	75	0	-0.04	0.57	1.07	1.08	-0.50	*	__D	Yes
ppvt100	495	2	1	97	1	0	-2.54	0.60	0.99	1.73	0.86		__C_	Yes
ppvt101	495	8	16	9	66	0	0.49	0.65	0.90	0.86	0.07		___	Yes
ppvt102	495	23	19	11	47	0	2.69	0.50	1.13	1.23	-0.08		___	Yes
ppvt103	495	4	7	83	6	0	-0.60	0.67	0.90	0.77	0.06		___	Yes
ppvt104	495	2	3	6	89	0	-1.13	0.63	0.96	0.77	-0.33		___	Yes
ppvt105	495	91	6	2	1	0	-1.35	0.62	0.98	0.95	-0.21		___	Yes
ppvt106	495	9	47	20	23	0	1.39	0.56	1.05	1.07	0.02		___	Yes
ppvt107	495	72	8	13	7	0	0.18	0.52	1.15	1.24	0.21		___	Yes
ppvt108	495	3	17	75	5	0	-0.04	0.64	0.98	0.89	-0.10		___	Yes
ppvt109	657	1	1	1	98	0	-2.80	0.62	1.05	1.18	-0.17		___	Yes
ppvt110	657	20	41	28	12	0	2.08	0.55	0.84	0.85	-0.04		___	Yes
ppvt111	657	4	28	65	3	0	0.85	0.58	0.95	0.89	0.19		___	Yes
ppvt112	657	3	15	78	4	0	0.07	0.62	0.96	0.83	-0.27		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender Male-Female		
ppvt113	657	42	8	39	11	0	2.01	0.30	1.27	1.41	-0.09		Yes	
ppvt114	657	79	4	3	14	0	0.00	0.57	1.04	1.06	0.10		Yes	
ppvt115	657	4	2	2	92	0	-1.24	0.62	0.98	1.17	-0.16		Yes	
ppvt116	657	0	92	1	6	0	-1.28	0.56	1.11	1.14	-0.36		Yes	
ppvt117	657	55	30	12	4	0	1.38	0.46	1.11	1.15	0.05		Yes	
ppvt118	657	5	89	3	3	0	-0.81	0.69	0.88	0.75	0.26		Yes	
ppvt119	657	0	0	0	100	0	-4.30	0.63	1.01	1.21	-0.71		Yes	
ppvt120	657	10	19	59	12	0	1.19	0.54	0.98	0.97	-0.13		Yes	
ppvt121	636	4	93	2	1	0	-1.40	0.65	1.00	1.30	-0.41		Yes	
ppvt122	636	8	9	21	62	0	1.08	0.64	0.86	0.81	-0.16		Yes	
ppvt123	636	24	8	51	17	0	1.64	0.55	0.91	0.89	0.09		Yes	
ppvt124	636	14	51	25	10	0	1.62	0.59	0.85	0.83	0.14		Yes	
ppvt125	636	10	25	45	20	0	1.93	0.49	0.99	0.97	0.44	*	Yes	
ppvt126	636	49	34	7	11	0	1.73	0.38	1.18	1.29	-0.06		Yes	
ppvt127	636	1	95	2	1	0	-1.71	0.69	0.97	0.75	0.21		Yes	
ppvt128	636	86	2	3	9	0	-0.48	0.70	0.91	0.91	-0.07		Yes	
ppvt129	636	12	9	17	63	0	1.06	0.56	0.98	0.97	-0.11		Yes	
ppvt130	636	81	5	4	10	0	-0.09	0.75	0.85	0.65	-0.45	*	Yes	
ppvt131	636	26	9	24	42	0	2.06	0.49	0.95	0.94	0.13		Yes	
ppvt132	636	1	2	94	3	0	-1.42	0.69	0.98	0.85	0.21		Yes	
ppvt133	570	5	24	57	14	0	1.48	0.50	1.10	1.17	-0.12		Yes	
ppvt134	570	2	91	3	3	0	-0.82	0.76	0.94	0.76	0.06		Yes	
ppvt135	570	60	9	24	6	0	1.34	0.61	0.93	0.99	-0.45	*	Yes	
ppvt136	570	4	5	3	88	0	-0.39	0.78	0.87	0.68	-0.55	*	Yes	
ppvt137	570	66	4	28	3	0	1.08	0.62	0.96	0.97	0.57	*	Yes	
ppvt138	570	56	18	14	12	0	1.53	0.55	0.98	0.97	-0.19		Yes	
ppvt139	570	1	1	1	97	0	-2.00	0.79	1.00	0.67	-0.57		Yes	
ppvt140	570	8	13	37	42	0	2.20	0.51	0.92	0.95	-0.18		Yes	
ppvt141	570	8	80	7	5	0	0.24	0.73	0.88	0.81	0.06		Yes	
ppvt142	570	28	13	56	3	0	1.57	0.59	0.89	0.86	0.29		Yes	
ppvt143	570	5	93	2	0	0	-1.10	0.78	0.96	0.71	-0.15		Yes	
ppvt144	570	1	4	91	4	0	-0.77	0.71	1.06	1.30	-0.87	*	Yes	
ppvt145	544	0	1	1	98	0	-2.41	0.83	0.98	0.50	-0.76		Yes	
ppvt146	544	24	16	47	13	0	2.04	0.52	0.96	0.97	0.17		Yes	
ppvt147	544	1	97	1	1	0	-2.06	0.82	1.00	1.10	-2.20	*	Yes	
ppvt148	544	41	14	35	10	0	2.60	0.33	1.21	1.38	-0.17		Yes	
ppvt149	544	86	5	4	6	0	-0.14	0.74	1.00	0.90	0.47		Yes	
ppvt150	544	22	24	35	20	0	2.62	0.41	1.08	1.12	0.00		Yes	
ppvt151	544	7	15	6	72	0	0.83	0.68	0.93	0.92	-0.32		Yes	
ppvt152	544	38	52	5	5	0	2.45	0.38	1.15	1.25	-0.56	*	Yes	
ppvt153	544	10	39	24	28	0	2.44	0.39	1.15	1.23	0.00		Yes	
ppvt154	544	17	44	17	22	0	2.18	0.48	1.02	1.01	0.07		Yes	
ppvt155	544	93	4	1	2	0	-0.92	0.80	0.99	0.76	-0.13		Yes	
ppvt156	544	18	2	6	74	0	0.71	0.70	0.93	0.92	-0.07		Yes	
ppvt157	502	0	96	3	1	0	-1.40	0.83	0.95	1.01	-0.28		Yes	
ppvt158	502	17	25	14	44	0	3.84	0.25	1.06	1.45	0.00		Yes	
ppvt159	502	12	57	12	19	0	3.63	0.32	0.98	1.04	0.15		Yes	
ppvt160	502	3	9	85	4	0	0.02	0.69	1.13	1.35	-1.17	*	Yes	
ppvt161	502	14	2	3	81	0	0.33	0.68	1.10	1.25	-0.49	*	Yes	
ppvt162	502	70	1	3	27	0	1.00	0.67	0.98	1.03	-0.64	*	Yes	
ppvt163	502	11	56	26	7	0	1.72	0.61	0.87	0.83	-0.24		Yes	
ppvt164	502	77	7	2	14	0	0.60	0.71	0.96	0.97	-0.57	*	Yes	
ppvt165	502	2	3	3	92	0	-0.77	0.80	0.97	0.89	0.25		Yes	
ppvt166	502	2	93	3	2	0	-0.89	0.78	1.04	1.00	-0.48		Yes	
ppvt167	502	2	2	93	2	0	-0.89	0.80	1.00	0.90	0.74		Yes	
ppvt168	502	3	4	74	19	0	0.80	0.65	1.05	1.02	-0.16		Yes	
ppvt169	486	0	2	2	96	0	-1.47	0.82	0.98	0.91	-0.59		Yes	
ppvt170	486	4	89	5	1	0	-0.33	0.79	0.93	0.81	0.06		Yes	

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male-Female						
ppvt171	486	74	10	8	9	0	0.82	0.68	0.99	0.98	0.71	*	___D	Yes
ppvt172	486	56	11	29	5	0	3.07	0.31	1.18	1.42	0.41		___	Yes
ppvt173	486	49	17	15	19	0	3.67	0.28	1.08	1.29	0.23		___	Yes
ppvt174	486	3	89	1	7	0	-0.33	0.78	0.96	0.84	0.00		___	Yes
ppvt175	486	1	2	95	2	0	-1.27	0.82	0.99	0.85	-0.34		___	Yes
ppvt176	486	11	55	8	27	0	1.81	0.55	1.01	1.00	-0.30		___	Yes
ppvt177	486	49	8	24	19	0	2.07	0.50	1.08	1.08	0.34		___	Yes
ppvt178	486	0	12	87	1	0	-0.10	0.73	1.07	1.08	-0.73	*	___D	Yes
ppvt179	486	6	7	11	77	0	0.67	0.73	0.86	0.75	0.00		___	Yes
ppvt180	486	22	8	63	7	0	3.44	0.35	0.94	0.95	0.29		___	Yes
ppvt181	467	6	16	2	76	0	0.77	0.69	0.95	0.91	0.17		___	Yes
ppvt182	467	62	13	13	12	0	1.50	0.59	1.02	1.01	0.23		___	Yes
ppvt183	467	1	5	91	4	0	-0.47	0.76	1.03	1.05	0.52		___	Yes
ppvt184	467	3	2	6	88	0	-0.20	0.75	1.02	1.05	-1.24	*	___D	Yes
ppvt185	467	24	40	7	30	0	3.41	0.36	0.99	0.95	0.00		___	Yes
ppvt186	467	4	5	88	3	0	-0.15	0.76	0.99	1.03	-0.41		___	Yes
ppvt187	467	1	97	1	1	0	-1.55	0.81	1.02	1.17	-0.31		___	Yes
ppvt188	467	19	18	15	48	0	2.16	0.46	1.15	1.31	-0.24		___	Yes
ppvt189	467	16	61	11	12	0	1.54	0.57	1.08	1.09	0.23		___	Yes
ppvt190	467	3	3	93	2	0	-0.70	0.78	1.03	1.08	0.56		___	Yes
ppvt191	467	79	14	1	6	0	0.57	0.73	0.92	0.83	-0.02		___	Yes
ppvt192	467	17	73	6	4	0	0.92	0.66	1.02	0.96	0.14		___	Yes
ppvt193	460	46	12	4	38	0	2.63	0.41	1.09	1.14	-0.47	*	___D	Yes
ppvt194	460	15	74	7	4	0	0.91	0.64	1.10	1.14	-0.31		___	Yes
ppvt195	460	8	2	86	3	0	0.03	0.77	0.93	0.80	-0.06		___	Yes
ppvt196	460	72	3	18	7	0	1.02	0.63	1.10	1.25	0.44	*	___D	Yes
ppvt197	460	62	6	14	18	0	1.55	0.66	0.80	0.75	0.30		___	Yes
ppvt198	460	1	5	2	92	0	-0.63	0.78	1.00	0.95	-0.26		___	Yes
ppvt199	460	4	7	80	10	0	0.55	0.68	1.08	1.09	0.39		___	Yes
ppvt200	460	73	2	3	22	0	0.95	0.68	0.94	0.86	-0.16		___	Yes
ppvt201	460	29	28	4	38	0	2.62	0.41	1.12	1.17	0.09		___	Yes
ppvt202	460	3	95	2	1	0	-1.03	0.80	0.99	0.98	-0.88	*	___D	Yes
ppvt203	460	4	42	53	2	0	1.94	0.58	0.90	0.88	0.34		___	Yes
ppvt204	460	28	58	4	10	0	1.71	0.59	0.94	0.92	-0.08		___	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 27. Item statistics by item in the Peabody Picture Vocabulary Test for the Older Cohort, India – Telugu

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
		Male-Female												
ppvt001	109	2	0	0	98	0	-4.31	0.08	0.91	1.01	1.01		A__	No
ppvt002	109	1	0	99	0	0	-5.01	0.02	1.01	1.90	0.34		A_C_	No
ppvt003	109	100	0	0	0	0	-6.23		0.00	0.00	0.00		_BC_	No
ppvt004	109	98	0	1	1	0	-4.31	0.12	0.90	0.40	1.01		_C_	Yes
ppvt005	109	0	0	20	80	0	-1.61	0.24	1.06	1.15	-0.38		___	Yes
ppvt006	109	100	0	0	0	0	-6.23		0.00	0.00	0.00		_BC_	No
ppvt007	109	4	94	2	0	0	-3.15	0.12	1.04	1.16	-1.72		___	Yes
ppvt008	109	19	8	66	6	0	-0.84	0.34	0.97	0.98	0.05		___	Yes
ppvt009	109	0	1	3	96	0	-3.58	0.13	0.99	0.66	-0.94		___	Yes
ppvt010	109	1	78	9	12	0	-1.49	0.29	0.88	0.81	0.65		___	Yes
ppvt011	109	0	98	2	0	0	-4.31	0.03	1.04	2.49	1.01		A_C_	No
ppvt012	109	0	6	93	1	0	-2.83	0.16	0.94	0.88	-0.41		___	Yes
ppvt013	205	5	90	2	3	0	-2.14	0.26	0.90	0.69	1.71	*	_D	Yes
ppvt014	205	96	0	3	0	0	-3.08	0.18	1.00	0.81	0.55		___	Yes
ppvt015	205	0	1	99	0	0	-4.23	0.09	1.00	0.92	-1.48		A__	No
ppvt016	205	97	1	0	1	0	-3.35	0.17	0.90	0.44	2.46		_C_	Yes
ppvt017	205	2	2	95	0	0	-2.97	0.20	0.93	0.61	0.13		___	Yes
ppvt018	205	6	82	3	9	0	-1.44	0.32	0.91	0.92	0.05		___	Yes
ppvt019	205	87	1	7	4	0	-1.88	0.25	0.97	0.95	0.59		___	Yes
ppvt020	205	9	13	6	72	0	-0.83	0.43	0.84	0.78	-0.19		___	Yes
ppvt021	205	12	15	61	13	0	-0.27	0.42	0.99	0.97	0.66	*	_D	Yes
ppvt022	205	11	1	4	83	0	-1.51	0.26	1.00	1.16	-0.67		___	Yes
ppvt023	205	21	72	3	4	0	-0.80	0.34	1.00	1.07	0.20		___	Yes
ppvt024	205	16	9	17	58	0	-0.14	0.41	1.11	1.13	-0.02		___	Yes
ppvt025	347	99	1	0	0	0	-4.27	0.13	0.92	0.46	1.80		_C_	Yes
ppvt026	347	1	1	98	1	0	-3.55	0.17	0.90	0.50	0.32		_C_	Yes
ppvt027	347	9	16	23	52	0	0.35	0.53	1.00	1.04	-0.56	*	_D	Yes
ppvt028	347	0	1	1	98	0	-3.55	0.15	0.94	0.81	0.32		___	Yes
ppvt029	347	80	2	3	15	0	-1.08	0.28	1.10	1.27	0.31		___	Yes
ppvt030	347	0	60	39	0	0	-0.03	0.42	1.12	1.14	0.00		___	Yes
ppvt031	347	2	7	89	2	0	-1.82	0.31	0.93	0.82	-0.48		___	Yes
ppvt032	347	13	68	9	10	0	-0.42	0.40	1.08	1.11	-0.30		___	Yes
ppvt033	347	2	91	3	5	0	-2.05	0.28	0.91	0.77	0.79		___	Yes
ppvt034	347	73	9	13	6	0	-0.65	0.42	0.97	0.94	0.34		___	Yes
ppvt035	347	5	8	5	81	0	-1.18	0.40	0.92	0.81	-1.06	*	_D	Yes
ppvt036	347	1	11	87	1	0	-1.64	0.28	1.01	1.01	0.03		___	Yes
ppvt037	362	21	64	9	5	0	-0.20	0.45	1.05	1.03	0.16		___	Yes
ppvt038	362	0	0	1	99	0	-4.29	0.09	0.94	0.89	1.80		A__	No
ppvt039	362	1	19	78	2	0	-0.91	0.41	0.97	0.90	-0.15		___	Yes
ppvt040	362	29	3	12	56	0	1.44	0.66	0.93	0.90	0.32		___	Yes
ppvt041	362	13	80	2	6	0	-1.06	0.43	0.89	0.80	-0.77	*	_D	Yes
ppvt042	362	12	17	11	59	0	2.55	0.65	1.06	1.55	0.38		_C_	Yes
ppvt043	362	23	17	44	16	0	0.72	0.52	1.09	1.16	-0.23		___	Yes
ppvt044	362	18	33	10	38	0	0.98	0.54	1.11	1.14	0.39		___	Yes
ppvt045	362	44	17	9	30	0	1.36	0.57	1.11	1.19	-0.47		___	Yes
ppvt046	362	4	51	40	4	0	0.41	0.58	0.96	0.94	-0.24		___	Yes
ppvt047	362	2	1	96	1	0	-3.07	0.22	0.89	0.54	0.45		___	Yes
ppvt048	362	92	5	2	2	0	-2.15	0.23	1.02	1.03	0.28		___	Yes
ppvt049	380	44	21	21	14	0	1.93	0.66	1.00	1.06	0.40		___	Yes
ppvt050	380	41	52	4	2	0	0.91	0.57	1.04	1.04	0.02		___	Yes
ppvt051	380	60	25	8	7	0	0.07	0.61	0.85	0.82	-0.25		___	Yes
ppvt052	380	6	88	3	3	0	-1.66	0.34	0.96	1.00	0.09		___	Yes
ppvt053	380	51	12	21	16	0	2.30	0.67	1.02	1.08	0.48		___	Yes
ppvt054	380	14	54	21	10	0	0.31	0.50	1.09	1.09	0.65	*	_D	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt055	380	78	20	2	1	0	-0.86	0.47	0.91	0.84	0.00		___	Yes
ppvt056	380	6	28	8	58	0	0.13	0.56	0.91	0.88	0.04		___	Yes
ppvt057	380	26	51	12	12	0	0.46	0.49	1.10	1.10	0.45		___	Yes
ppvt058	380	13	7	57	24	0	0.20	0.51	1.02	1.02	-0.10		___	Yes
ppvt059	380	6	2	91	2	0	-1.95	0.27	0.99	1.07	-0.07		___	Yes
ppvt060	380	7	4	6	83	0	-1.24	0.35	1.02	1.03	-0.12		___	Yes
ppvt061	426	6	8	8	79	0	-0.82	0.57	0.80	0.69	0.54	*	___D	Yes
ppvt062	426	55	25	13	7	0	0.39	0.56	1.00	0.99	0.23		___	Yes
ppvt063	426	4	78	10	9	0	-0.78	0.38	1.09	1.24	-0.53	*	___D	Yes
ppvt064	426	1	1	96	1	0	-2.93	0.36	0.88	0.61	0.31		___	Yes
ppvt065	426	3	12	50	35	0	1.29	0.56	1.11	1.31	0.39		___	Yes
ppvt066	426	11	15	31	43	0	1.47	0.56	1.15	1.38	-0.45		___	Yes
ppvt067	426	11	44	12	32	0	0.87	0.49	1.19	1.30	0.33		___	Yes
ppvt068	426	92	3	3	3	0	-2.00	0.40	0.93	0.79	1.47	*	___D	Yes
ppvt069	426	6	0	8	85	0	-1.30	0.35	1.05	1.32	-0.17		___	Yes
ppvt070	426	4	92	3	2	0	-2.04	0.41	0.90	0.69	0.53		___	Yes
ppvt071	426	2	8	87	3	0	-1.47	0.40	0.99	1.01	-0.06		___	Yes
ppvt072	426	52	17	8	23	0	0.53	0.62	0.91	0.89	-0.16		___	Yes
ppvt073	553	8	43	44	5	0	1.11	0.45	1.18	1.31	0.73	*	___D	Yes
ppvt074	553	1	1	0	97	0	-2.98	0.39	0.90	0.51	-0.85		___	Yes
ppvt075	553	13	2	82	3	0	-0.90	0.46	0.96	0.90	-0.44		___	Yes
ppvt076	553	50	11	33	6	0	0.81	0.62	0.90	0.88	0.80	*	___D	Yes
ppvt077	553	88	5	6	2	0	-1.36	0.43	0.99	0.93	0.69	*	___D	Yes
ppvt078	553	0	3	95	1	0	-2.46	0.39	0.94	0.79	-1.10	*	___D	Yes
ppvt079	553	4	78	12	6	0	-0.59	0.36	1.08	1.21	-0.29		___	Yes
ppvt080	553	62	7	26	5	0	0.24	0.57	0.91	0.89	0.26		___	Yes
ppvt081	553	29	51	11	9	0	0.74	0.52	1.05	1.04	0.00		___	Yes
ppvt082	553	9	5	5	80	0	-0.72	0.52	0.90	0.81	0.18		___	Yes
ppvt083	553	3	10	75	11	0	-0.44	0.49	0.96	0.89	0.00		___	Yes
ppvt084	553	7	22	12	60	0	0.33	0.35	1.23	1.32	0.88	*	___D	Yes
ppvt085	594	28	37	20	15	0	1.44	0.49	1.11	1.17	-0.28		___	Yes
ppvt086	594	34	14	24	28	0	1.90	0.54	1.07	1.11	-0.13		___	Yes
ppvt087	594	91	2	3	3	0	-1.73	0.35	0.99	1.22	-0.11		___	Yes
ppvt088	594	5	8	8	79	0	-0.63	0.55	0.86	0.79	0.27		___	Yes
ppvt089	594	15	32	30	23	0	1.81	0.48	1.13	1.23	0.00		___	Yes
ppvt090	594	71	13	5	10	0	-0.16	0.52	0.93	0.88	0.08		___	Yes
ppvt091	594	12	14	7	67	0	0.05	0.44	1.03	1.04	-0.11		___	Yes
ppvt092	594	6	47	45	2	0	0.97	0.56	0.97	1.00	-0.31		___	Yes
ppvt093	594	2	1	96	0	0	-2.66	0.40	0.92	0.83	0.00		___	Yes
ppvt094	594	14	44	37	6	0	1.15	0.43	1.16	1.21	-0.13		___	Yes
ppvt095	594	4	6	84	6	0	-0.94	0.45	0.97	0.91	0.16		___	Yes
ppvt096	594	79	9	5	7	0	-0.63	0.52	0.91	0.82	-0.16		___	Yes
ppvt097	622	14	55	9	21	0	0.69	0.56	0.96	0.95	0.79	*	___D	Yes
ppvt098	622	8	49	2	41	0	0.98	0.46	1.12	1.17	-0.20		___	Yes
ppvt099	622	0	6	18	76	0	-0.34	0.50	0.97	0.96	-0.55	*	___D	Yes
ppvt100	622	32	12	30	26	0	1.89	0.64	0.91	0.92	0.06		___	Yes
ppvt101	622	2	10	3	86	0	-1.09	0.49	0.92	0.94	-0.17		___	Yes
ppvt102	622	16	3	9	72	0	2.79	0.54	1.07	1.34	0.05		___	Yes
ppvt103	622	21	38	29	12	0	1.97	0.57	1.01	1.09	0.21		___	Yes
ppvt104	622	0	3	2	94	0	-2.11	0.41	0.98	1.41	-0.80	*	___D	Yes
ppvt105	622	81	10	4	5	0	-0.67	0.52	0.92	0.91	0.11		___	Yes
ppvt106	622	9	69	12	9	0	0.02	0.52	0.98	0.97	0.75	*	___D	Yes
ppvt107	622	38	10	46	6	0	1.48	0.54	1.03	1.11	0.12		___	Yes
ppvt108	622	4	17	64	5	10	0.29	0.50	1.04	1.06	0.08		___	Yes
ppvt109	764	1	0	1	98	0	-3.00	0.58	0.97	1.24	-0.47		___	Yes
ppvt110	764	18	80	1	1	0	-0.23	0.52	1.03	1.04	-0.54	*	___D	Yes
ppvt111	764	2	16	74	9	0	0.16	0.58	0.94	0.91	-0.05		___	Yes
ppvt112	764	3	13	68	16	0	0.48	0.60	0.90	0.82	-0.09		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
		Male-Female												
ppvt113	764	43	13	33	11	0	1.67	0.39	1.08	1.14	-0.21		___	Yes
ppvt114	764	60	5	12	22	0	0.86	0.55	0.95	0.94	0.37	*	___D	Yes
ppvt115	764	6	1	0	93	0	-1.49	0.55	1.03	1.10	-0.05		___	Yes
ppvt116	764	2	80	4	14	0	-0.21	0.62	0.91	0.92	0.06		___	Yes
ppvt117	764	46	23	26	5	0	1.52	0.54	0.91	0.90	0.60	*	___D	Yes
ppvt118	764	6	82	7	5	0	-0.35	0.66	0.87	0.70	0.74	*	___D	Yes
ppvt119	764	4	10	5	82	0	-0.36	0.66	0.85	0.73	0.06		___	Yes
ppvt120	764	5	12	51	25	8	1.33	0.53	0.94	0.91	-0.05		___	Yes
ppvt121	702	9	75	13	3	0	0.25	0.60	0.98	0.99	0.26		___	Yes
ppvt122	702	5	12	18	65	0	0.75	0.52	1.01	1.02	-0.41	*	___D	Yes
ppvt123	702	7	4	81	7	0	-0.17	0.68	0.92	0.82	0.08		___	Yes
ppvt124	702	47	13	33	7	0	3.63	0.21	1.04	1.24	-0.07		___	Yes
ppvt125	702	7	17	67	9	0	0.68	0.51	1.05	1.12	0.00		___	Yes
ppvt126	702	99	0	0	1	0	-3.24	0.72	1.01	1.55	-0.25		___C	Yes
ppvt127	702	30	21	19	30	0	2.97	0.35	0.98	0.98	0.39		___	Yes
ppvt128	702	96	0	3	1	0	-1.90	0.71	0.98	1.09	-1.46	*	___D	Yes
ppvt129	702	9	1	10	79	0	-0.03	0.65	0.94	0.87	-0.11		___	Yes
ppvt130	702	73	11	11	6	0	0.38	0.60	0.97	1.01	0.35		___	Yes
ppvt131	702	12	1	3	84	0	-0.36	0.62	1.02	1.05	-0.20		___	Yes
ppvt132	702	3	2	91	1	2	-1.07	0.68	0.98	1.00	-0.39		___	Yes
ppvt133	686	0	2	97	1	0	-2.14	0.74	0.99	0.99	-0.24		___	Yes
ppvt134	686	42	28	13	16	0	2.56	0.45	0.91	0.88	-0.16		___	Yes
ppvt135	686	43	7	48	1	0	1.82	0.39	1.12	1.12	-0.77	*	___D	Yes
ppvt136	686	7	3	6	84	0	-0.31	0.62	1.04	1.16	-0.21		___	Yes
ppvt137	686	75	3	20	2	0	0.26	0.58	1.05	1.11	0.19		___	Yes
ppvt138	686	42	19	14	25	0	1.85	0.46	1.00	1.01	0.00		___	Yes
ppvt139	686	61	11	8	21	0	3.01	0.29	1.05	1.15	-0.30		___	Yes
ppvt140	686	0	12	29	59	0	1.07	0.52	1.04	1.05	-0.05		___	Yes
ppvt141	686	10	57	15	18	0	1.16	0.54	1.01	0.99	-0.22		___	Yes
ppvt142	686	17	12	68	4	0	0.66	0.58	1.01	1.03	0.00		___	Yes
ppvt143	686	26	70	2	2	0	0.55	0.63	0.94	0.86	-0.28		___	Yes
ppvt144	686	7	12	50	18	13	1.51	0.60	0.87	0.85	0.14		___	Yes
ppvt145	595	2	2	2	95	0	-1.48	0.79	0.98	0.85	0.00		___	Yes
ppvt146	595	12	17	58	14	0	1.28	0.53	1.06	1.06	-0.15		___	Yes
ppvt147	595	2	94	1	2	0	-1.38	0.77	0.99	0.95	-1.05	*	___D	Yes
ppvt148	595	14	42	32	12	0	2.46	0.36	1.10	1.16	-0.31		___	Yes
ppvt149	595	61	32	5	3	0	1.11	0.57	1.02	1.07	0.18		___	Yes
ppvt150	595	9	3	85	3	0	-0.24	0.71	0.99	1.00	-0.50	*	___D	Yes
ppvt151	595	4	32	3	61	0	1.15	0.51	1.13	1.16	-0.62	*	___D	Yes
ppvt152	595	31	22	23	24	0	2.52	0.40	1.01	1.04	0.46	*	___D	Yes
ppvt153	595	2	87	6	5	0	-0.43	0.71	1.01	1.03	-0.15		___	Yes
ppvt154	595	1	66	1	33	0	0.90	0.64	0.96	0.96	0.29		___	Yes
ppvt155	595	84	8	1	6	0	-0.22	0.73	0.95	0.89	-0.06		___	Yes
ppvt156	595	16	5	2	74	3	0.46	0.68	0.95	0.98	-0.44	*	___D	Yes
ppvt157	577	1	97	2	1	0	-2.05	0.80	0.97	0.76	-0.27		___	Yes
ppvt158	577	17	45	20	19	0	3.46	0.24	1.08	1.33	-0.44		___	Yes
ppvt159	577	9	33	17	41	0	2.06	0.46	1.04	1.06	0.13		___	Yes
ppvt160	577	9	11	53	27	0	1.54	0.60	0.90	0.88	-0.04		___	Yes
ppvt161	577	24	1	2	73	0	0.58	0.62	1.06	1.17	0.47	*	___D	Yes
ppvt162	577	49	8	9	34	0	1.69	0.42	1.18	1.23	-0.60	*	___D	Yes
ppvt163	577	24	37	9	30	0	2.24	0.48	0.96	0.94	-0.05		___	Yes
ppvt164	577	87	5	2	6	0	-0.44	0.71	1.02	1.09	-0.36		___	Yes
ppvt165	577	13	7	11	69	0	0.77	0.68	0.91	0.86	-0.02		___	Yes
ppvt166	577	3	90	5	2	0	-0.75	0.75	1.01	0.95	0.28		___	Yes
ppvt167	577	2	1	96	1	0	-1.83	0.78	1.01	1.18	0.58		___	Yes
ppvt168	577	9	8	59	20	5	1.27	0.59	0.98	0.96	-0.09		___	Yes
ppvt169	548	15	21	22	43	0	2.03	0.54	0.93	0.92	-0.14		___	Yes
ppvt170	548	17	34	24	26	0	2.45	0.47	0.94	0.94	0.00		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male-Female						
ppvt171	548	43	18	9	30	0	2.01	0.48	1.07	1.08	0.18		___	Yes
ppvt172	548	24	32	33	11	0	2.50	0.45	0.97	0.97	0.19		___	Yes
ppvt173	548	25	18	44	13	0	3.78	0.30	0.93	0.91	0.39		___	Yes
ppvt174	548	15	30	44	10	0	2.62	0.41	0.99	1.03	0.05		___	Yes
ppvt175	548	13	13	63	11	0	1.11	0.65	0.93	0.91	-0.17		___	Yes
ppvt176	548	11	52	6	31	0	1.62	0.57	0.97	0.96	-0.20		___	Yes
ppvt177	548	24	1	20	55	0	2.98	0.27	1.16	1.39	-1.15	*	___D	Yes
ppvt178	548	2	13	80	5	0	0.18	0.75	0.88	0.76	0.39		___	Yes
ppvt179	548	3	3	2	91	0	-0.86	0.77	0.96	0.98	0.79	*	___D	Yes
ppvt180	548	47	3	16	2	33	1.84	0.60	0.89	0.87	0.11		___	Yes
ppvt181	360	2	3	28	67	0	1.21	0.62	0.97	0.99	-0.07		___	Yes
ppvt182	360	15	63	20	2	0	3.90	0.28	0.98	1.00	-0.17		___	Yes
ppvt183	360	8	4	81	8	0	0.46	0.66	1.01	1.00	0.89	*	___D	Yes
ppvt184	360	0	0	2	98	0	-1.85	0.75	0.99	0.84	1.32		___	Yes
ppvt185	360	73	10	7	10	0	0.92	0.64	0.95	0.92	-0.20		___	Yes
ppvt186	360	1	1	98	1	0	-1.98	0.74	1.02	1.31	-1.11		___	Yes
ppvt187	360	48	29	11	13	0	2.99	0.41	0.94	0.90	0.57	*	___D	Yes
ppvt188	360	20	31	21	28	0	3.05	0.37	1.01	1.05	-0.30		___	Yes
ppvt189	360	12	60	14	14	0	1.54	0.54	1.08	1.10	-0.85	*	___D	Yes
ppvt190	360	4	19	73	4	0	0.91	0.63	1.01	1.00	-0.24		___	Yes
ppvt191	360	84	5	3	8	0	0.21	0.69	0.98	0.98	0.35		___	Yes
ppvt192	360	2	91	3	1	3	-0.49	0.73	0.95	0.99	-0.09		___	Yes
ppvt193	348	39	13	24	25	0	3.23	0.34	1.04	1.10	-0.13		___	Yes
ppvt194	348	19	59	7	15	0	1.61	0.55	1.05	1.05	0.00		___	Yes
ppvt195	348	5	3	91	2	0	-0.42	0.70	1.03	1.53	0.66		___C	Yes
ppvt196	348	60	5	5	30	0	1.59	0.54	1.08	1.11	0.00		___	Yes
ppvt197	348	63	8	11	18	0	1.46	0.59	1.00	1.01	0.12		___	Yes
ppvt198	348	10	7	11	72	0	1.02	0.63	0.95	0.88	0.29		___	Yes
ppvt199	348	13	9	67	11	0	1.27	0.60	0.98	1.00	0.19		___	Yes
ppvt200	348	40	9	32	20	0	2.49	0.47	0.92	0.91	-0.19		___	Yes
ppvt201	348	26	54	4	15	0	3.93	0.27	1.03	1.10	0.27		___	Yes
ppvt202	348	5	79	7	9	0	0.61	0.64	1.07	1.21	-0.52		___	Yes
ppvt203	348	6	28	41	25	0	2.45	0.46	0.99	1.01	0.02		___	Yes
ppvt204	348	29	33	11	27	0	2.80	0.41	1.00	1.01	-0.05		___	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 28. Item statistics by item in the Peabody Picture Vocabulary Test for the Older Cohort, Peru – Spanish

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt001	2	0	100	0	0	0	-6.46	c	c	c	c		Yes	
ppvt002	2	0	0	0	100	0	-6.46	c	c	c	c		Yes	
ppvt003	2	100	0	0	0	0	-6.46	c	c	c	c		Yes	
ppvt004	3	0	100	0	0	0	-6.76	c	c	c	c		Yes	
ppvt005	3	67	33	0	0	0	-4.83	0.09	3.43	9.90	c	ABC_	Yes	
ppvt006	3	0	0	0	100	0	-6.76	c	c	c	c		Yes	
ppvt007	3	67	33	0	0	0	-4.83	0.19	0.43	0.30	c	_BC_	Yes	
ppvt008	3	67	33	0	0	0	-4.83	0.19	0.43	0.30	c	_BC_	Yes	
ppvt009	4	0	0	0	100	0	-6.81	c	c	c	c		Yes	
ppvt010	4	25	75	0	0	0	-5.00	0.19	0.40	0.25	c	_BC_	Yes	
ppvt011	4	0	0	100	0	0	-6.81	c	c	c	c		Yes	
ppvt012	4	25	25	25	25	0	-1.45	0.27	1.92	1.19	c	_B_	Yes	
ppvt013	4	0	0	75	25	0	-5.00	0.19	0.40	0.25	c	_BC_	Yes	
ppvt014	4	0	0	100	0	0	-6.81	c	c	c	c		Yes	
ppvt015	4	0	100	0	0	0	-6.81	c	c	c	c		Yes	
ppvt016	4	25	75	0	0	0	-5.00	0.18	1.89	1.17	c	_B_	Yes	
ppvt017	4	50	50	0	0	0	-3.23	0.23	1.59	1.11	c	_B_	Yes	
ppvt018	4	25	50	0	25	0	-1.45	0.27	1.92	1.19	c	_B_	Yes	
ppvt019	4	0	50	0	50	0	-3.23	0.26	0.37	0.28	c	_BC_	Yes	
ppvt020	4	25	0	75	0	0	0.37	0.30	c	c	c		Yes	
ppvt021	4	25	0	0	75	0	-5.00	0.18	1.89	1.17	c	_B_	Yes	
ppvt022	5	0	0	100	0	0	-6.82	c	c	c	c		Yes	
ppvt023	5	20	80	0	0	0	-5.03	0.19	0.39	0.21	c	_BC_	Yes	
ppvt024	5	20	0	20	60	0	-3.37	0.26	0.34	0.25	c	_BC_	Yes	
ppvt025	4	0	0	50	50	0	-2.01	0.30	0.41	0.36	c	_BC_	Yes	
ppvt026	3	0	0	67	33	0	-2.22	0.30	0.49	0.44	c	_BC_	Yes	
ppvt027	4	0	0	100	0	0	-4.14	0.26	c	c	c		Yes	
ppvt028	4	0	25	25	50	0	-1.42	0.30	0.98	0.95	c		Yes	
ppvt029	5	0	100	0	0	0	-4.71	0.26	c	c	c		Yes	
ppvt030	6	0	17	0	83	0	-3.38	0.27	1.40	1.94	c	_C_	Yes	
ppvt031	6	0	0	100	0	0	-4.79	0.26	c	c	c		Yes	
ppvt032	6	17	83	0	0	0	-3.38	0.27	1.43	2.32	c	_C_	Yes	
ppvt033	6	17	17	0	67	0	-2.35	0.30	1.21	1.30	c		Yes	
ppvt034	7	0	0	0	100	0	-4.86	0.26	c	c	c		Yes	
ppvt035	8	63	25	0	13	0	-1.84	0.31	1.14	1.07	c		Yes	
ppvt036	10	0	0	0	100	0	-4.97	0.26	c	c	c		Yes	
ppvt037	15	80	0	7	13	0	-2.34	0.33	0.63	0.48	c	_C_	Yes	
ppvt038	16	0	88	6	6	0	-2.96	0.27	1.32	1.16	c		Yes	
ppvt039	16	13	0	88	0	0	-2.96	0.32	0.53	0.26	c	_C_	Yes	
ppvt040	19	0	5	95	0	0	-3.86	0.29	0.72	0.20	c	_C_	Yes	
ppvt041	20	0	0	95	5	0	-3.92	0.27	1.15	1.05	c		Yes	
ppvt042	22	14	82	0	5	0	-2.26	0.32	0.91	0.67	c		Yes	
ppvt043	24	79	13	4	4	0	-2.10	0.33	0.87	1.05	c		Yes	
ppvt044	26	0	73	19	8	0	-1.69	0.36	0.82	0.69	c		Yes	
ppvt045	28	64	7	18	11	0	-1.00	0.35	1.22	1.29	c		Yes	
ppvt046	32	3	6	6	84	0	-2.19	0.30	1.31	1.83	1.60	_C_	Yes	
ppvt047	36	0	89	3	8	0	-2.54	0.32	1.04	0.72	1.31		Yes	
ppvt048	42	5	86	10	0	0	-2.16	0.31	1.24	1.41	0.83		Yes	
ppvt049	43	88	9	0	2	0	-2.40	0.34	0.88	0.76	1.59		Yes	
ppvt050	62	6	6	79	8	0	-1.41	0.39	0.91	0.79	0.75		Yes	
ppvt051	63	5	3	5	87	0	-2.07	0.34	1.06	0.88	0.02		Yes	
ppvt052	63	3	84	13	0	0	-1.78	0.37	0.94	0.88	0.05		Yes	
ppvt053	75	4	11	79	7	0	-1.22	0.39	0.96	0.97	-1.19		Yes	
ppvt054	91	3	9	87	1	0	-1.72	0.39	0.86	0.64	-0.40		Yes	

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item- fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt055	95	2	0	6	92	0	-2.26	0.33	1.02	1.01	-0.57		___	Yes
ppvt056	94	6	10	10	74	0	-0.71	0.45	0.86	0.86	-0.57		___	Yes
ppvt057	95	94	2	1	2	1	-2.51	0.36	0.94	0.65	0.20		___	Yes
ppvt058	95	12	29	51	8	0	0.57	0.52	0.97	0.93	-0.38		___	Yes
ppvt059	105	98	2	0	0	0	-3.72	0.33	0.99	2.74	0.13		__C_	Yes
ppvt060	107	93	2	2	4	0	-2.23	0.37	0.90	0.81	0.13		___	Yes
ppvt061	116	10	32	9	48	1	0.71	0.56	0.96	0.92	-0.02		___	Yes
ppvt062	116	50	31	9	9	0	0.66	0.48	1.39	1.58	0.22		__C_	Yes
ppvt063	119	8	9	6	76	0	-0.66	0.44	1.05	1.18	-0.85		___	Yes
ppvt064	164	4	95	0	1	0	-2.38	0.38	0.96	0.96	-1.14		___	Yes
ppvt065	166	0	1	99	1	0	-3.85	0.36	0.99	0.65	0.06	*	__D	Yes
ppvt066	167	2	5	87	5	1	-1.26	0.40	1.03	1.44	-1.08	*	__D	Yes
ppvt067	206	73	7	18	1	0	-0.01	0.50	1.05	1.00	0.76		___	Yes
ppvt068	208	8	84	4	4	0	-0.75	0.44	1.08	1.11	-0.45		___	Yes
ppvt069	238	4	8	13	74	0	0.08	0.57	0.81	0.65	-0.22		___	Yes
ppvt070	256	2	93	3	2	0	-1.66	0.37	1.09	1.58	-0.44		__C_	Yes
ppvt071	263	2	10	87	2	0	-0.78	0.51	0.80	0.54	0.05		___	Yes
ppvt072	281	56	19	12	14	0	1.30	0.59	1.02	1.08	-0.24		___	Yes
ppvt073	311	1	4	3	92	0	-1.19	0.45	0.93	0.66	-0.42		___	Yes
ppvt074	313	89	5	4	3	0	-0.81	0.49	0.89	0.53	0.19		___	Yes
ppvt075	324	20	14	48	17	1	1.85	0.64	0.97	1.10	0.22		___	Yes
ppvt076	334	10	81	3	6	0	0.03	0.55	0.91	0.76	-0.56		___	Yes
ppvt077	340	8	12	64	16	0	1.18	0.57	1.07	1.14	0.29		___	Yes
ppvt078	342	16	54	23	6	1	1.73	0.60	1.07	1.13	0.33		___	Yes
ppvt079	352	8	7	83	2	0	0.03	0.53	0.98	0.95	0.05		___	Yes
ppvt080	353	11	10	10	67	1	1.13	0.49	1.22	1.52	-0.48	*	__CD	Yes
ppvt081	365	20	45	27	8	0	2.32	0.56	1.22	1.29	-0.74		___	Yes
ppvt082	372	1	2	88	9	0	-0.23	0.55	0.98	1.04	0.57		___	Yes
ppvt083	394	9	69	8	14	0	1.31	0.64	0.92	0.92	0.37		___	Yes
ppvt084	402	60	5	20	16	0	1.87	0.63	0.99	1.00	-0.16		___	Yes
ppvt085	410	6	18	69	7	1	1.46	0.57	1.13	1.19	0.19		___	Yes
ppvt086	403	11	75	8	6	0	1.15	0.68	0.87	0.82	0.10		___	Yes
ppvt087	402	9	57	12	21	0	2.15	0.67	0.93	0.91	0.35		___	Yes
ppvt088	402	4	72	16	8	0	1.43	0.69	0.88	0.77	0.12		___	Yes
ppvt089	394	19	20	5	56	0	2.35	0.66	0.97	0.94	-0.14		___	Yes
ppvt090	405	4	4	31	60	1	2.22	0.67	0.97	0.94	-0.27		___	Yes
ppvt091	390	16	7	25	49	2	2.83	0.64	1.01	1.00	-0.02		___	Yes
ppvt092	378	3	14	52	30	0	2.75	0.68	0.95	0.96	-0.25		___	Yes
ppvt093	401	7	13	14	65	0	2.24	0.66	1.03	1.03	0.00	*	__D	Yes
ppvt094	410	16	3	6	75	0	1.80	0.74	0.88	0.81	0.69		___	Yes
ppvt095	399	0	1	8	91	0	0.45	0.74	1.01	1.02	0.20		___	Yes
ppvt096	399	10	13	19	58	2	2.76	0.64	1.01	1.00	0.43		___	Yes
ppvt097	392	12	64	17	8	0	2.52	0.65	1.06	1.07	-0.05		___	Yes
ppvt098	390	17	60	1	22	0	2.73	0.61	1.13	1.17	-0.36		___	Yes
ppvt099	387	1	2	91	5	0	0.64	0.80	0.88	0.65	-0.37	*	__D	Yes
ppvt100	386	4	4	88	3	0	1.04	0.78	0.96	0.81	-0.74		___	Yes
ppvt101	386	34	24	37	6	0	4.02	0.57	0.98	0.99	0.21		___	Yes
ppvt102	392	27	23	26	23	1	4.66	0.52	0.95	1.03	0.18		___	Yes
ppvt103	390	61	20	5	13	0	2.78	0.70	0.90	0.88	-0.19		___	Yes
ppvt104	381	75	14	7	4	0	2.14	0.70	1.02	1.04	0.29	*	__D	Yes
ppvt105	378	37	2	15	46	1	3.97	0.57	0.97	0.97	0.70		___	Yes
ppvt106	371	21	60	4	16	0	2.99	0.60	1.11	1.11	-0.26		___	Yes
ppvt107	358	32	57	9	2	0	3.15	0.65	0.96	0.96	-0.20		___	Yes
ppvt108	344	14	9	73	3	0	2.43	0.73	0.89	0.80	-0.07	*	__D	Yes
ppvt109	321	9	13	60	18	0	3.18	0.65	0.95	0.93	0.60		___	Yes
ppvt110	318	17	28	33	20	2	4.43	0.49	1.06	1.11	0.09		___	Yes
ppvt111	315	4	5	6	85	0	1.75	0.75	0.95	0.83	-0.47	*	__D	Yes
ppvt112	312	26	61	6	8	0	3.25	0.63	0.99	1.00	-0.95	*	__D	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male-Female						
ppvt113	301	60	23	11	6	0	3.33	0.63	0.97	0.94	-0.66		_____	Yes
ppvt114	309	9	39	6	45	1	4.34	0.48	1.06	1.08	0.00		_____	Yes
ppvt115	301	5	25	47	23	0	4.01	0.55	0.97	0.96	0.17		_____	Yes
ppvt116	294	12	9	51	28	0	3.83	0.55	1.05	1.03	-0.03		_____	Yes
ppvt117	287	2	86	6	6	0	1.88	0.74	0.93	0.78	0.38		_____	Yes
ppvt118	285	19	17	41	22	1	4.41	0.49	0.99	0.97	0.38		_____	Yes
ppvt119	280	73	5	6	16	0	2.88	0.65	0.99	0.97	0.44		_____	Yes
ppvt120	266	16	14	41	29	1	4.50	0.46	1.11	1.12	0.43		_____	Yes
ppvt121	260	13	76	10	1	0	2.79	0.65	0.99	1.03	0.37		_____	Yes
ppvt122	254	19	24	29	27	0	5.14	0.36	1.17	1.28	-0.19		_____	Yes
ppvt123	247	64	8	16	13	0	3.49	0.57	1.07	1.03	0.05		_____	Yes
ppvt124	243	5	3	50	42	0	4.49	0.48	1.00	0.98	-0.37		_____	Yes
ppvt125	240	40	8	24	26	2	5.38	0.34	1.13	1.30	0.43		_____	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Table 29. Item statistics by item in the Peabody Picture Vocabulary Test for the Older Cohort, Vietnam – Tieng Viet Nam

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt001	5	0	0	0	100	0	-4.50	c	c	c	c		Yes	
ppvt002	5	0	0	100	0	0	-4.50	c	c	c	c		Yes	
ppvt003	5	100	0	0	0	0	-4.50	c	c	c	c		Yes	
ppvt004	5	60	0	20	20	0	-2.07	0.20	0.92	0.83	c		Yes	
ppvt005	5	0	0	20	80	0	-3.11	0.14	0.95	0.80	c		Yes	
ppvt006	5	80	0	0	20	0	-3.11	0.14	0.97	0.83	c		Yes	
ppvt007	5	0	100	0	0	0	-4.50	c	c	c	c		Yes	
ppvt008	5	20	20	60	0	0	-2.07	0.20	0.90	0.81	c		Yes	
ppvt009	5	20	20	20	40	0	-1.16	0.25	0.77	0.75	c		Yes	
ppvt010	5	0	40	20	40	0	-1.16	0.25	0.81	0.81	c		Yes	
ppvt011	5	60	40	0	0	0	-1.16	0.25	0.77	0.75	c		Yes	
ppvt012	5	20	0	60	20	0	-2.07	0.20	0.95	0.85	c		Yes	
ppvt013	5	20	80	0	0	0	-3.11	0.14	0.95	0.80	c		Yes	
ppvt014	5	80	0	20	0	0	-3.11	0.14	0.97	0.83	c		Yes	
ppvt015	5	40	0	40	20	0	-1.16	0.25	0.81	0.81	c		Yes	
ppvt016	5	20	40	20	20	0	-0.03	0.29	0.48	0.33	c	_BC_	Yes	
ppvt017	5	40	20	40	0	0	-1.16	0.25	0.77	0.75	c		Yes	
ppvt018	5	60	0	40	0	0	1.48	0.30	0.00	0.00	c	_BC_	Yes	
ppvt019	5	40	40	0	20	0	-1.16	0.25	0.79	0.78	c		Yes	
ppvt020	5	20	40	0	40	0	-1.16	0.23	1.60	1.62	c	_BC_	Yes	
ppvt021	5	0	20	80	0	0	-3.11	0.14	0.97	0.83	c		Yes	
ppvt022	5	60	0	0	40	0	-1.16	0.22	1.62	1.65	c	_BC_	Yes	
ppvt023	5	40	40	20	0	0	-1.16	0.25	0.81	0.81	c		Yes	
ppvt024	5	40	0	0	60	0	-2.07	0.17	1.53	2.18	c	_BC_	Yes	
ppvt025	18	72	17	6	6	0	-1.57	0.25	1.24	1.33	c		Yes	
ppvt026	18	6	0	94	0	0	-3.69	0.14	0.90	0.37	c	_C_	Yes	
ppvt027	18	0	6	6	89	0	-2.88	0.20	0.80	0.39	c	_C_	Yes	
ppvt028	18	0	0	0	100	0	-4.98	c	c	c	c		Yes	
ppvt029	18	100	0	0	0	0	-4.98	c	c	c	c		Yes	
ppvt030	18	0	94	6	0	0	-3.69	0.14	0.88	0.35	c	_C_	Yes	
ppvt031	18	0	6	83	11	0	-2.35	0.25	0.66	0.38	c	_C_	Yes	
ppvt032	18	0	94	0	6	0	-3.69	0.14	0.90	0.37	c	_C_	Yes	
ppvt033	18	11	67	11	11	0	-1.24	0.33	0.75	0.64	c		Yes	
ppvt034	18	83	11	0	6	0	-2.35	0.25	0.67	0.38	c	_C_	Yes	
ppvt035	18	11	6	22	61	0	-0.93	0.34	0.70	0.60	c		Yes	
ppvt036	18	0	17	83	0	0	-2.35	0.23	0.79	0.47	c	_C_	Yes	
ppvt037	25	20	52	24	4	0	-0.29	0.35	1.12	1.19	c		Yes	
ppvt038	25	8	8	0	84	0	-2.20	0.29	0.55	0.31	c	_C_	Yes	
ppvt039	25	8	16	64	12	0	-0.90	0.36	0.79	0.70	c		Yes	
ppvt040	25	60	16	12	12	0	-0.69	0.33	1.21	1.25	c		Yes	
ppvt041	25	48	48	0	4	0	-0.09	0.38	1.23	1.26	c		Yes	
ppvt042	25	44	28	16	12	0	0.11	0.36	1.28	1.42	c		Yes	
ppvt043	25	0	0	92	8	0	-3.08	0.18	1.02	0.88	c		Yes	
ppvt044	25	4	16	8	72	0	-1.36	0.33	0.81	0.68	c		Yes	
ppvt045	25	0	16	0	84	0	-2.20	0.21	1.04	2.03	c	_C_	Yes	
ppvt046	25	12	80	4	4	0	-1.89	0.27	1.07	0.94	c		Yes	
ppvt047	25	0	0	100	0	0	-5.13	c	c	c	c		Yes	
ppvt048	25	84	8	0	8	0	-2.20	0.25	0.82	0.99	c		Yes	
ppvt049	33	0	6	88	6	0	-2.41	0.28	0.66	0.33	1.29	_C_	Yes	
ppvt050	33	70	18	9	3	0	-1.07	0.33	1.02	1.29	0.77		Yes	
ppvt051	33	76	9	6	9	0	-1.44	0.31	1.08	1.18	0.82		Yes	
ppvt052	33	15	70	9	6	0	-1.07	0.36	0.86	0.94	0.77		Yes	
ppvt053	33	45	27	9	18	0	1.71	0.49	1.14	1.12	0.98		Yes	
ppvt054	33	0	76	18	6	0	-1.44	0.31	0.98	1.27	0.00		Yes	

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male				Female		
ppvt055	33	73	27	0	0	0	-1.25	0.34	1.10	0.97	-0.32		___	Yes
ppvt056	33	6	30	24	39	0	0.46	0.44	1.13	1.03	-0.35		___	Yes
ppvt057	33	12	76	6	6	0	-1.44	0.34	0.88	0.75	0.82		___	Yes
ppvt058	33	12	0	82	6	0	-1.87	0.31	0.70	0.47	-0.02		__C_	Yes
ppvt059	33	6	3	64	27	0	-0.74	0.35	1.05	1.35	0.75		___	Yes
ppvt060	33	6	9	9	76	0	-1.44	0.37	0.64	0.51	0.82		___	Yes
ppvt061	63	14	0	0	86	0	-1.68	0.33	0.91	0.58	-0.19		___	Yes
ppvt062	63	41	35	14	10	0	0.91	0.47	1.27	1.63	-0.03		__C_	Yes
ppvt063	63	2	97	2	0	0	-3.50	0.20	0.79	0.17	1.70		__C_	Yes
ppvt064	63	5	3	92	0	0	-2.44	0.26	0.97	0.57	0.53		___	Yes
ppvt065	63	10	2	11	78	0	-1.04	0.32	1.21	1.30	-0.26		___	Yes
ppvt066	63	21	14	56	10	0	0.21	0.46	1.03	1.07	0.65		___	Yes
ppvt067	63	14	71	5	10	0	-0.63	0.37	1.24	1.31	-0.94		___	Yes
ppvt068	63	94	0	5	2	0	-2.71	0.25	0.82	0.42	1.28		__C_	Yes
ppvt069	63	3	6	2	89	0	-2.02	0.29	0.95	0.66	0.45		___	Yes
ppvt070	63	11	83	6	0	0	-1.40	0.36	0.77	0.72	0.95		___	Yes
ppvt071	63	0	2	98	0	0	-4.24	0.01	1.08	5.47	0.97		A_C_	No
ppvt072	63	90	5	2	3	0	-2.22	0.29	0.86	0.48	-1.86		__C_	Yes
ppvt073	134	44	32	4	20	0	2.01	0.53	1.21	1.43	0.37		___	Yes
ppvt074	134	1	0	2	96	0	-2.76	0.24	0.88	0.70	0.39		___	Yes
ppvt075	134	5	1	93	1	0	-2.07	0.32	0.86	0.56	-0.31		___	Yes
ppvt076	134	70	10	12	8	0	0.05	0.49	0.92	0.91	-0.15		___	Yes
ppvt077	134	93	2	3	2	0	-1.94	0.28	0.85	0.94	-0.02		___	Yes
ppvt078	134	4	6	85	4	0	-1.02	0.43	0.82	0.59	-0.86		___	Yes
ppvt079	134	1	93	5	0	0	-2.07	0.27	1.02	0.69	-1.52		___	Yes
ppvt080	134	81	5	12	2	0	-0.64	0.39	1.05	1.01	-0.44		___	Yes
ppvt081	134	4	85	3	8	0	-1.02	0.42	0.85	0.65	-0.56		___	Yes
ppvt082	134	5	1	5	88	0	-1.33	0.34	0.97	1.12	-0.69		___	Yes
ppvt083	134	1	14	82	2	0	-0.76	0.46	0.83	0.64	0.27		___	Yes
ppvt084	134	14	4	3	79	0	-0.53	0.43	0.94	0.88	0.69	*	__D	Yes
ppvt085	144	40	25	17	18	0	2.50	0.55	1.22	2.19	1.09		__C_	Yes
ppvt086	144	24	13	37	25	1	2.50	0.58	1.13	1.27	0.03		___	Yes
ppvt087	144	54	12	21	13	1	0.97	0.50	1.19	1.37	0.45		___	Yes
ppvt088	144	6	11	6	76	1	-0.28	0.50	0.89	0.82	0.15		___	Yes
ppvt089	144	23	13	51	14	0	1.14	0.51	1.13	1.17	0.22		___	Yes
ppvt090	144	88	4	3	4	0	-1.29	0.33	0.97	1.05	-0.52		___	Yes
ppvt091	144	17	20	36	26	1	2.45	0.53	1.23	2.81	0.00		__C_	Yes
ppvt092	144	10	40	47	3	0	1.66	0.52	1.17	1.67	0.35		__C_	Yes
ppvt093	144	38	5	56	1	0	0.86	0.44	1.30	1.47	0.53		___	Yes
ppvt094	144	13	65	15	7	1	0.43	0.47	1.11	1.18	0.62		___	Yes
ppvt095	144	2	1	81	16	0	-0.58	0.41	1.06	0.97	-0.25		___	Yes
ppvt096	144	82	9	2	7	0	-0.69	0.42	0.96	0.82	-0.26		___	Yes
ppvt097	491	2	88	8	2	0	-0.32	0.50	0.90	0.84	0.21		___	Yes
ppvt098	491	15	81	2	2	0	0.33	0.48	1.00	0.91	0.39		___	Yes
ppvt099	491	1	2	8	88	0	-0.32	0.41	1.08	0.94	-0.07		___	Yes
ppvt100	491	3	3	90	3	0	-0.57	0.51	0.91	0.78	0.14		___	Yes
ppvt101	491	1	3	1	95	0	-1.35	0.46	0.89	0.71	0.45		___	Yes
ppvt102	491	67	7	3	23	0	1.23	0.42	1.09	1.10	-0.22		___	Yes
ppvt103	491	2	1	96	1	0	-1.56	0.44	0.91	0.72	0.08		___	Yes
ppvt104	491	1	2	0	97	0	-1.95	0.45	0.86	0.46	-0.20		__C_	Yes
ppvt105	491	91	7	2	1	0	-0.62	0.45	1.01	0.88	-0.43		___	Yes
ppvt106	491	4	76	16	3	0	0.69	0.39	1.13	1.26	-0.20		___	Yes
ppvt107	491	91	3	6	0	0	-0.65	0.39	1.09	1.17	0.00		___	Yes
ppvt108	491	2	10	88	0	0	-0.23	0.46	0.98	1.00	0.28		___	Yes
ppvt109	916	7	9	14	69	0	1.57	0.42	0.99	0.98	-0.15		___	Yes
ppvt110	916	16	79	3	1	0	0.92	0.29	1.11	1.09	0.00		___	Yes
ppvt111	916	5	18	73	3	0	1.31	0.45	0.96	0.98	0.07	*	__D	Yes
ppvt112	916	2	6	88	4	0	0.18	0.43	0.97	0.89	-0.46		___	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT
INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
								Male-Female						
ppvt113	916	29	22	26	23	0	3.56	0.14	1.18	1.43	0.14		___	Yes
ppvt114	916	93	3	2	2	0	-0.53	0.43	0.95	0.86	-0.42	*	___D	Yes
ppvt115	916	5	5	11	79	0	0.92	0.41	0.99	1.02	-0.52		___	Yes
ppvt116	916	5	91	2	3	0	-0.14	0.47	0.91	0.77	-0.35		___	Yes
ppvt117	916	84	7	7	2	0	0.58	0.39	1.02	0.97	0.07		___	Yes
ppvt118	916	1	99	0	0	0	-2.55	0.37	1.03	0.67	0.54		___	Yes
ppvt119	916	1	1	1	98	0	-2.07	0.50	0.84	0.22	-1.01		__C_	Yes
ppvt120	916	6	7	81	5	1	0.77	0.45	0.94	0.97	0.18		___	Yes
ppvt121	896	8	88	4	0	0	0.32	0.49	0.99	0.97	0.13		___	Yes
ppvt122	896	3	3	13	82	0	0.88	0.42	1.03	1.09	-0.07		___	Yes
ppvt123	896	4	8	84	4	0	0.71	0.51	0.93	0.87	-0.09		___	Yes
ppvt124	896	31	19	28	21	0	4.23	0.12	1.18	1.56	-0.09		__C_	Yes
ppvt125	896	7	15	72	6	0	1.47	0.46	0.96	0.91	0.26		___	Yes
ppvt126	896	78	12	4	6	0	1.10	0.50	0.93	0.88	-0.09	*	___D	Yes
ppvt127	896	0	97	2	0	0	-1.41	0.52	1.02	1.18	-0.92		___	Yes
ppvt128	896	79	3	7	10	0	1.06	0.48	0.96	0.97	-0.24		___	Yes
ppvt129	896	1	1	3	95	0	-0.64	0.54	0.97	0.80	-0.58		___	Yes
ppvt130	896	79	5	11	5	0	1.07	0.48	0.97	0.92	0.29		___	Yes
ppvt131	896	12	4	5	78	0	1.10	0.41	1.03	1.10	0.08		___	Yes
ppvt132	896	1	1	97	1	0	-1.28	0.60	0.94	0.55	0.28		___	Yes
ppvt133	883	1	3	95	2	0	-0.61	0.62	0.97	0.77	0.20		___	Yes
ppvt134	883	1	96	1	1	0	-0.98	0.62	0.97	0.89	-0.31		___	Yes
ppvt135	883	79	3	10	7	0	1.08	0.54	0.94	0.84	-0.15		___	Yes
ppvt136	883	5	37	28	30	0	3.62	0.14	1.25	1.36	0.11		___	Yes
ppvt137	883	47	30	11	12	0	2.74	0.29	1.16	1.20	0.20		___	Yes
ppvt138	883	56	11	9	24	0	2.35	0.39	0.99	0.96	0.00		___	Yes
ppvt139	883	20	6	5	69	0	1.71	0.43	1.02	0.97	-0.18		___	Yes
ppvt140	883	2	20	17	60	0	2.14	0.39	1.03	1.00	0.00		___	Yes
ppvt141	883	3	83	7	6	0	0.83	0.54	0.94	0.85	-0.07		___	Yes
ppvt142	883	42	5	50	3	0	2.61	0.31	1.08	1.09	0.24		___	Yes
ppvt143	883	3	92	4	1	0	-0.08	0.57	0.98	0.75	-0.11		___	Yes
ppvt144	883	3	3	88	6	1	0.39	0.56	0.96	0.97	-0.12		___	Yes
ppvt145	852	6	13	10	70	0	1.71	0.45	1.06	1.06	0.02		___	Yes
ppvt146	852	9	19	64	7	0	2.01	0.52	0.93	0.94	-0.20		___	Yes
ppvt147	852	2	95	2	1	0	-0.55	0.67	0.98	0.88	-0.13		___	Yes
ppvt148	852	6	8	79	8	0	1.22	0.54	0.97	0.93	-0.34		___	Yes
ppvt149	852	65	11	6	18	0	1.99	0.45	1.05	1.04	0.00	*	___D	Yes
ppvt150	852	10	4	80	6	0	1.09	0.53	0.98	0.99	-0.39		___	Yes
ppvt151	852	1	4	2	93	0	-0.14	0.65	0.98	0.86	-0.07		___	Yes
ppvt152	852	73	9	3	15	0	1.57	0.51	1.00	1.06	0.19		___	Yes
ppvt153	852	4	83	10	3	0	0.87	0.54	1.01	1.04	0.18		___	Yes
ppvt154	852	3	75	9	13	0	1.42	0.53	0.97	0.93	0.16		___	Yes
ppvt155	852	86	4	4	6	0	0.66	0.61	0.94	1.17	0.09	*	___D	Yes
ppvt156	852	21	4	6	69	1	1.76	0.47	1.02	1.06	-0.39		___	Yes
ppvt157	820	8	73	16	3	0	1.64	0.61	0.92	0.88	0.00	*	___D	Yes
ppvt158	820	49	24	5	21	0	2.78	0.50	0.88	0.85	0.56	*	___D	Yes
ppvt159	820	5	27	24	44	0	3.00	0.44	0.92	0.92	0.33		___	Yes
ppvt160	820	19	31	36	15	0	3.38	0.34	1.05	1.07	0.00		___	Yes
ppvt161	820	4	4	1	91	0	0.15	0.63	1.03	1.12	0.26		___	Yes
ppvt162	820	50	10	33	7	0	2.71	0.44	1.00	0.98	-0.07	*	___D	Yes
ppvt163	820	11	46	23	20	0	2.91	0.46	0.93	0.92	-0.33		___	Yes
ppvt164	820	39	22	12	27	0	3.25	0.33	1.11	1.15	0.10	*	___D	Yes
ppvt165	820	11	3	7	78	0	1.30	0.58	0.99	0.96	0.36		___	Yes
ppvt166	820	6	90	3	1	0	0.35	0.70	0.94	0.72	0.13		___	Yes
ppvt167	820	6	3	87	4	0	0.60	0.66	0.97	0.85	0.39		___	Yes
ppvt168	820	7	8	62	20	3	2.18	0.56	0.91	0.88	-0.29	*	___D	Yes
ppvt169	711	3	5	1	92	0	0.20	0.78	0.94	0.75	-0.57		___	Yes
ppvt170	711	9	70	18	3	0	1.93	0.61	0.95	0.92	-0.31	*	___D	Yes

PSYCHOMETRIC CHARACTERISTICS OF COGNITIVE DEVELOPMENT AND ACHIEVEMENT INSTRUMENTS IN ROUND 3 OF YOUNG LIVES

	N	Response options					Difficulty ² (IRT)	CTT item-fit indicator	IRT item-fit indicator		Bias analysis ³		Warnings	Item kept for the analysis ⁴
		1 (%)	2 (%)	3 (%)	4 (%)	NR ¹ (%)			Item-test correlation	In-fit (IRT)	Out-fit (IRT)	Gender		
										Male-Female				
ppvt171	711	81	9	5	5	0	1.26	0.67	0.97	0.92	0.42		___	Yes
ppvt172	711	13	11	48	28	0	2.94	0.46	0.99	1.00	-0.24	*	___D	Yes
ppvt173	711	33	18	6	44	0	3.16	0.47	0.87	0.86	0.49		___	Yes
ppvt174	711	1	67	32	1	0	2.11	0.51	1.14	1.32	-0.08	*	___D	Yes
ppvt175	711	6	6	83	5	0	1.11	0.68	0.99	0.97	-0.43	*	___D	Yes
ppvt176	711	2	83	4	11	0	1.13	0.69	0.96	0.92	-0.55		___	Yes
ppvt177	711	59	1	3	38	0	2.48	0.51	1.01	0.98	0.21		___	Yes
ppvt178	711	1	11	80	8	0	1.38	0.67	0.96	0.89	-0.09		___	Yes
ppvt179	711	2	4	2	91	0	0.32	0.78	0.95	0.81	-0.29		___	Yes
ppvt180	711	54	2	41	2	2	2.69	0.51	0.97	0.95	-0.08		___	Yes
ppvt181	684	5	20	8	68	0	2.08	0.59	0.98	0.95	-0.14	*	___D	Yes
ppvt182	684	82	12	5	1	0	1.25	0.69	0.98	0.90	1.04		___	Yes
ppvt183	684	0	2	97	1	0	-0.63	0.82	0.98	0.78	0.38		___	Yes
ppvt184	684	3	4	3	90	0	0.58	0.74	1.01	1.04	0.11		___	Yes
ppvt185	684	72	13	8	7	0	1.87	0.62	0.96	0.94	0.15		___	Yes
ppvt186	684	2	3	93	2	0	0.16	0.79	0.96	0.78	0.10		___	Yes
ppvt187	684	1	98	1	0	0	-1.30	0.83	0.99	0.78	-0.43		___	Yes
ppvt188	684	13	18	9	61	0	2.43	0.51	1.07	1.13	-0.22		___	Yes
ppvt189	684	3	90	4	3	0	0.60	0.77	0.92	0.75	-0.25		___	Yes
ppvt190	684	2	2	94	2	0	-0.04	0.79	0.98	0.88	0.13		___	Yes
ppvt191	684	86	5	1	8	0	0.92	0.73	0.95	0.82	0.25		___	Yes
ppvt192	684	35	59	3	2	1	2.53	0.50	1.06	1.06	-0.18		___	Yes
ppvt193	669	44	13	14	29	0	3.94	0.30	1.12	1.19	-0.05	*	___D	Yes
ppvt194	669	3	86	3	7	0	0.96	0.72	0.97	0.85	-0.56		___	Yes
ppvt195	669	5	1	91	2	0	0.42	0.76	0.97	0.82	0.16		___	Yes
ppvt196	669	92	5	0	2	0	0.30	0.76	1.01	1.03	0.10		___	Yes
ppvt197	669	72	5	17	6	0	1.91	0.63	0.94	0.89	0.20		___	Yes
ppvt198	669	29	9	10	51	0	2.90	0.48	0.98	0.97	-0.07		___	Yes
ppvt199	669	2	4	89	6	0	0.73	0.73	0.99	0.91	0.42		___	Yes
ppvt200	669	89	6	2	3	0	0.68	0.73	0.99	1.03	-0.02		___	Yes
ppvt201	669	2	35	8	54	0	2.74	0.44	1.16	1.20	0.02		___	Yes
ppvt202	669	19	65	6	11	0	2.27	0.55	1.01	0.99	-0.18		___	Yes
ppvt203	669	1	1	97	1	0	-0.68	0.81	0.98	0.74	-0.17		___	Yes
ppvt204	669	4	93	1	1	0	0.16	0.77	0.98	0.93	-0.46		___	Yes

Note: Percentages in bold indicate the correct answer for the item.

1. NR: No Response.

2. Item difficulty according to the Rasch (IRT 1 parameter) model estimation.

3. Difference in difficulty between two groups that are compared, adjusting for total ability (the DIF analysis used the method of Mantel-Haenszel).

4. The criteria for keeping an item are: Item does not have a warning by correlation item-test or in-fit statistic.

Warnings: A: Item-test correlation lower than 0.10. B: In-fit out of the range 0.5 to 1.5. C: Out-fit out of the range 0.5 to 1.5. D: The difference by gender is significant at 5%.

Appendix 4. Raw and Rasch scores distributions

Figure 1. Distribution of the Raw scores for the EGRA reading comprehension test – Ethiopia

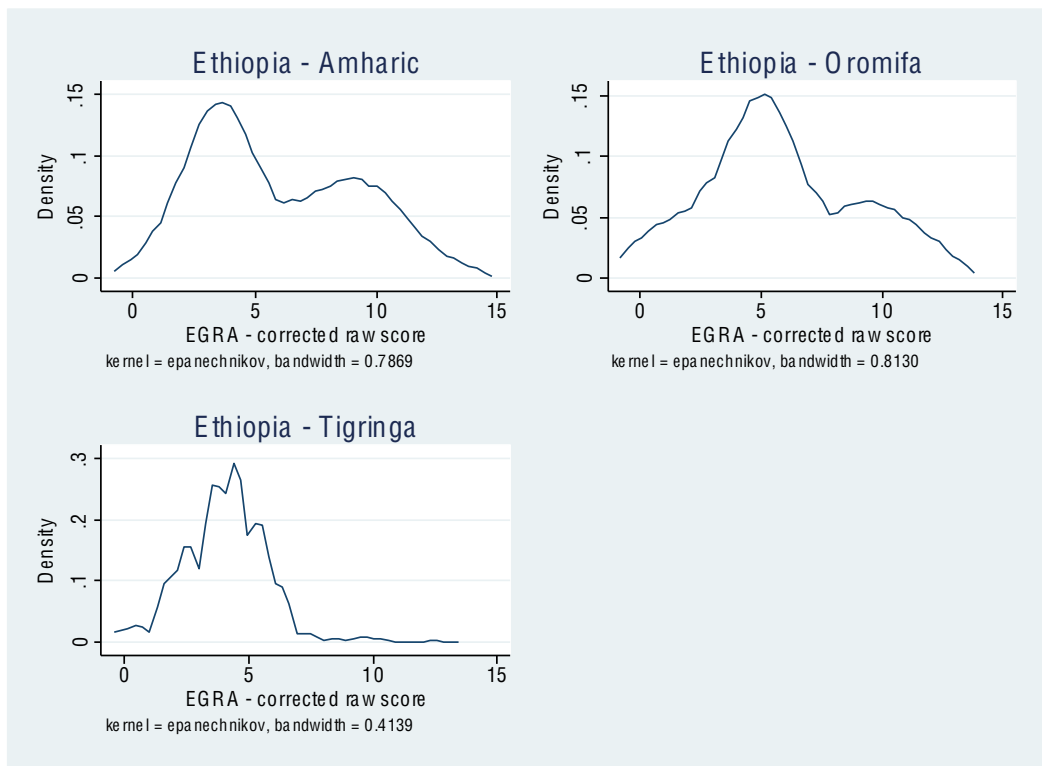


Figure 2. Distribution of the Rasch scores for the EGRA reading comprehension test – Ethiopia

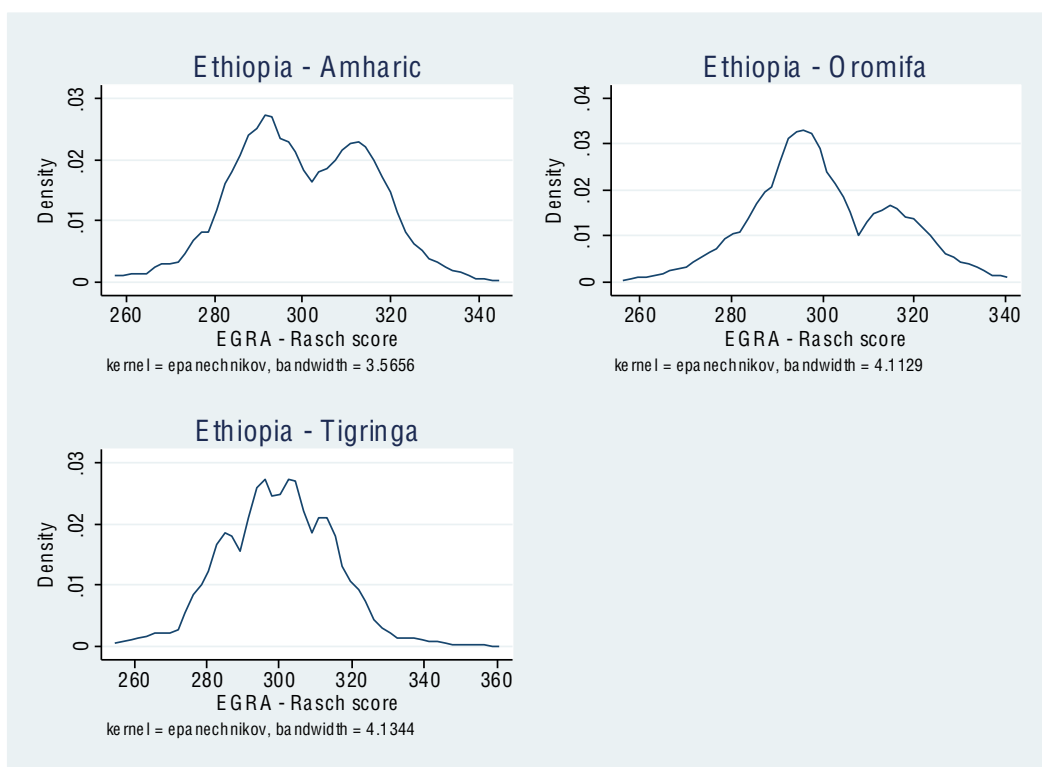


Figure 3. Distribution of the Raw scores for the EGRA reading comprehension test – India

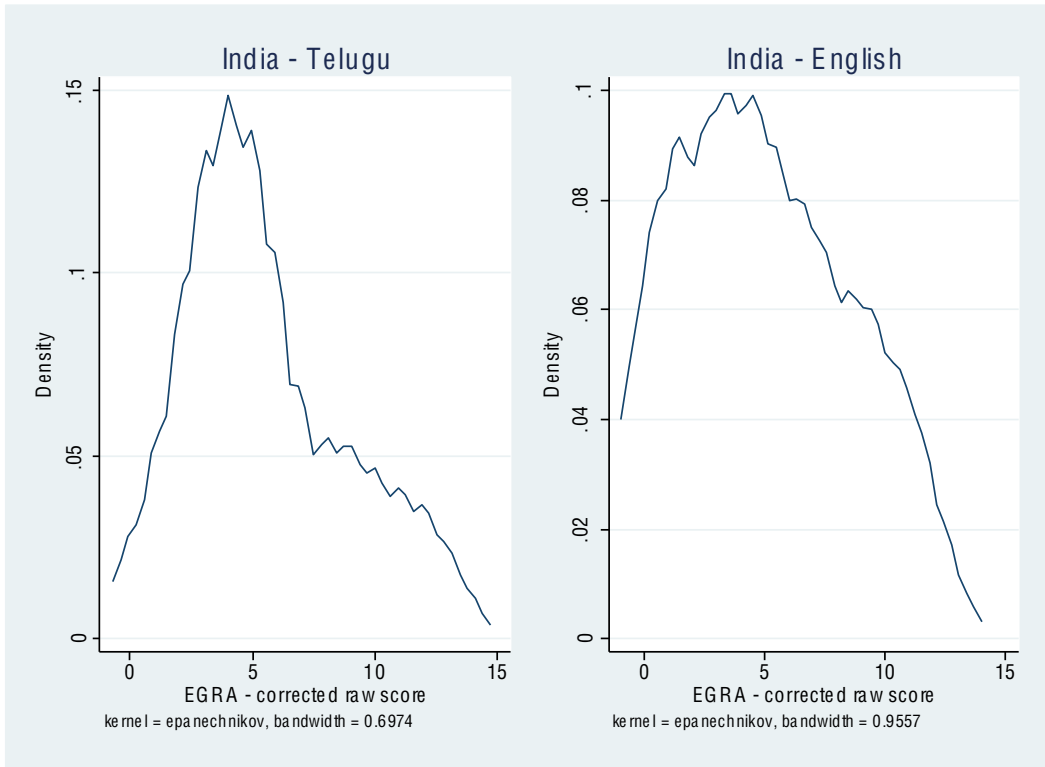


Figure 4. Distribution of the Rasch scores for the EGRA reading comprehension test – India

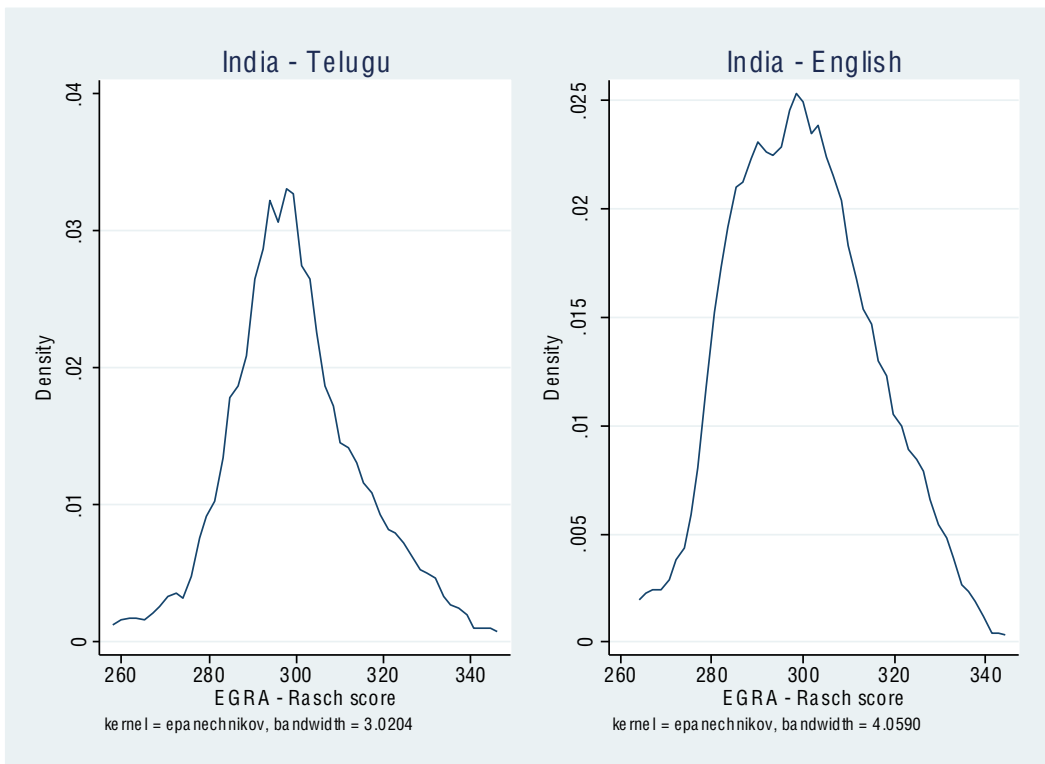


Figure 5. *Distribution of the Raw scores for the EGRA reading comprehension test – Peru*

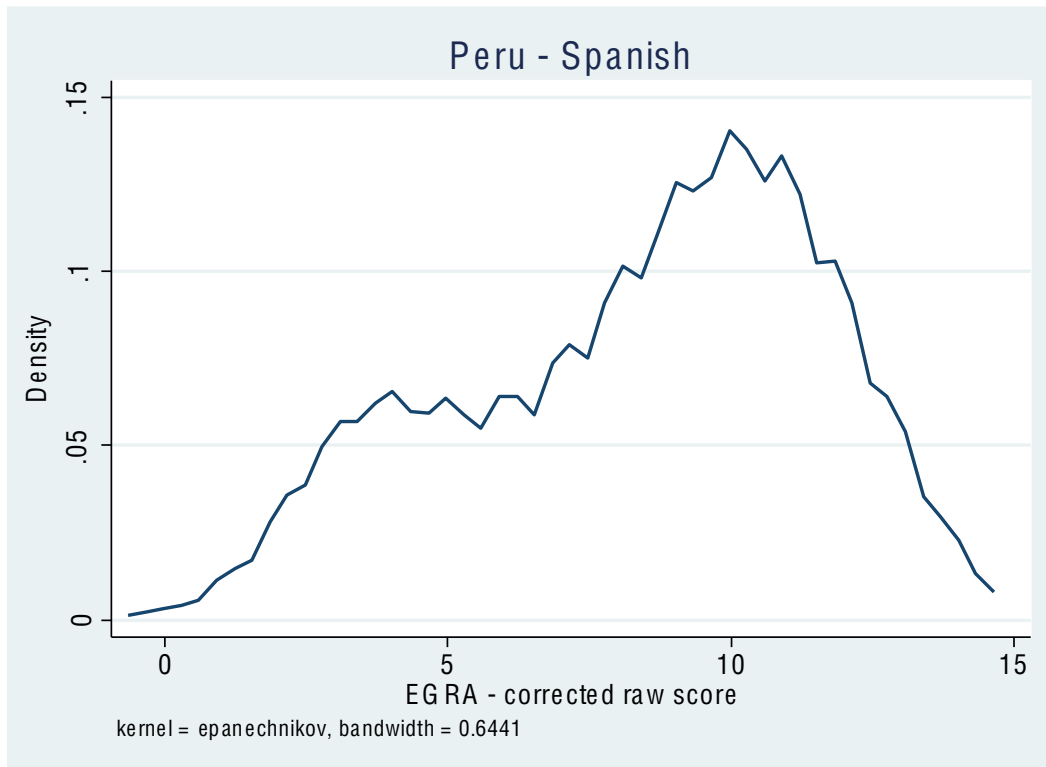


Figure 6. *Distribution of the Rasch scores for the EGRA reading comprehension test – Peru*

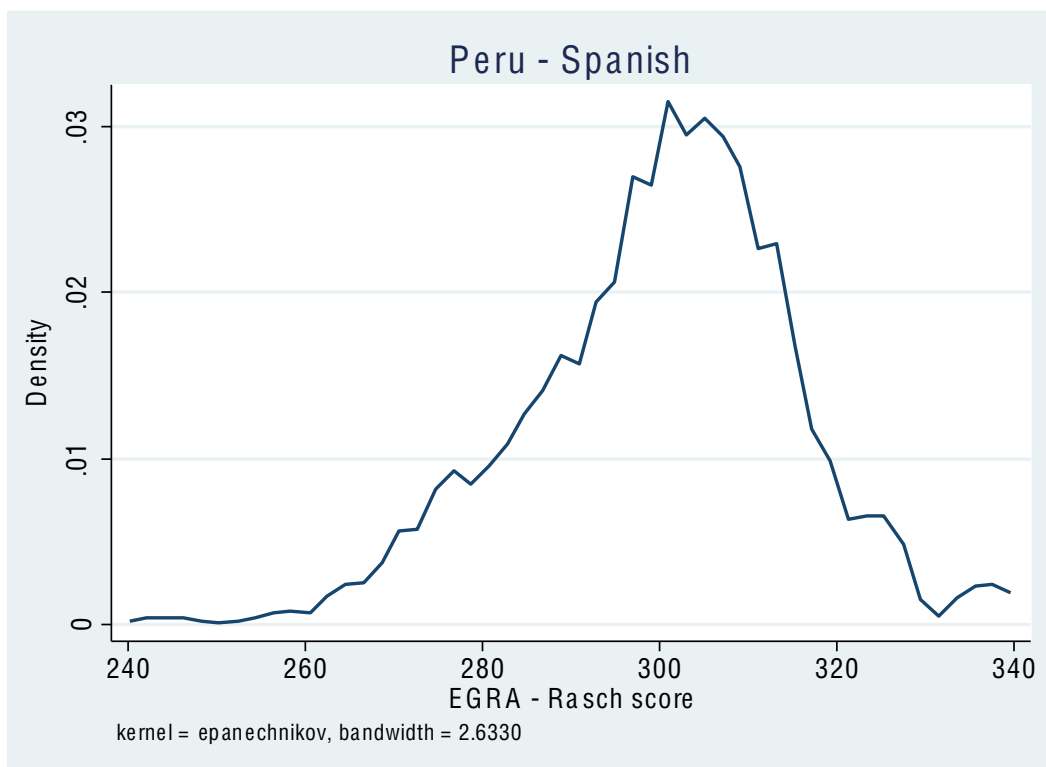


Figure 7. Distribution of the Raw scores for the EGRA reading comprehension test – Vietnam

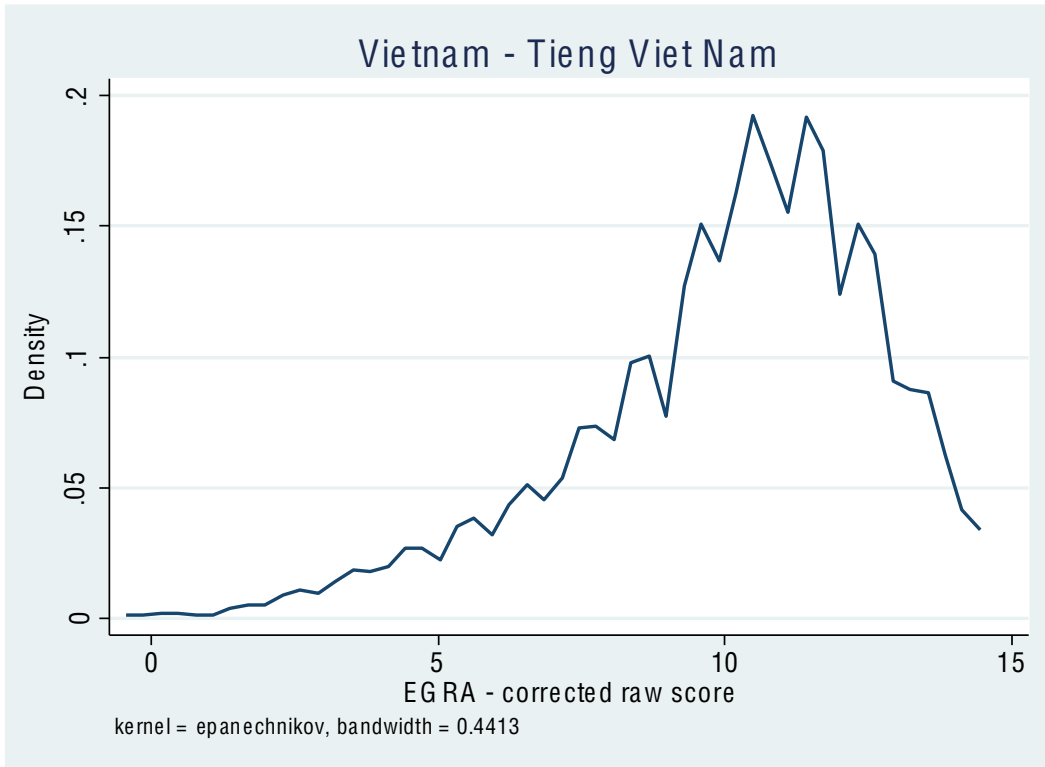


Figure 8. Distribution of the Rasch scores for the EGRA reading comprehension test – Vietnam

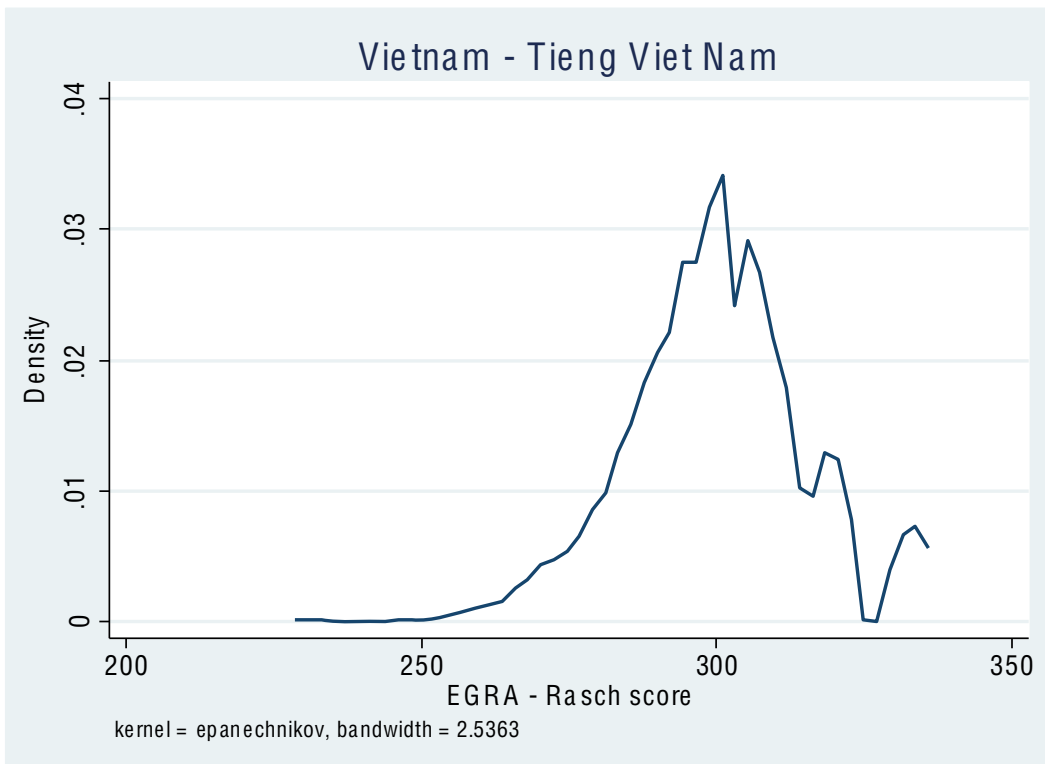


Figure 9. Distribution of the Raw scores for the Maths achievement test of the Younger Cohort – Ethiopia

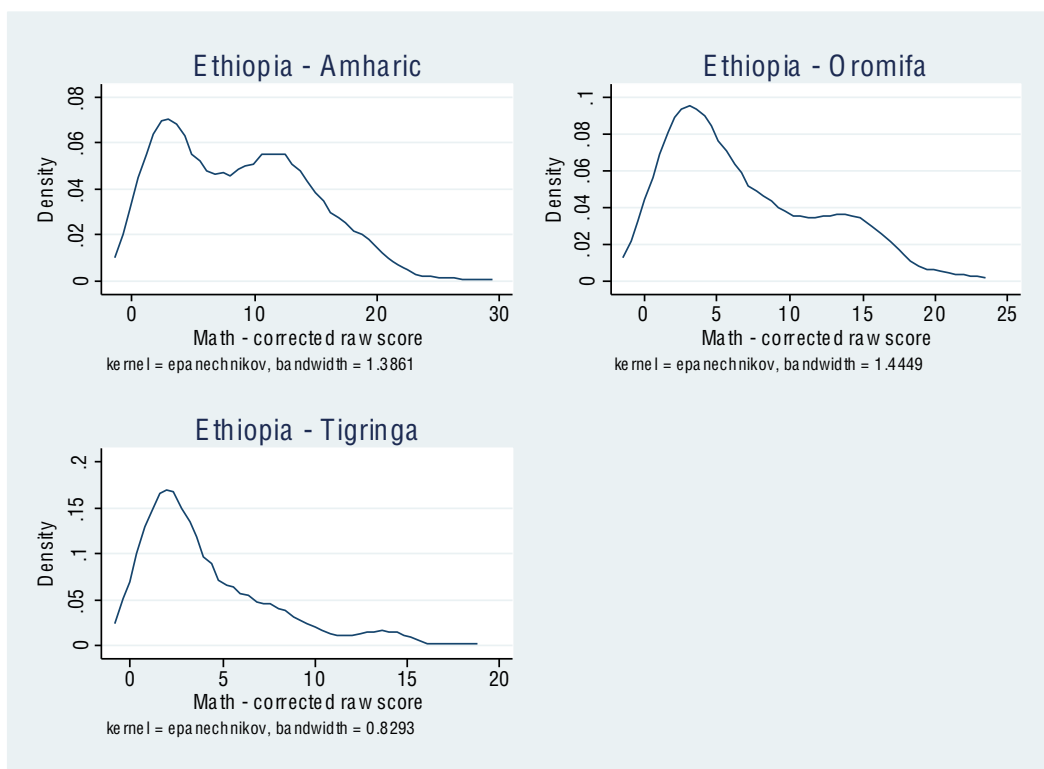


Figure 10. Distribution of the Rasch scores for the Maths achievement test of the Younger Cohort – Ethiopia

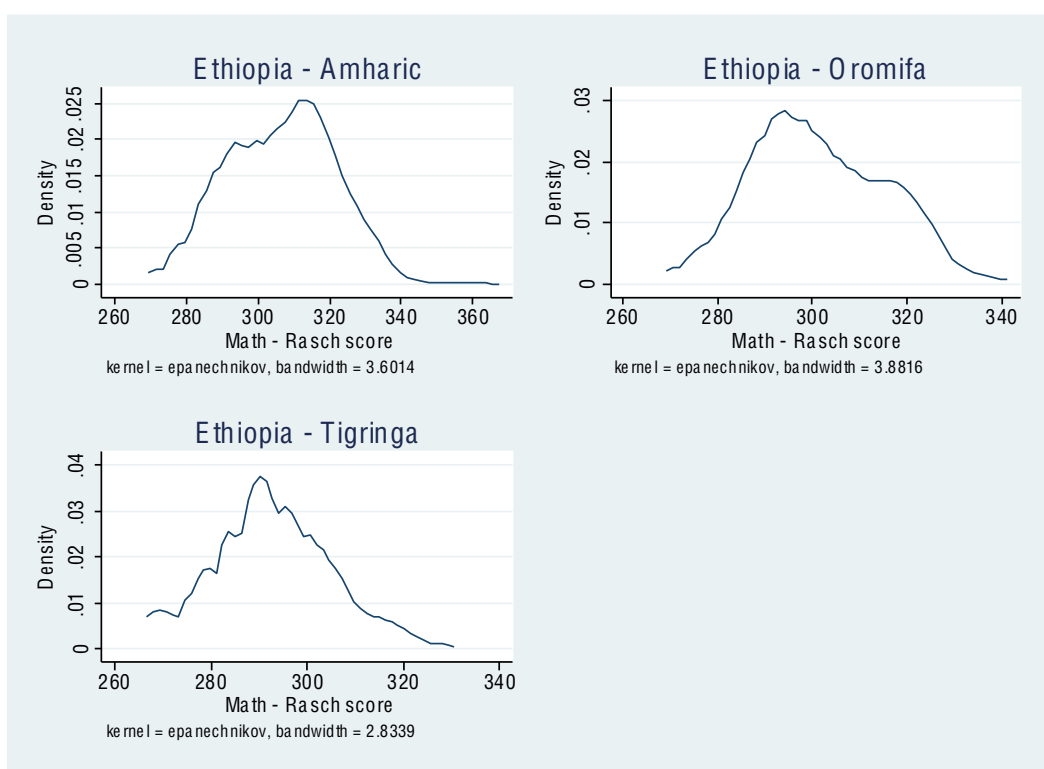


Figure 11. Distribution of the Raw scores for the Maths achievement test of the Younger Cohort – India

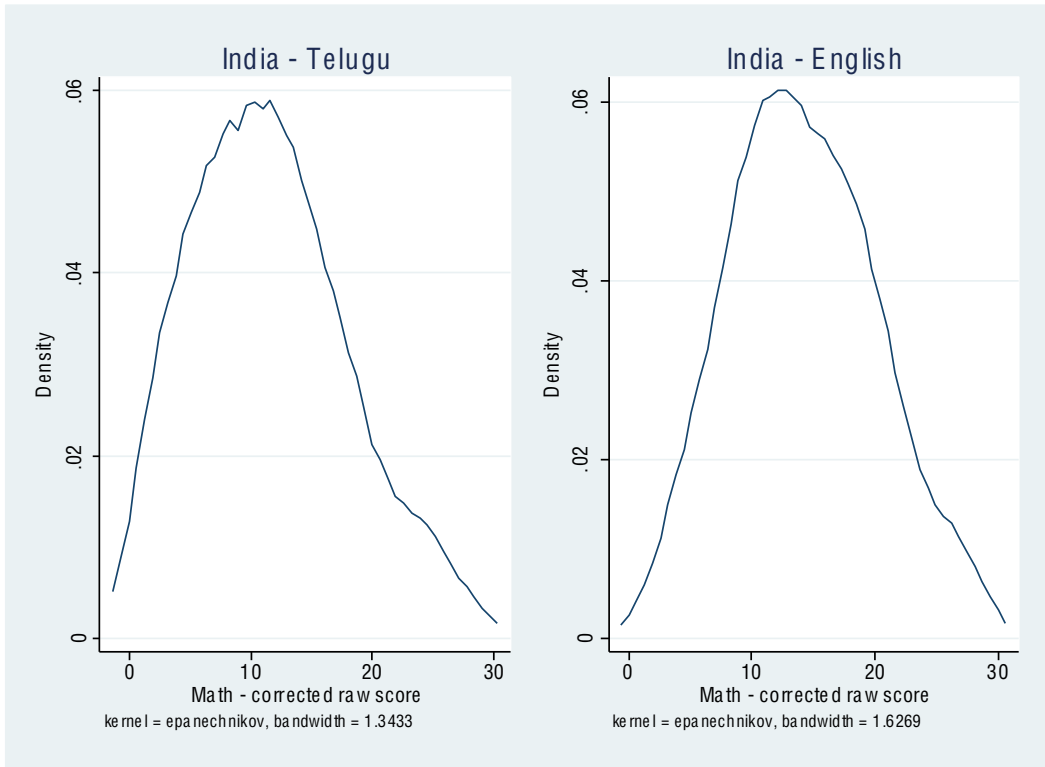


Figure 12. Distribution graph of the Rasch scores for the Maths achievement test of the Younger Cohort – India

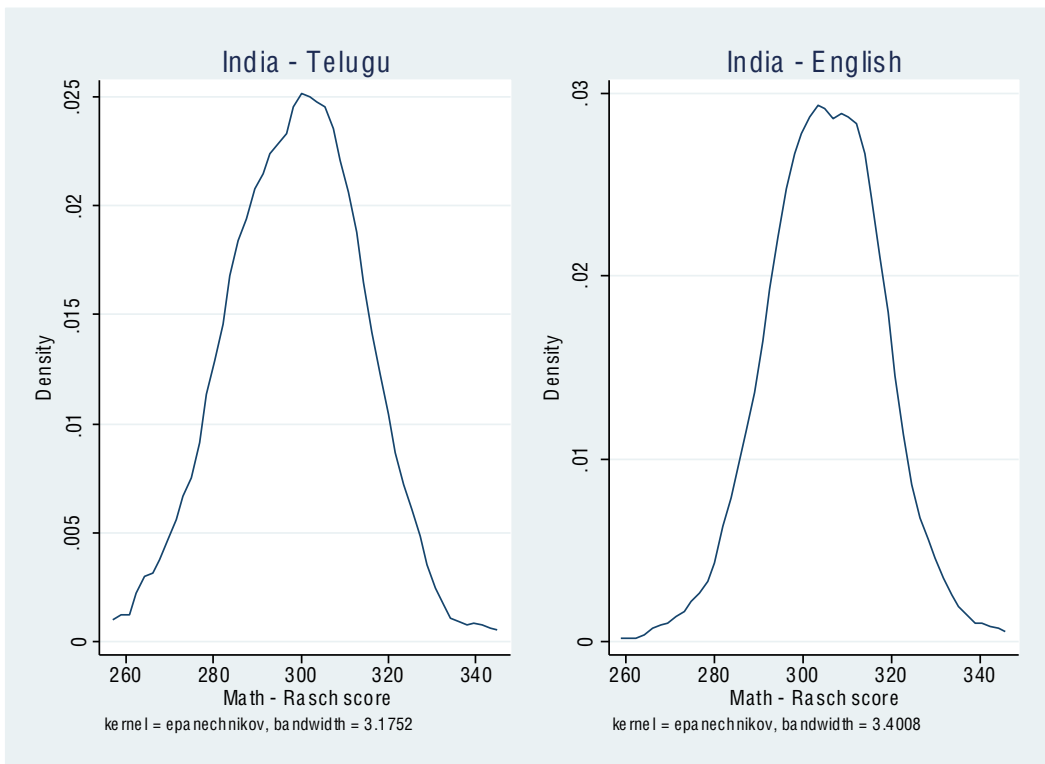


Figure 13. Distribution graph of the Raw scores for the Maths achievement test of the Younger Cohort – Peru

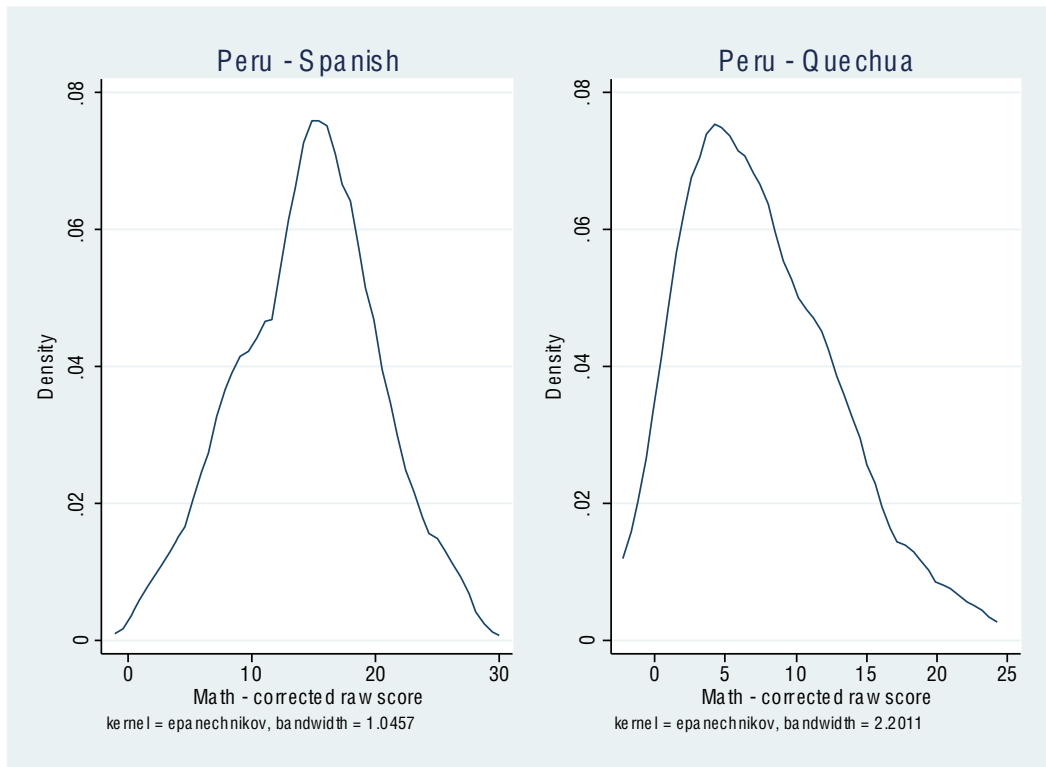


Figure 14. Distribution graph of the Rasch scores for the Maths achievement test of the Younger Cohort – Peru

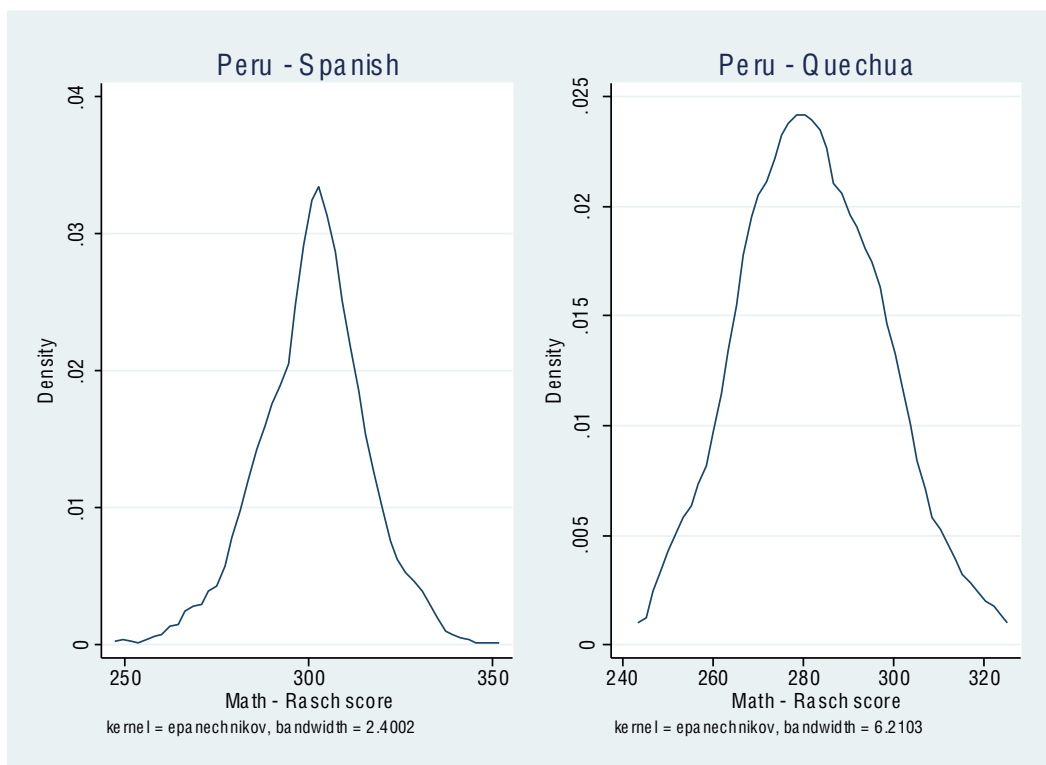


Figure 15. Distribution graph of the Raw scores for the Maths achievement test of the Younger Cohort – Vietnam

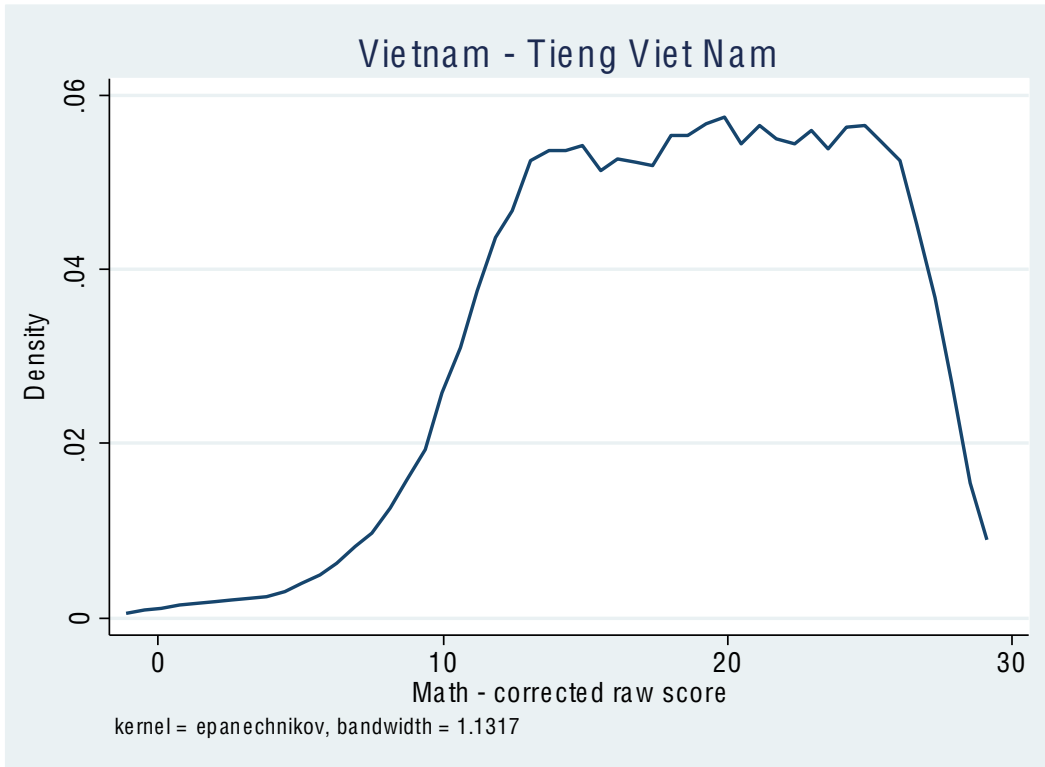


Figure 16. Distribution graph of the Rasch scores for the Maths achievement test of the Younger Cohort – Vietnam

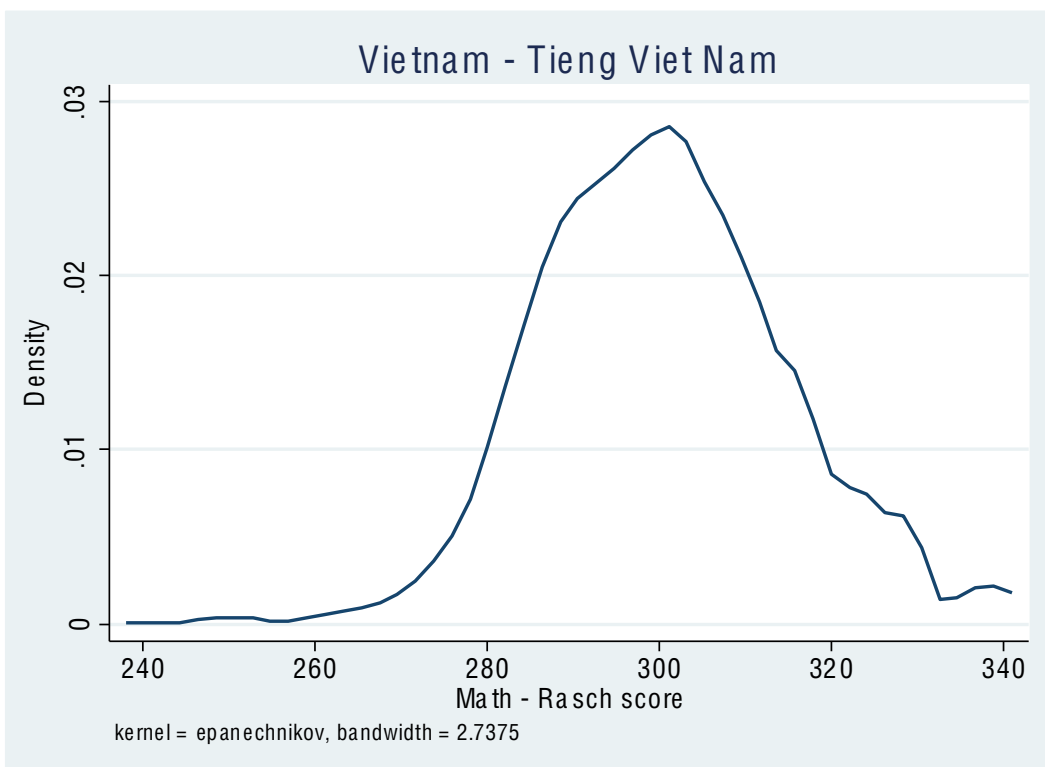


Figure 17. Distribution graph of the Raw scores for the PPVT test of the Younger Cohort – Ethiopia

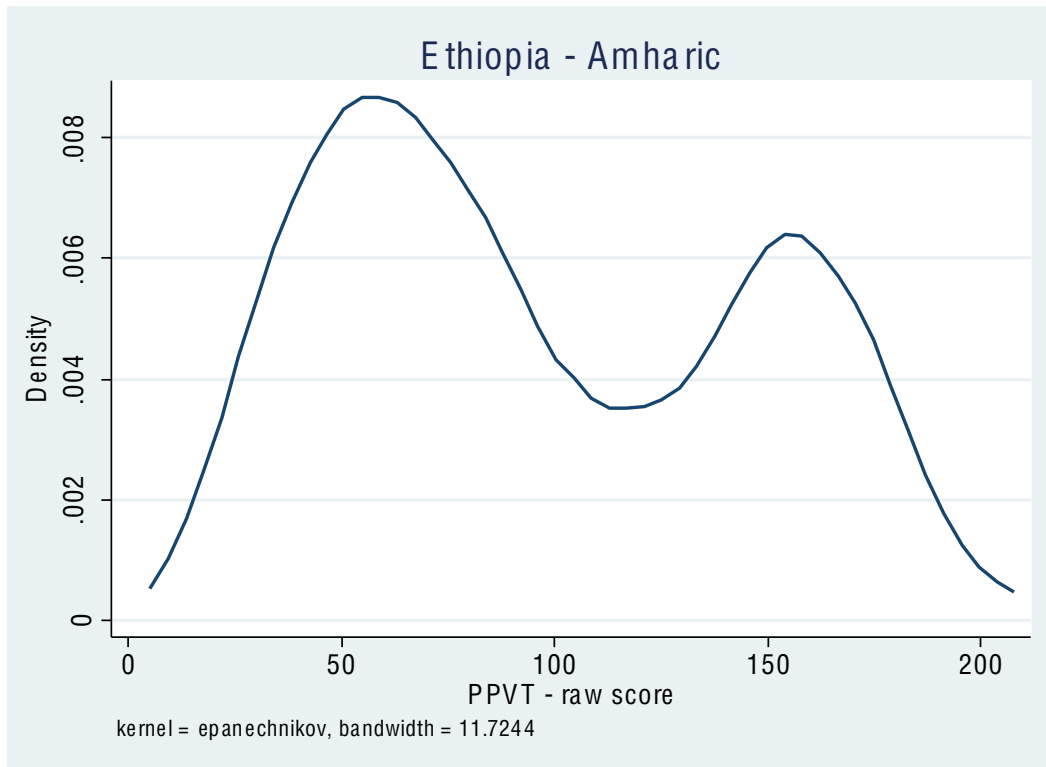


Figure 18. Distribution graph of the Rasch scores for the PPVT test of the Younger Cohort – Ethiopia

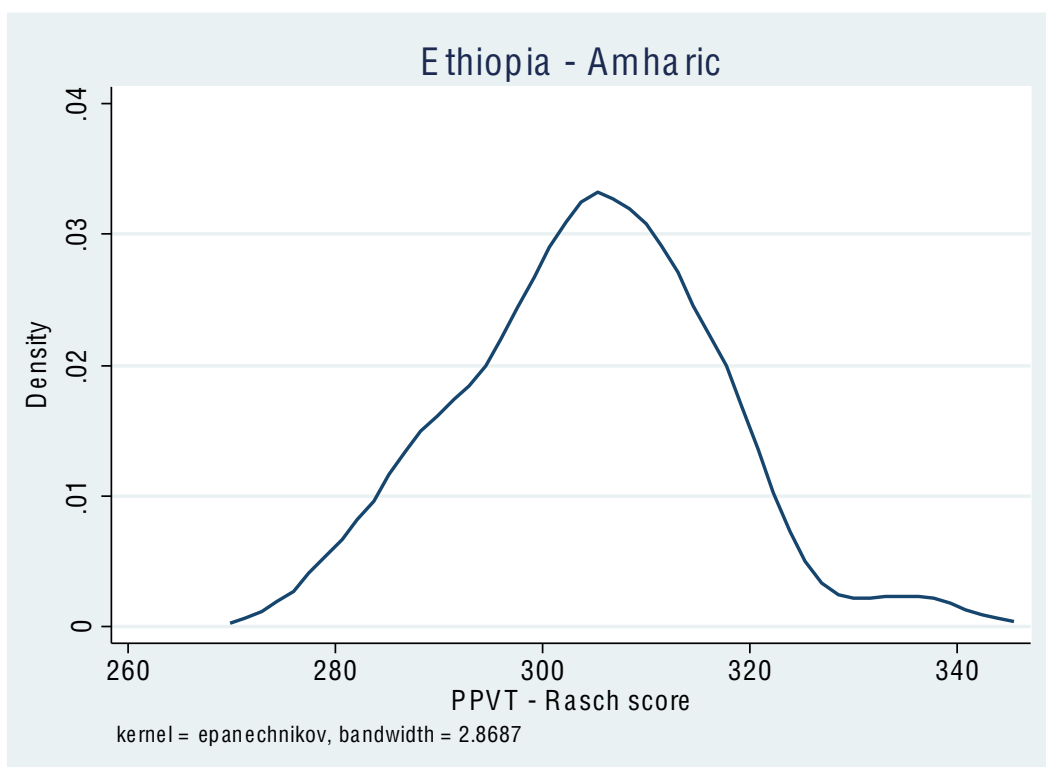


Figure 19. Distribution graph of the raw scores for the PPVT test of the Younger Cohort – India

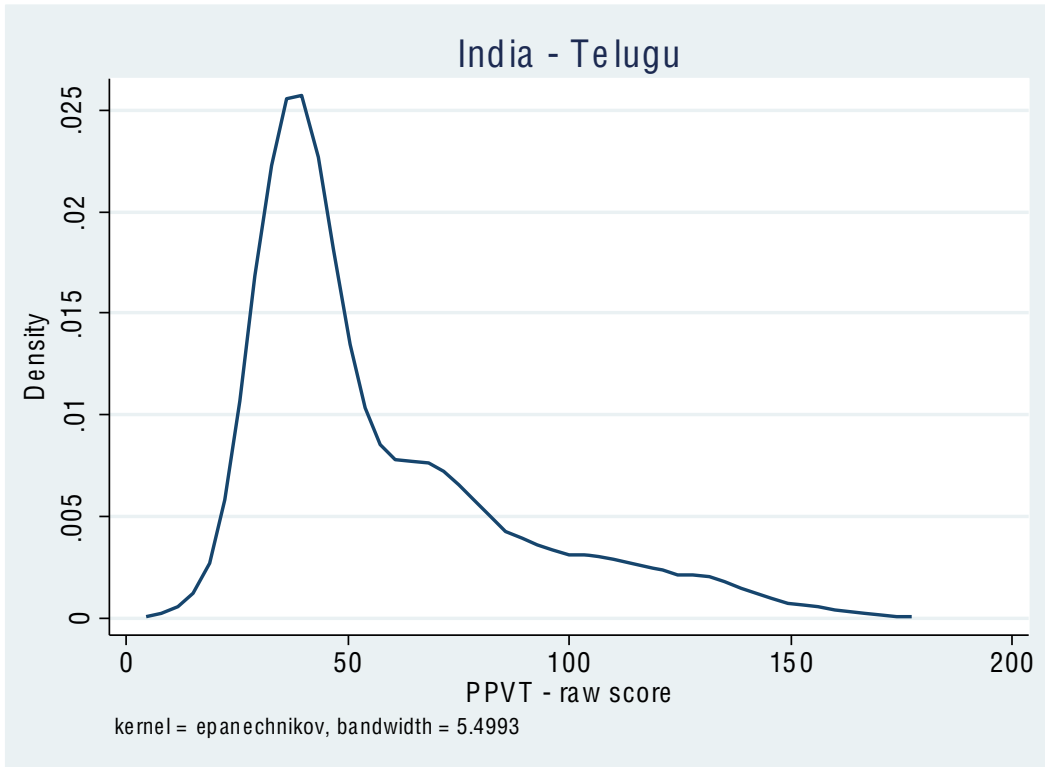


Figure 20. Distribution graph of the Rasch scores for the PPVT test of the Younger Cohort – India

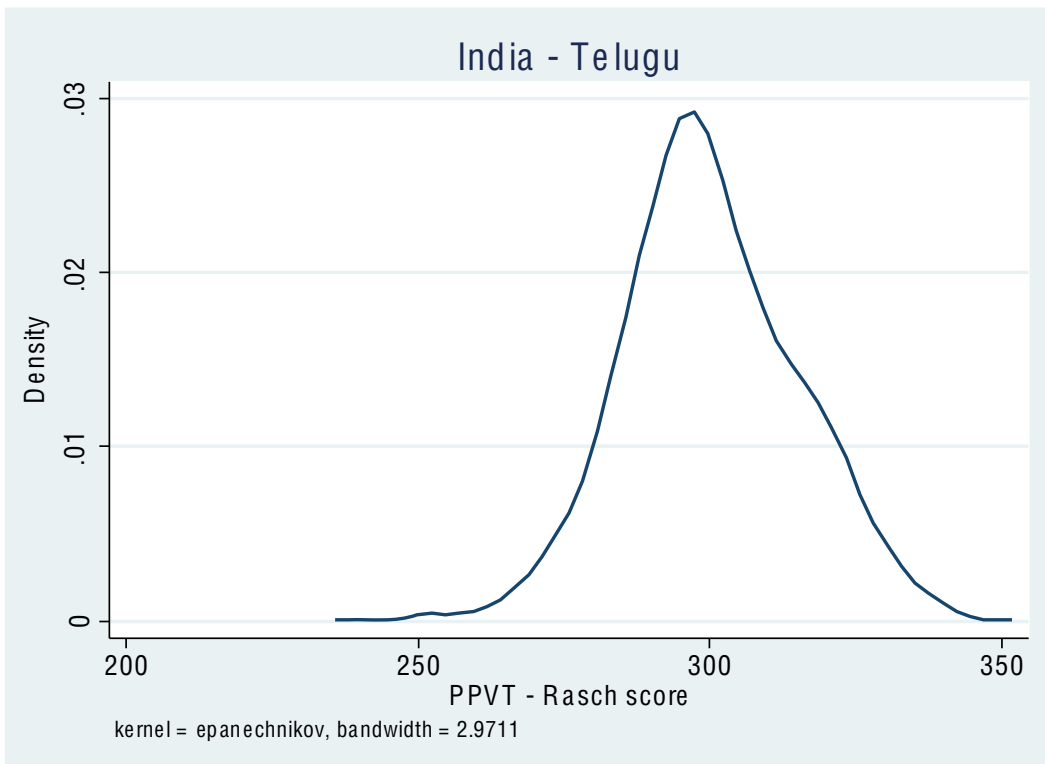


Figure 21. Distribution graph of the raw scores for the PPVT test of the Younger Cohort – Peru

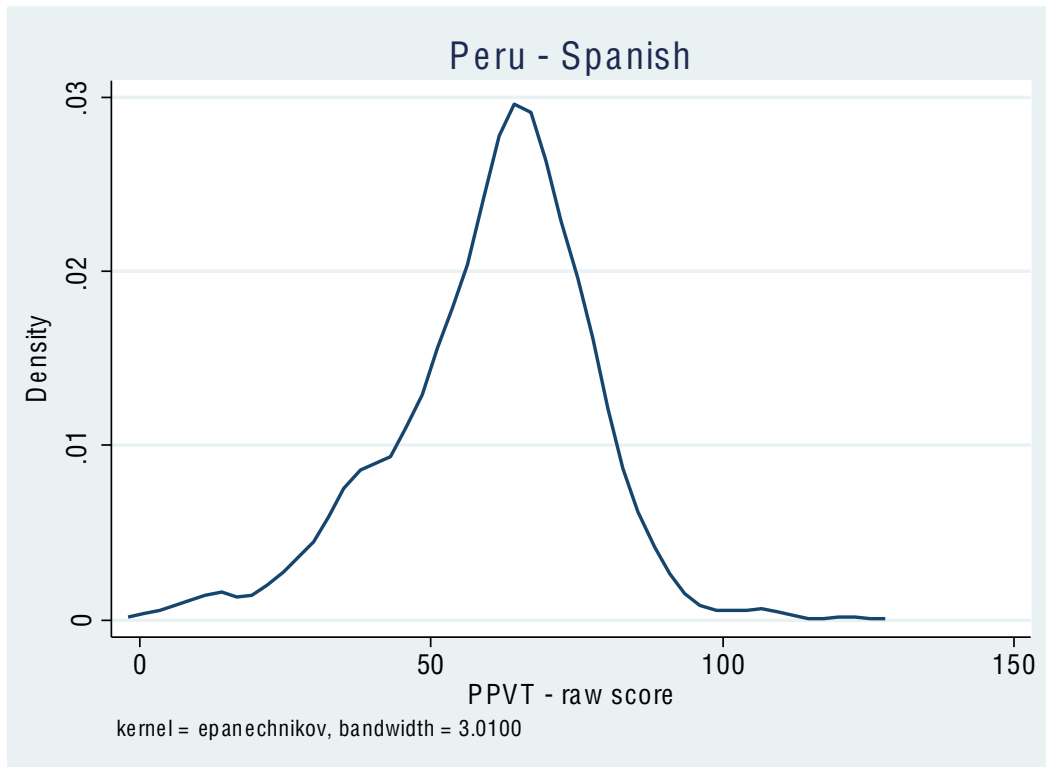


Figure 22. Distribution graph of the Rasch scores for the PPVT test of the Younger Cohort – Peru

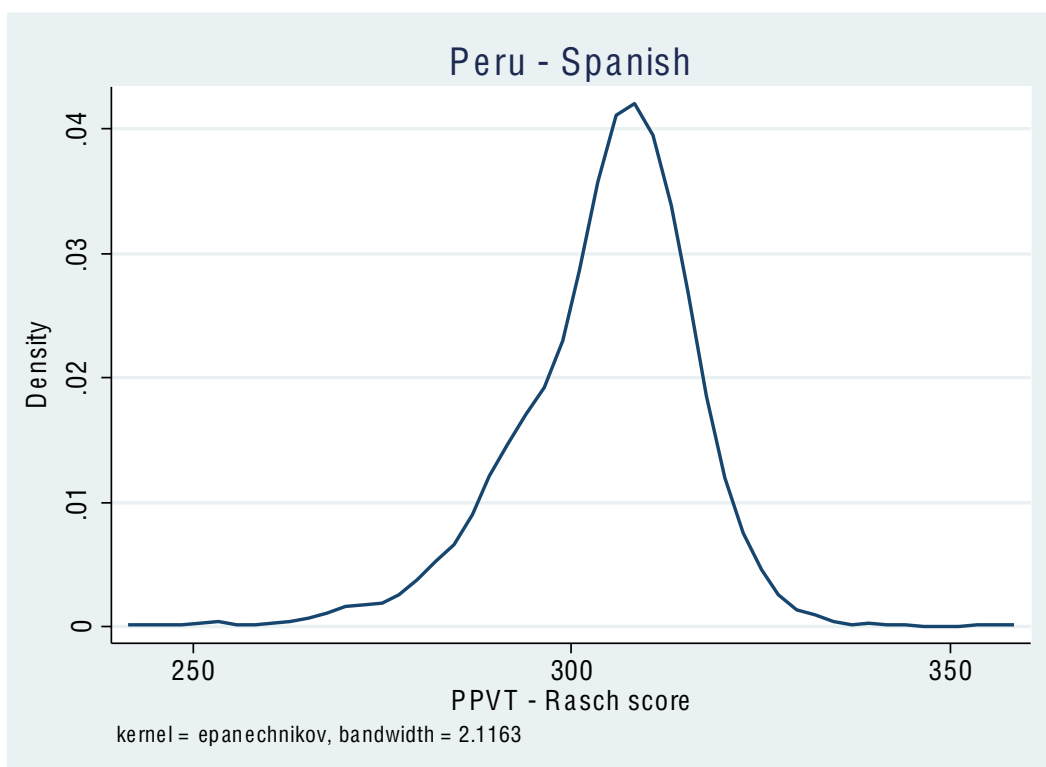


Figure 23. Distribution graph of the raw scores for the PPVT test of the Younger Cohort – Vietnam

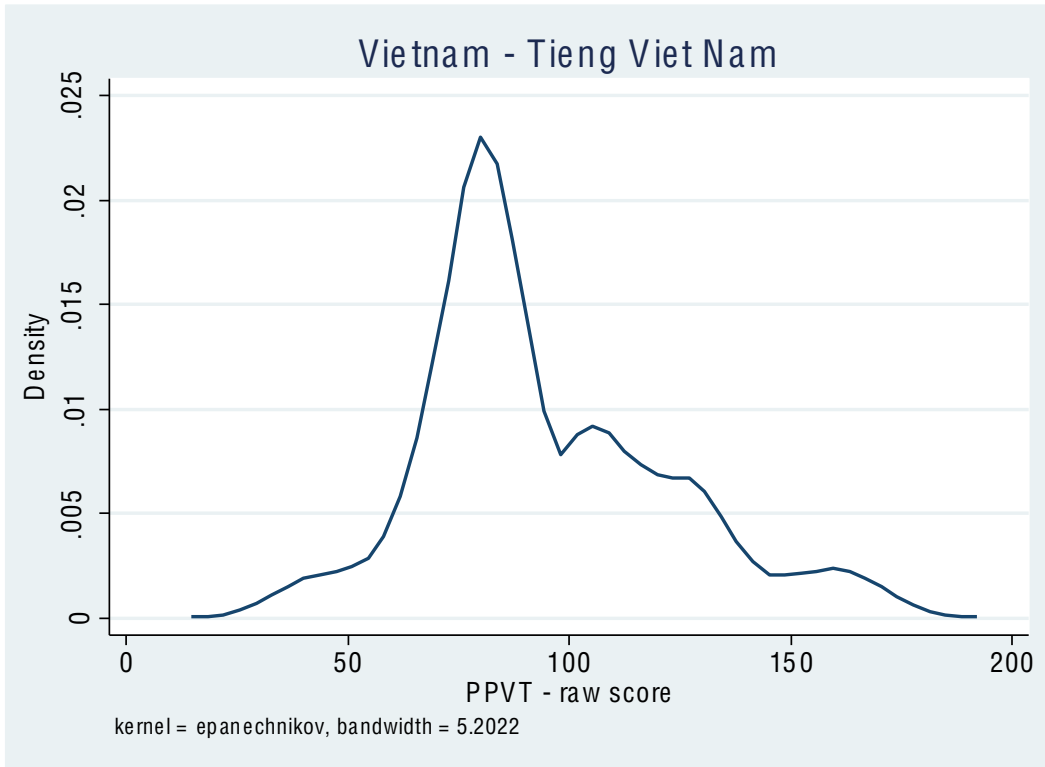


Figure 24. Distribution graph of the Rasch scores for the PPVT test of the Younger Cohort – Vietnam

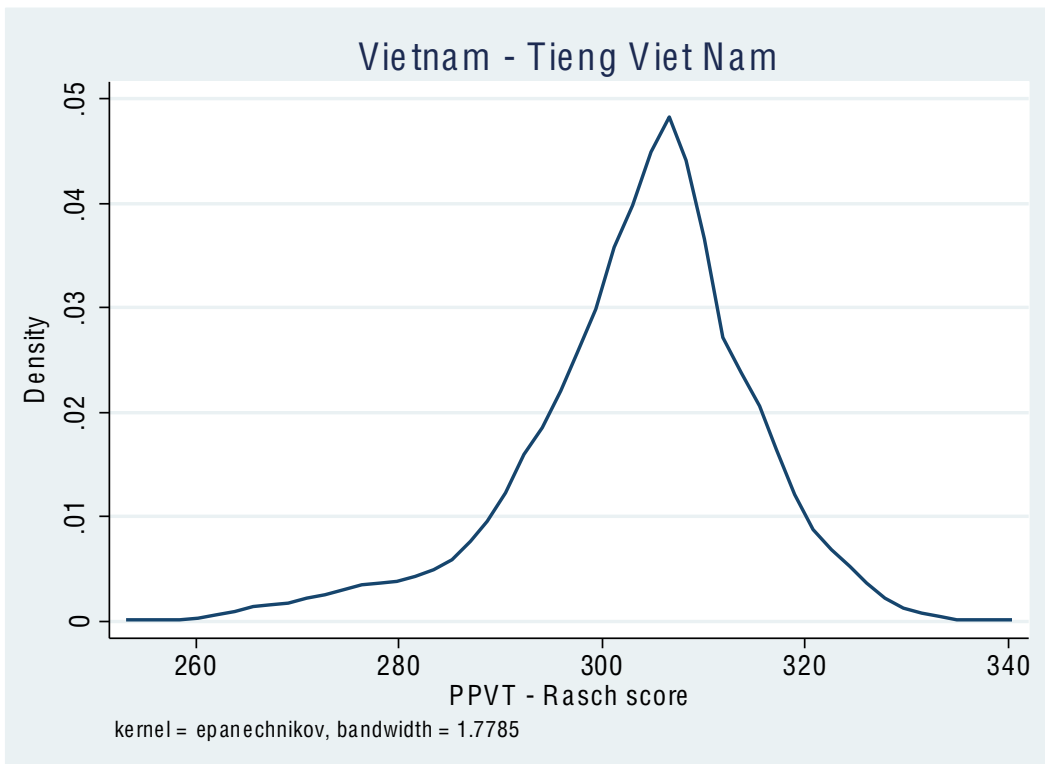


Figure 25. Distribution graph of the Raw scores for the Cloze test of the Older Cohort – Ethiopia

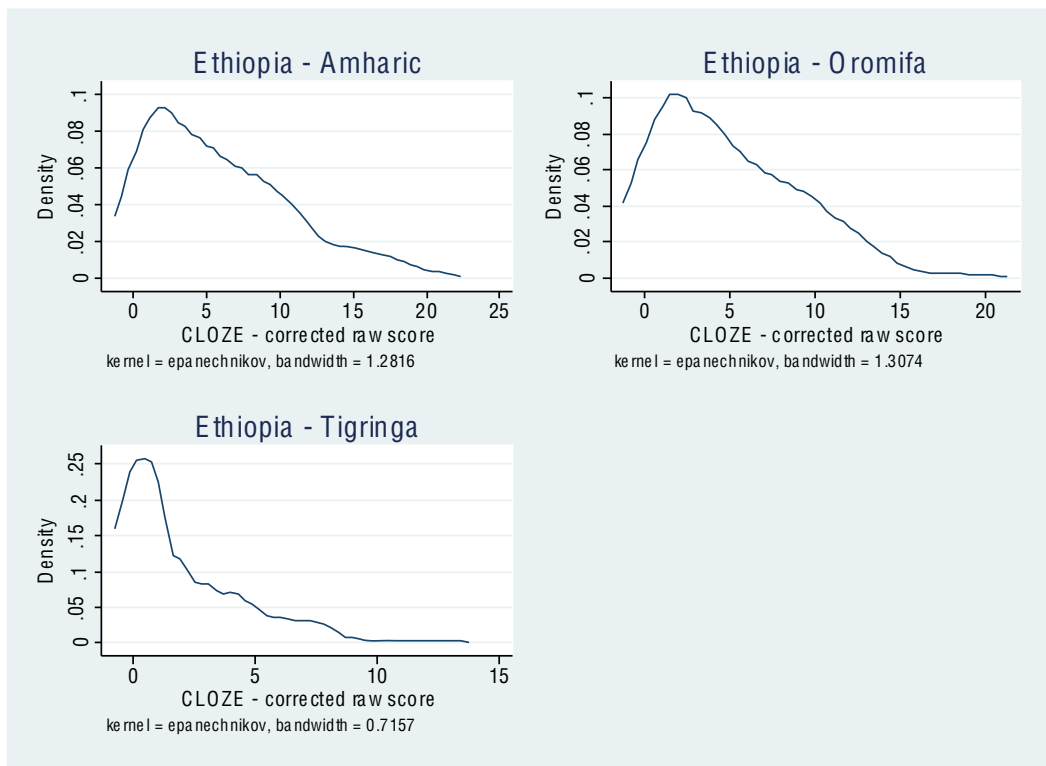


Figure 26. Distribution graph of the Rasch scores for the Cloze test of the Older Cohort – Ethiopia

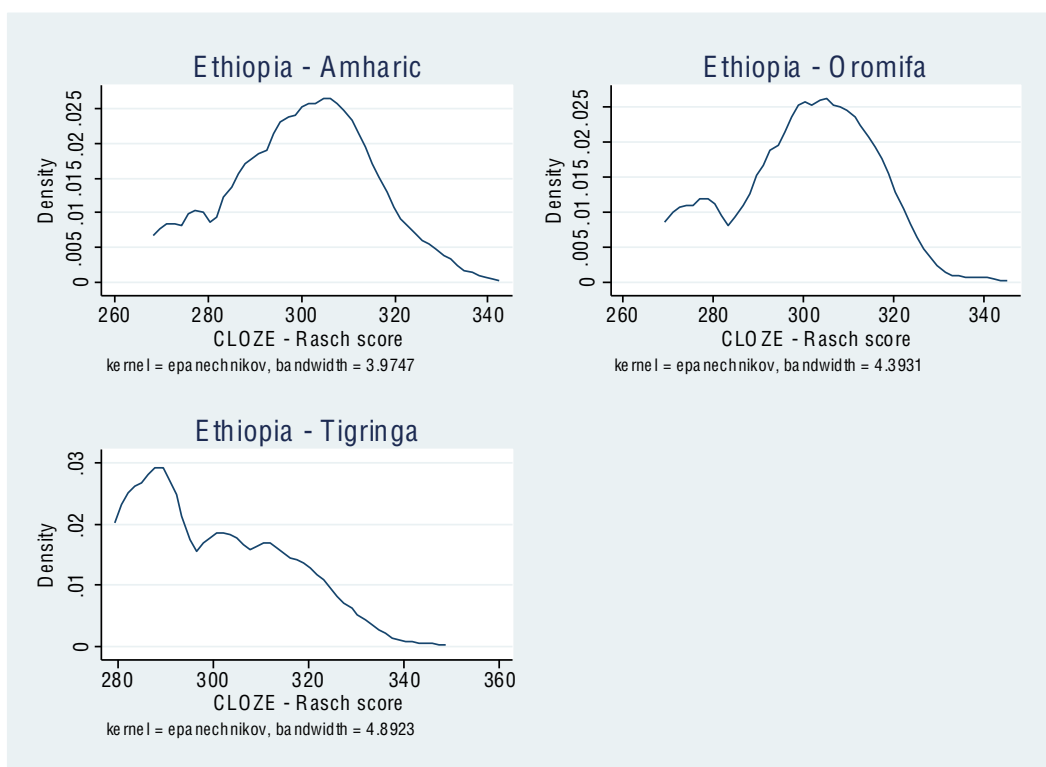


Figure 27. Distribution graph of the Raw scores for the Cloze test of the Older Cohort – India

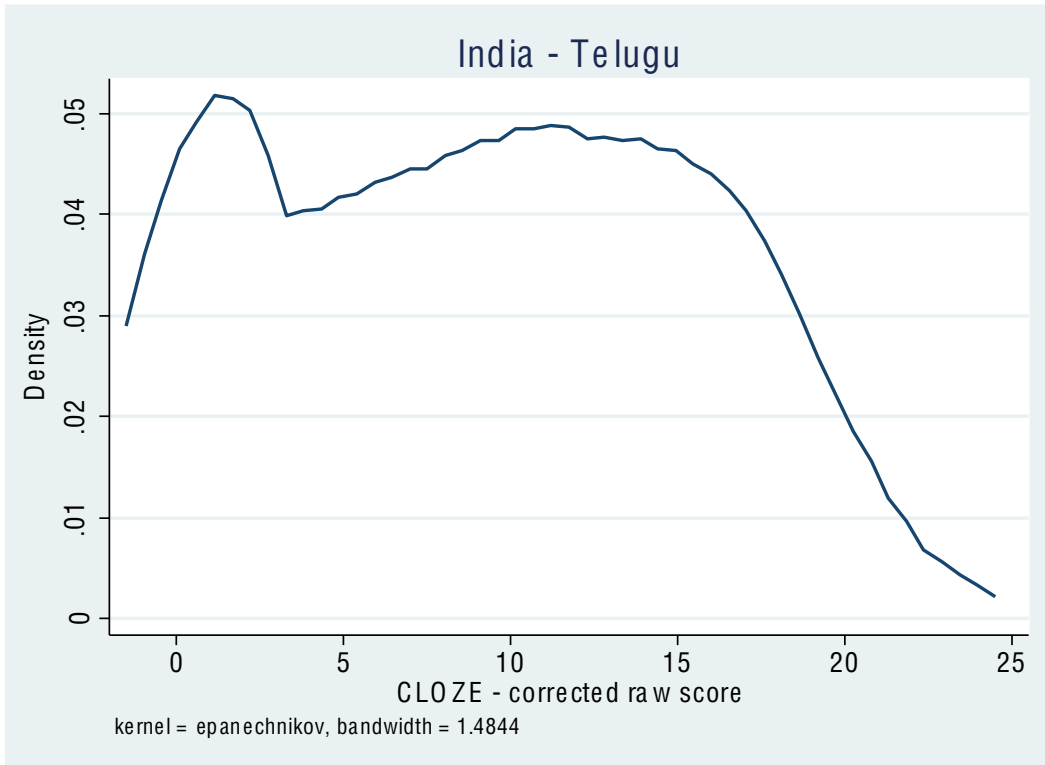


Figure 28. Distribution graph of the Rasch scores for the Cloze test of the Older Cohort – India

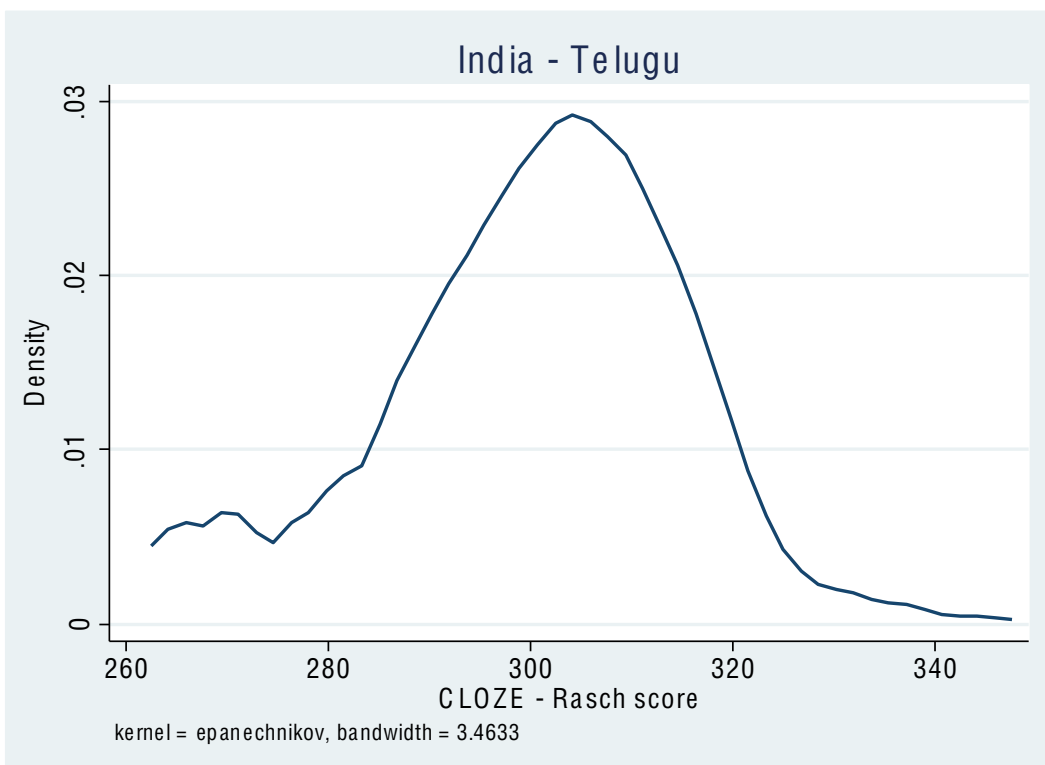


Figure 29. Distribution graph of the Raw scores for the Cloze test of the Older Cohort – Peru

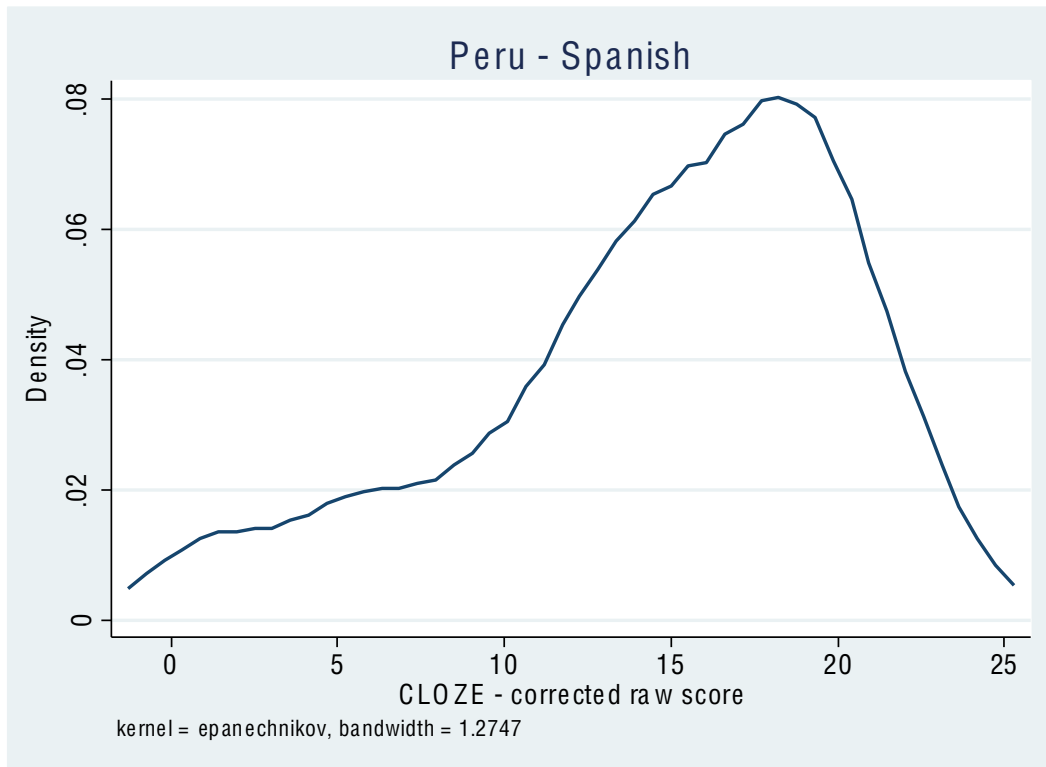


Figure 30. Distribution graph of the Rasch scores for the Cloze test of the Older Cohort – Peru

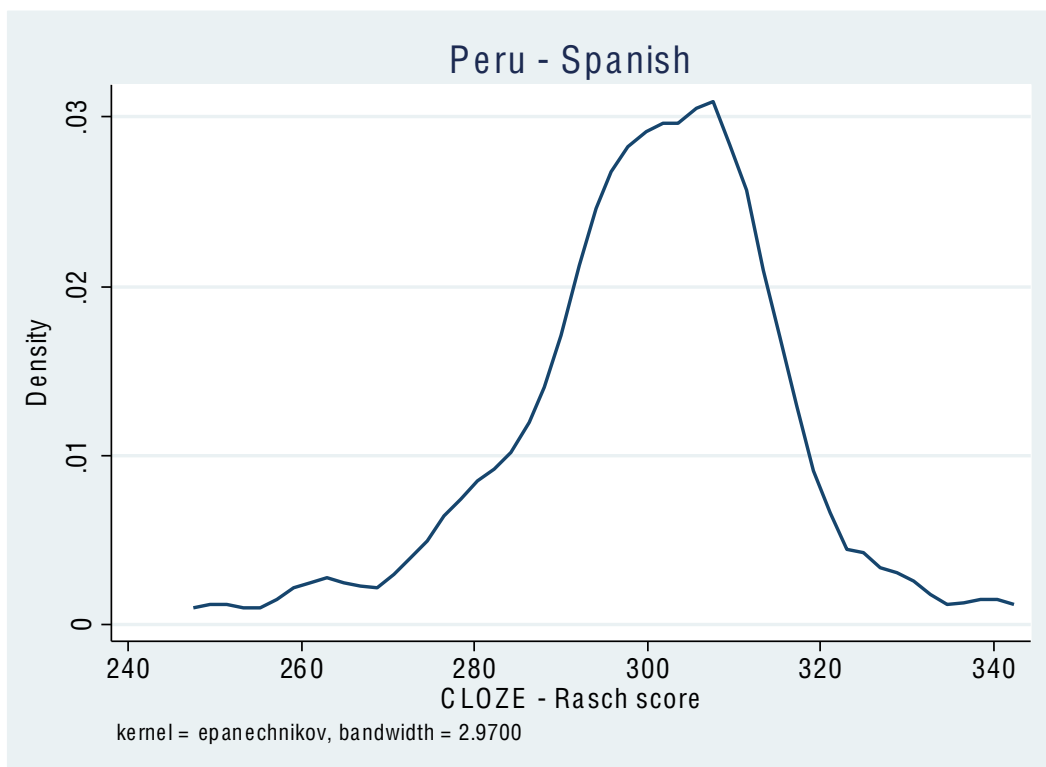


Figure 31. Distribution graph of the Raw scores for the Cloze test of the Older Cohort – Vietnam

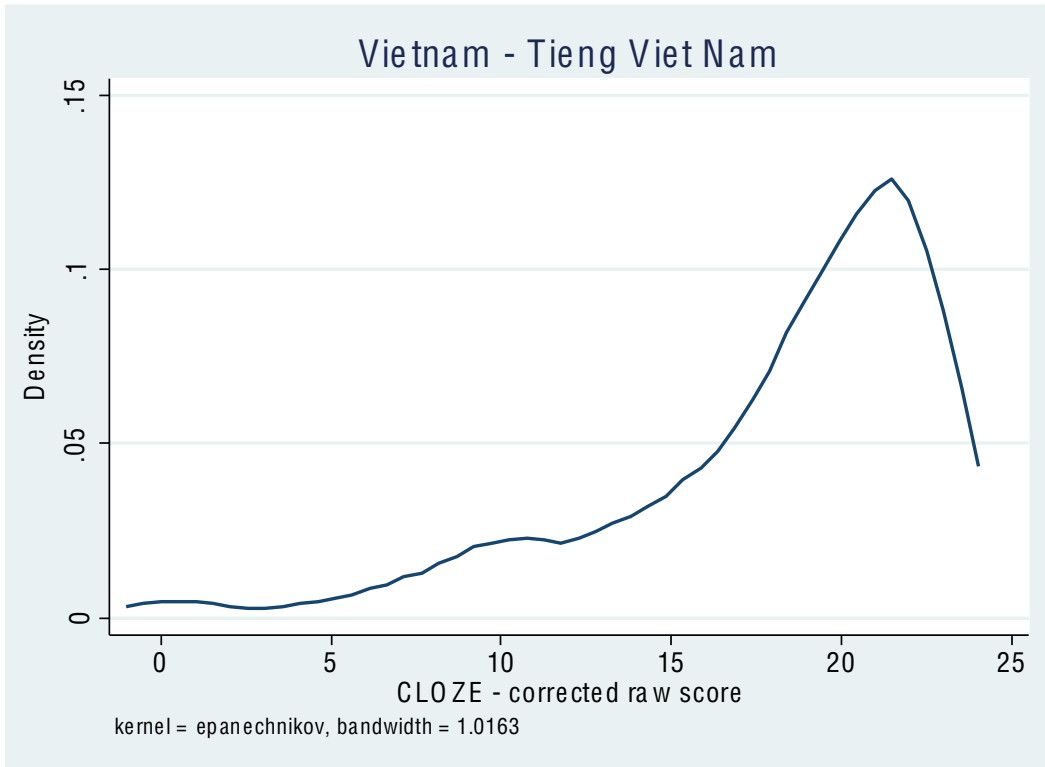


Figure 32. Distribution graph of the Rasch scores for the Cloze test of the Older Cohort – Vietnam

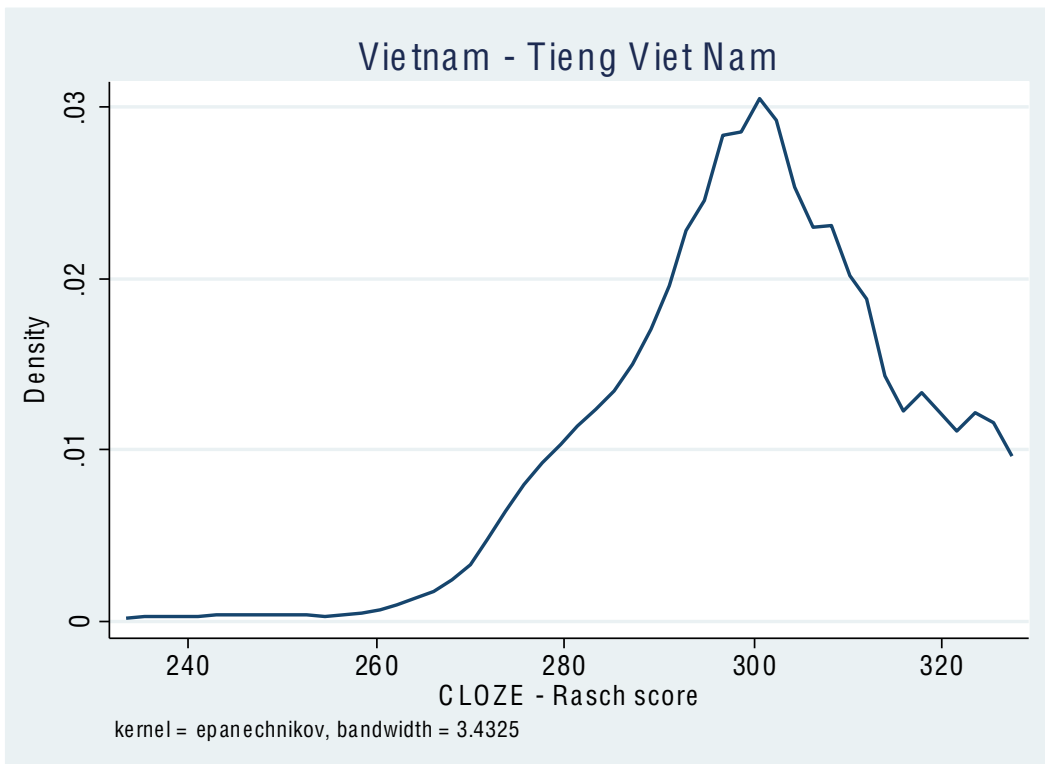


Figure 33. Distribution graph of the Raw scores for the Maths test of the Older Cohort – Ethiopia

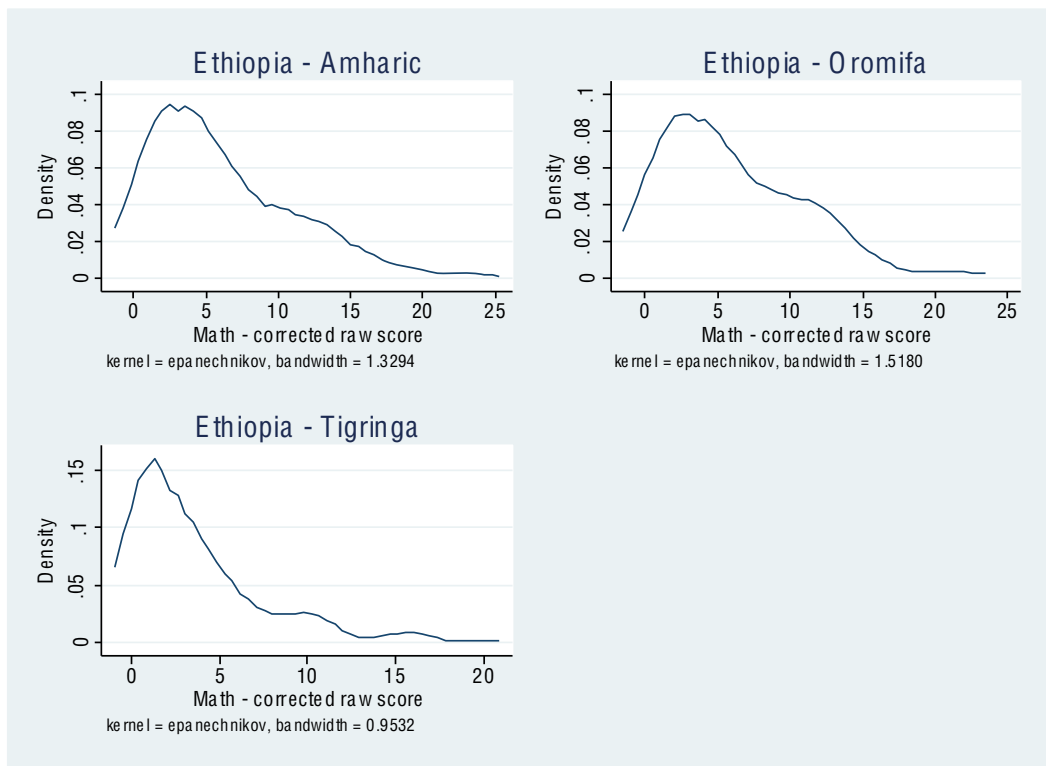


Figure 34. Distribution graph of the Rasch scores for the Maths test of the Older Cohort – Ethiopia

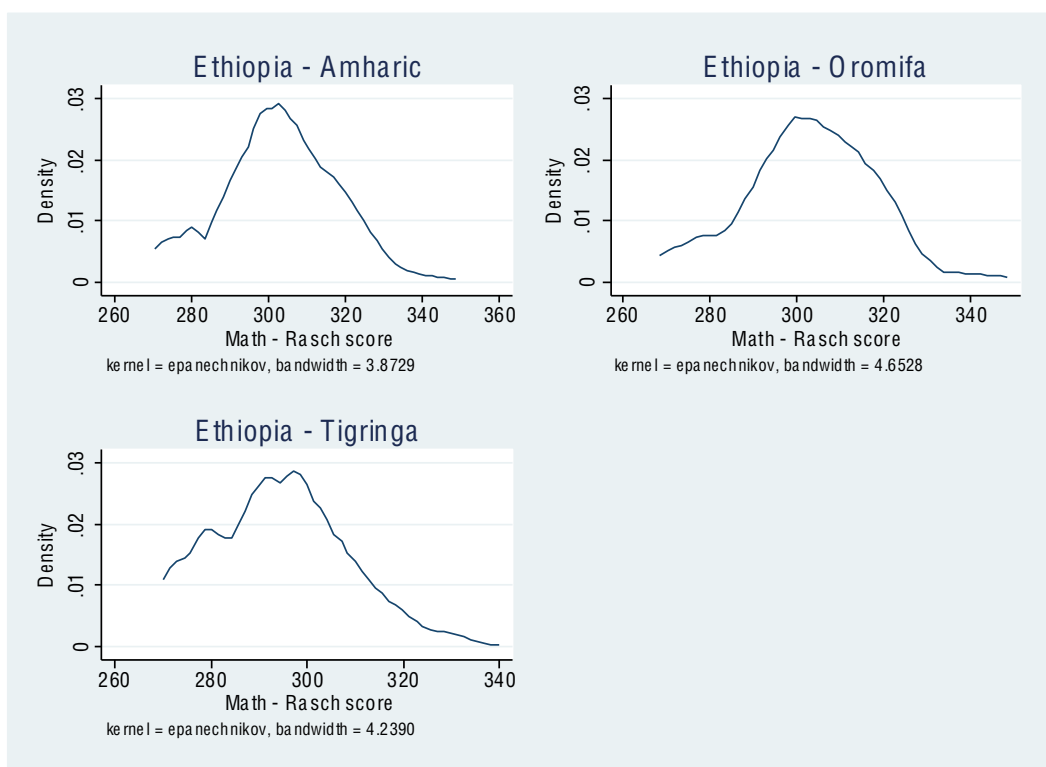


Figure 35. Distribution graph of the Raw scores for the Maths test of the Older Cohort – India

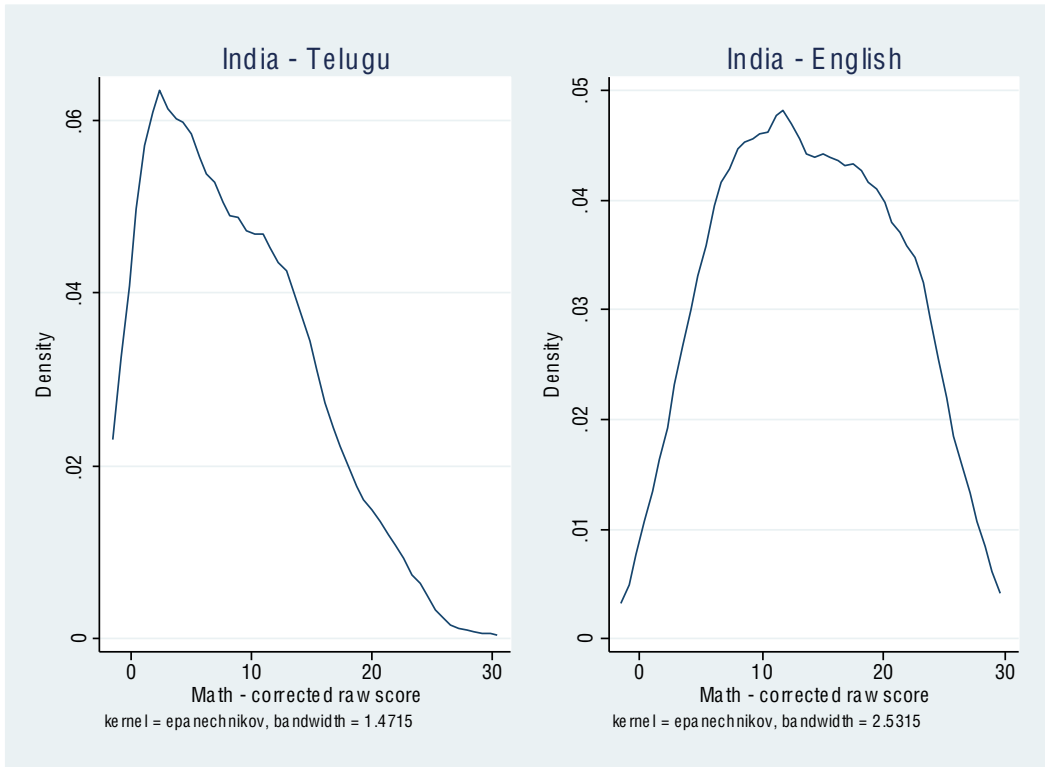


Figure 36. Distribution graph of the Rasch scores for the Maths test of the Older Cohort – India

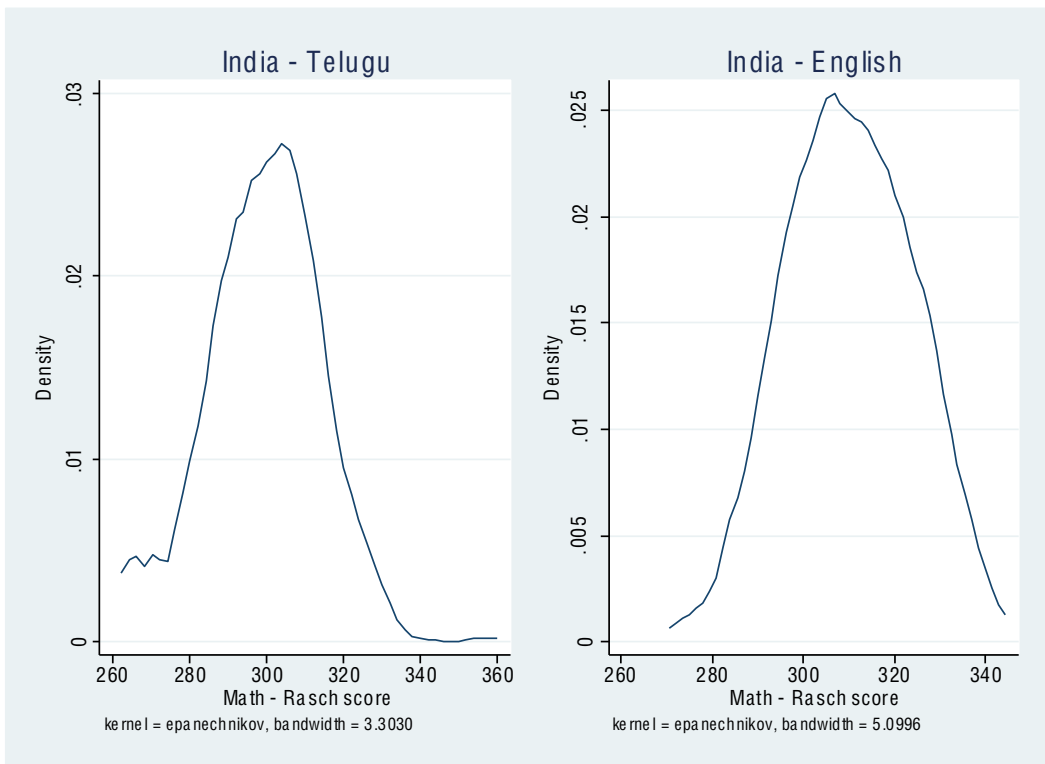


Figure 37. Distribution graph of the Raw scores for the Maths test of the Older Cohort – Peru

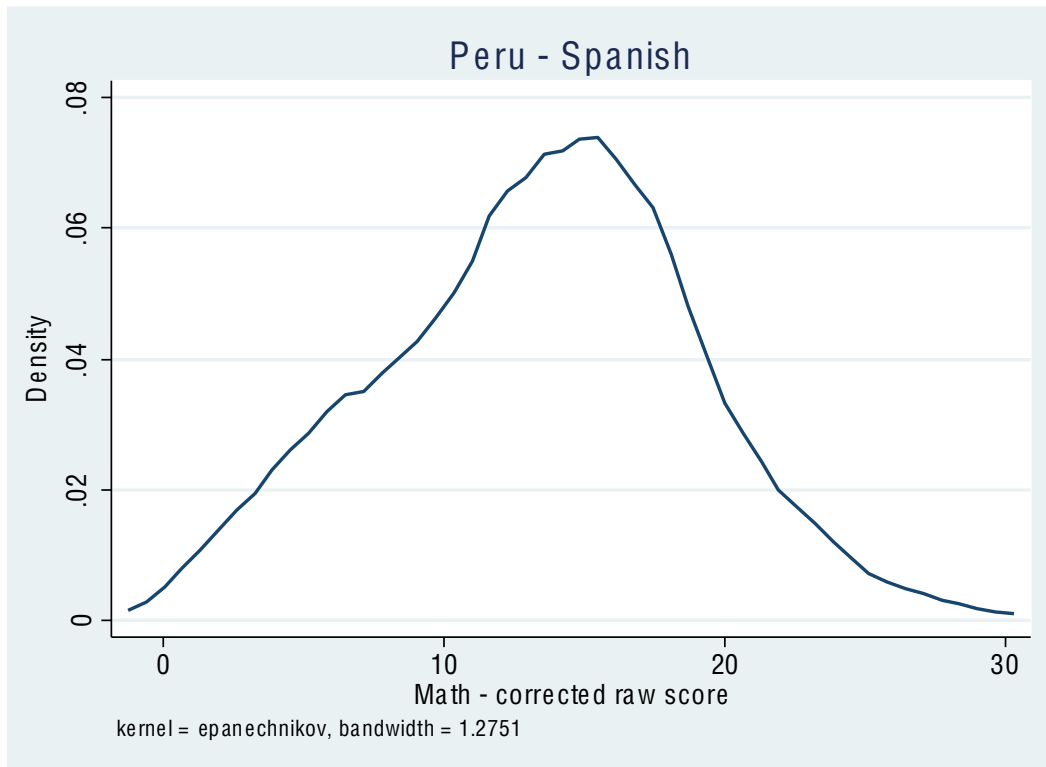


Figure 38. Distribution graph of the Rasch scores for the Maths test of the Older Cohort – Peru

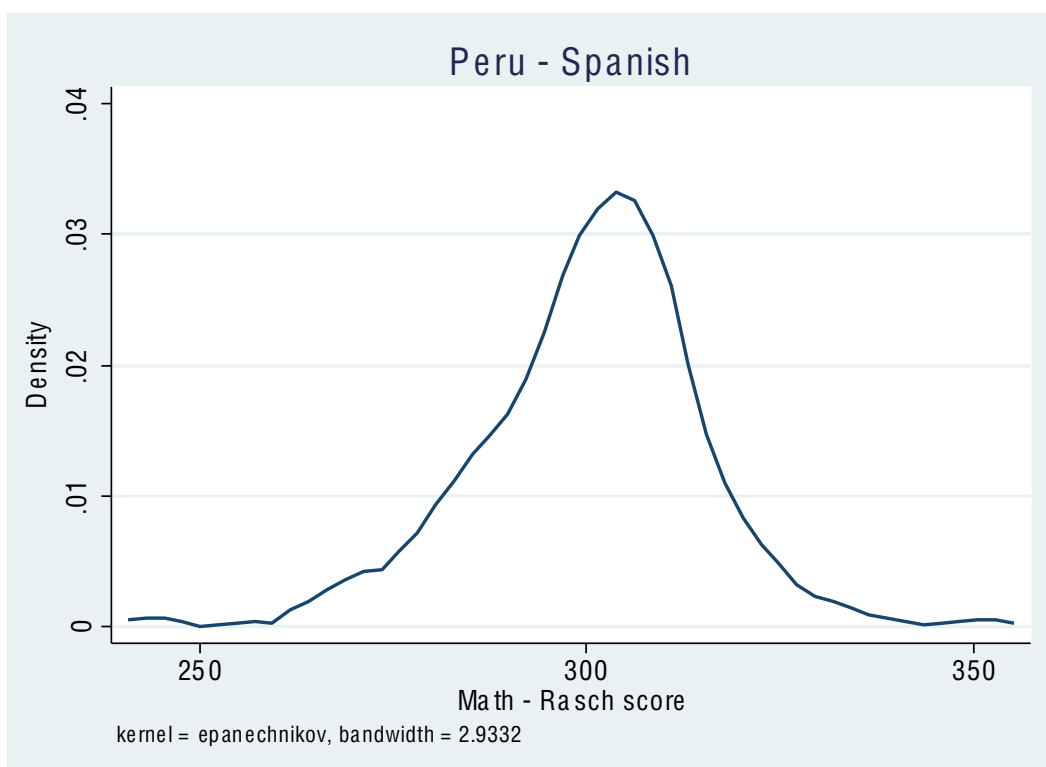


Figure 39. *Distribution graph of the Raw scores for the Maths test of the Older Cohort – Vietnam*

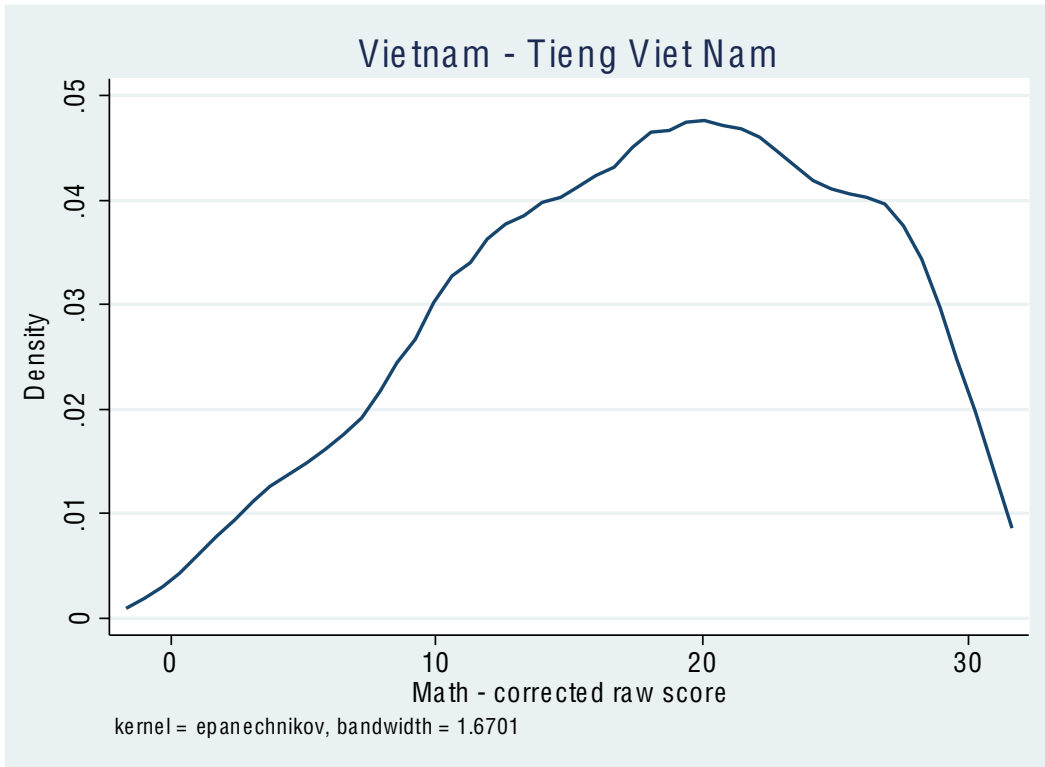


Figure 40. *Distribution graph of the Rasch scores for the Maths test of the Older Cohort – Vietnam*

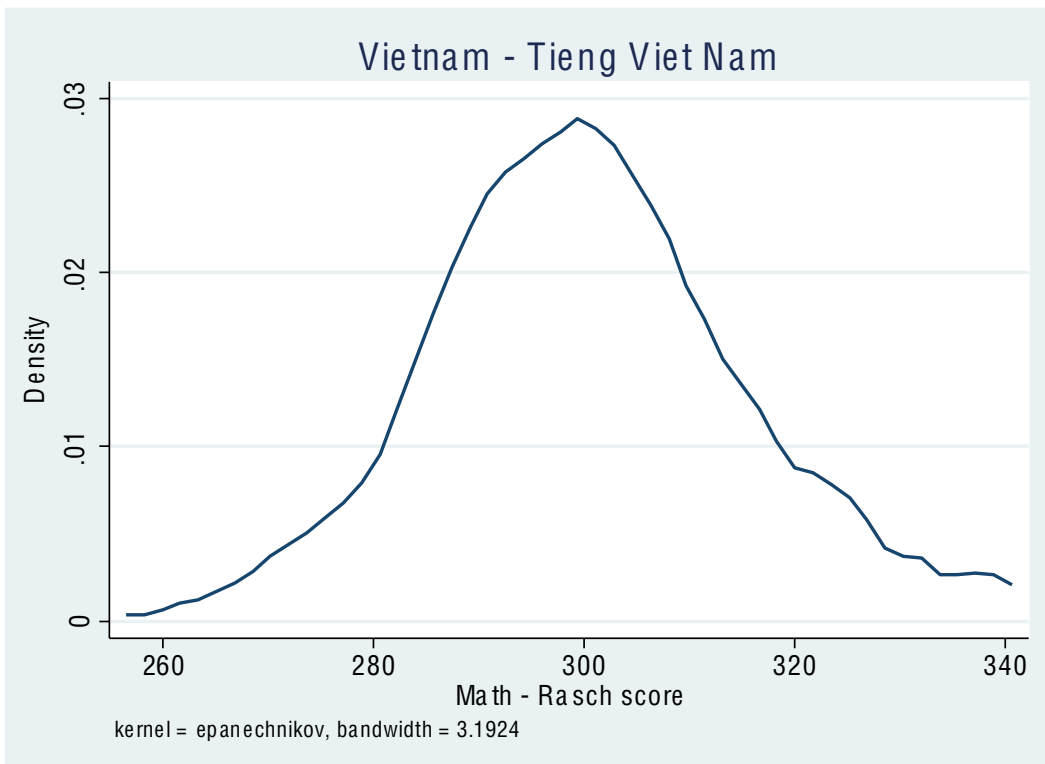


Figure 41. Distribution graph of the Raw scores for the PPVT test of the Older Cohort – Ethiopia

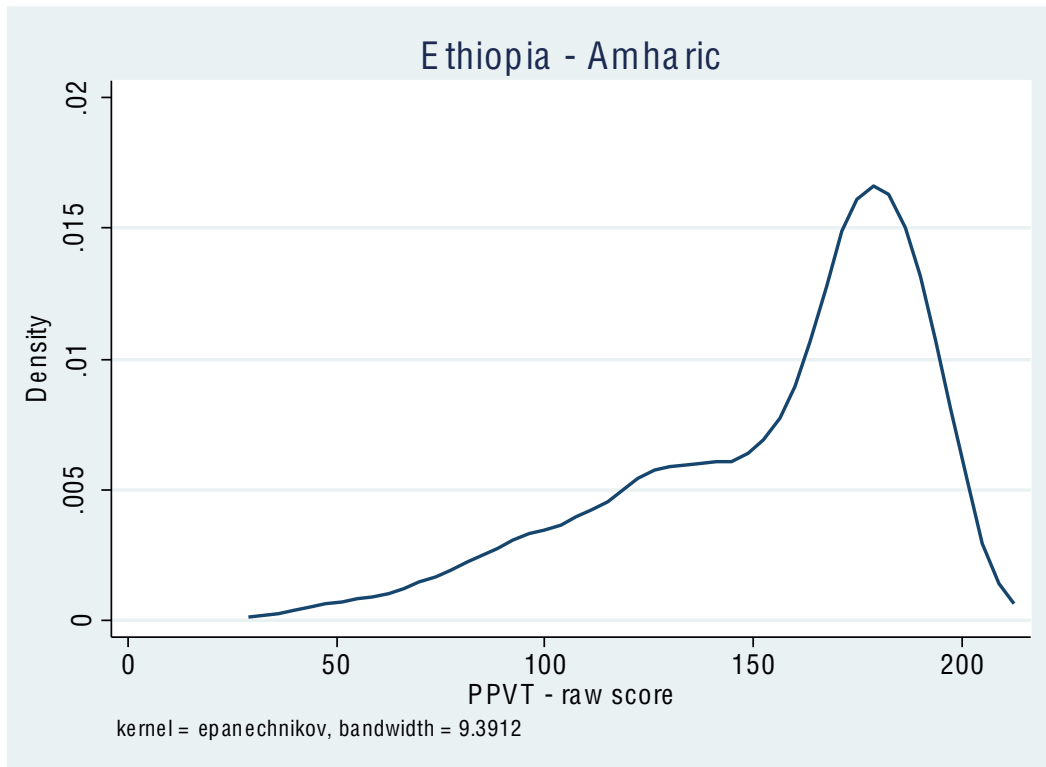


Figure 42. Distribution graph of the Rasch scores for the PPVT test of the Older Cohort – Ethiopia

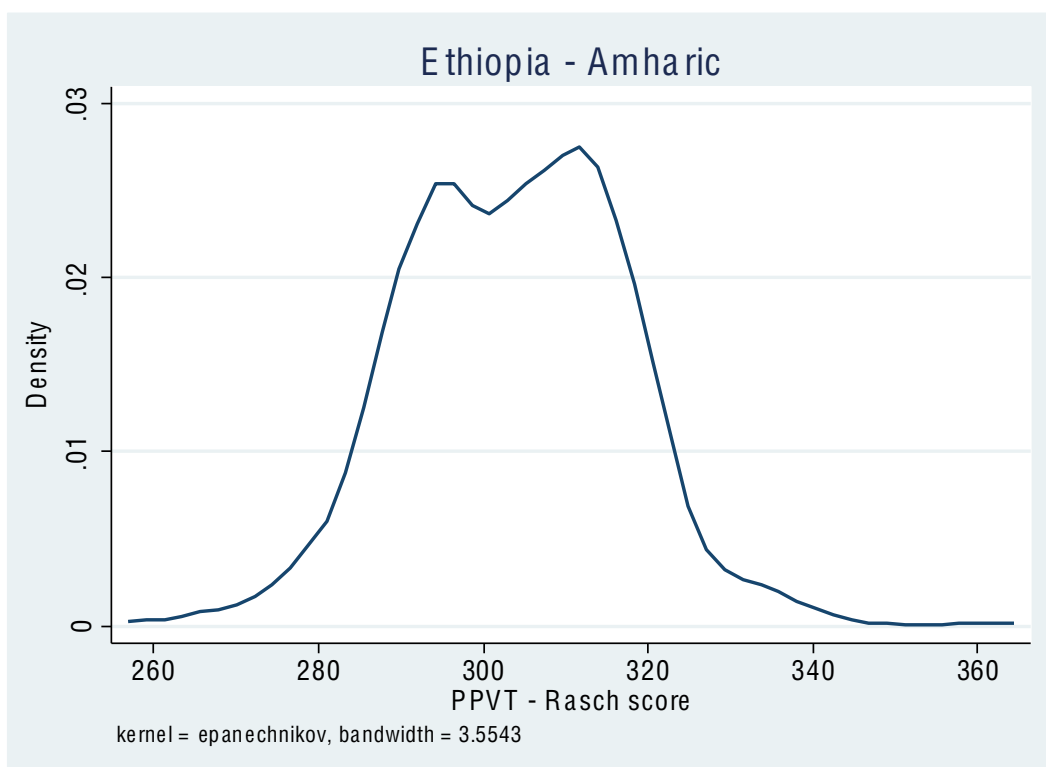


Figure 43. Distribution graph of the Raw scores for the PPVT test of the Older Cohort – India

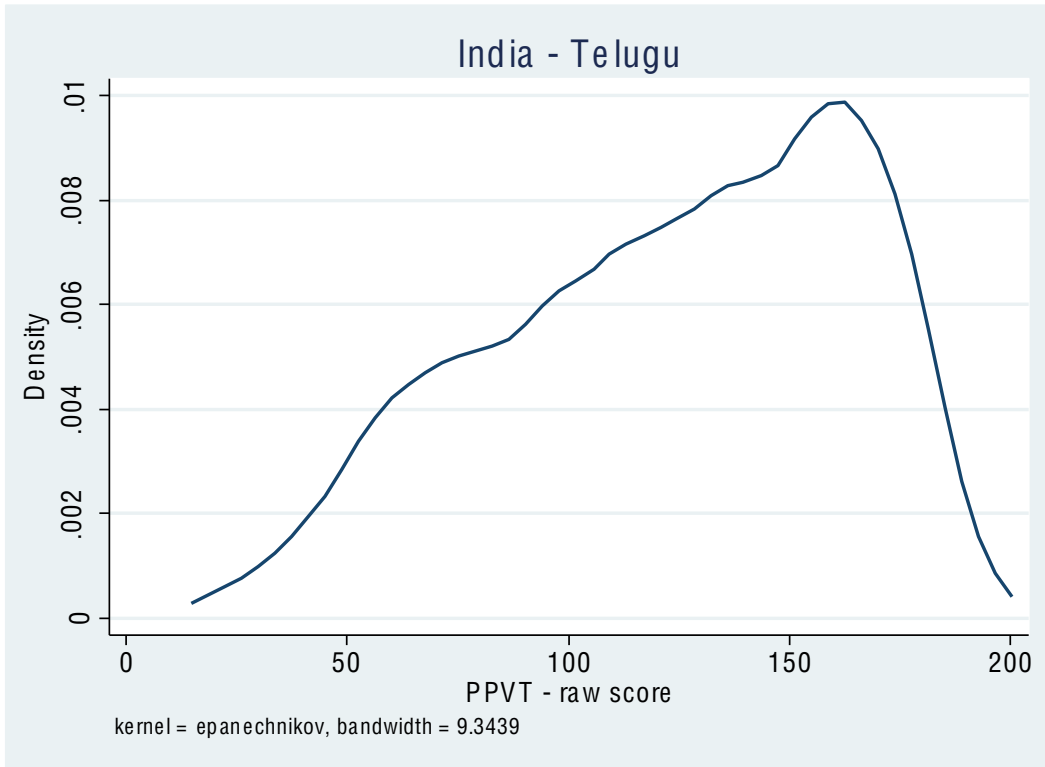


Figure 44. Distribution graph of the Rasch scores for the PPVT test of the Older Cohort – India

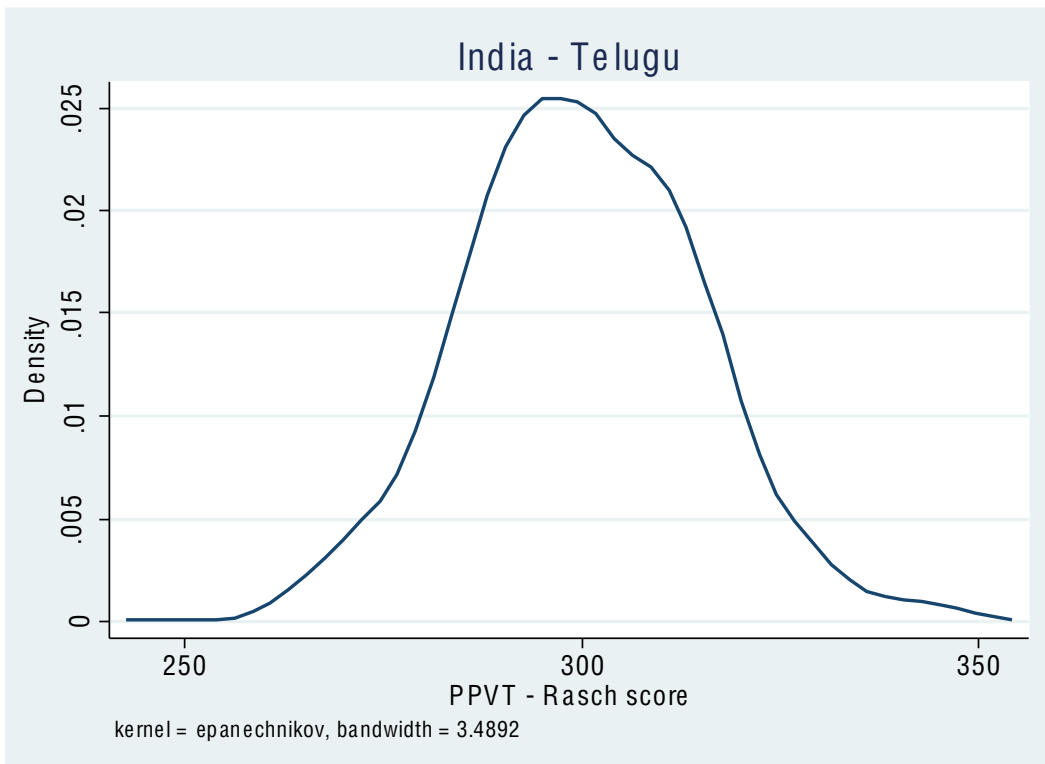


Figure 45. Distribution graph of the Raw scores for the PPVT test of the Older Cohort – Peru

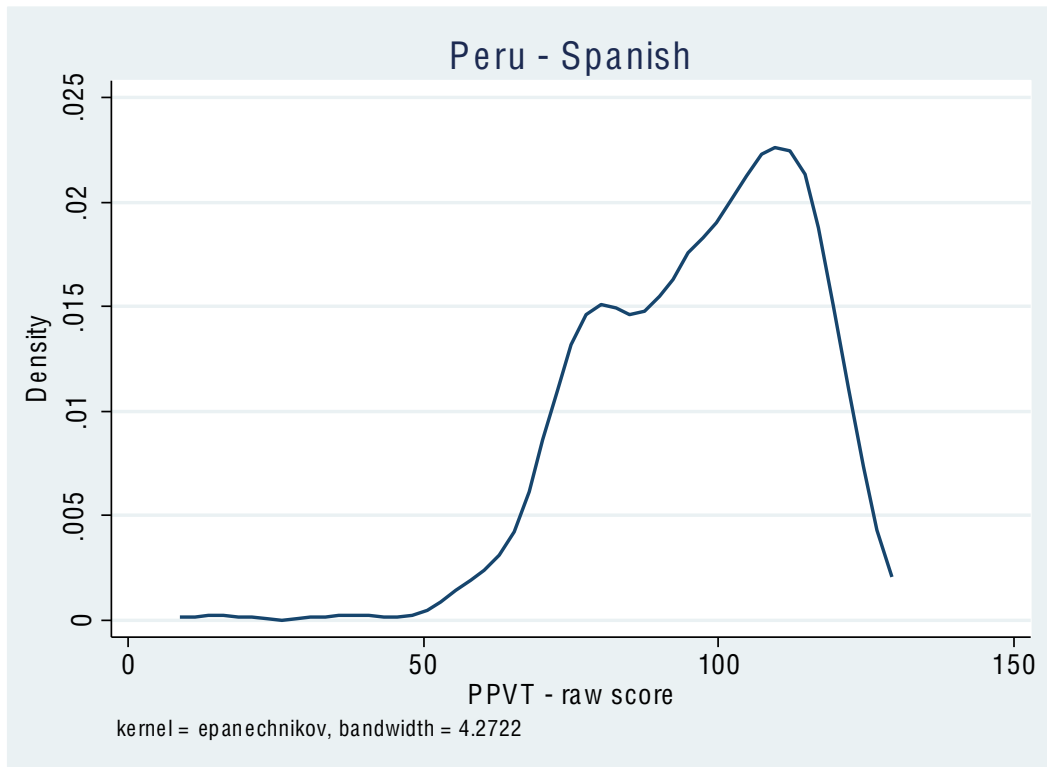


Figure 46. Distribution graph of the Rasch scores for the PPVT test of the Older Cohort – Peru

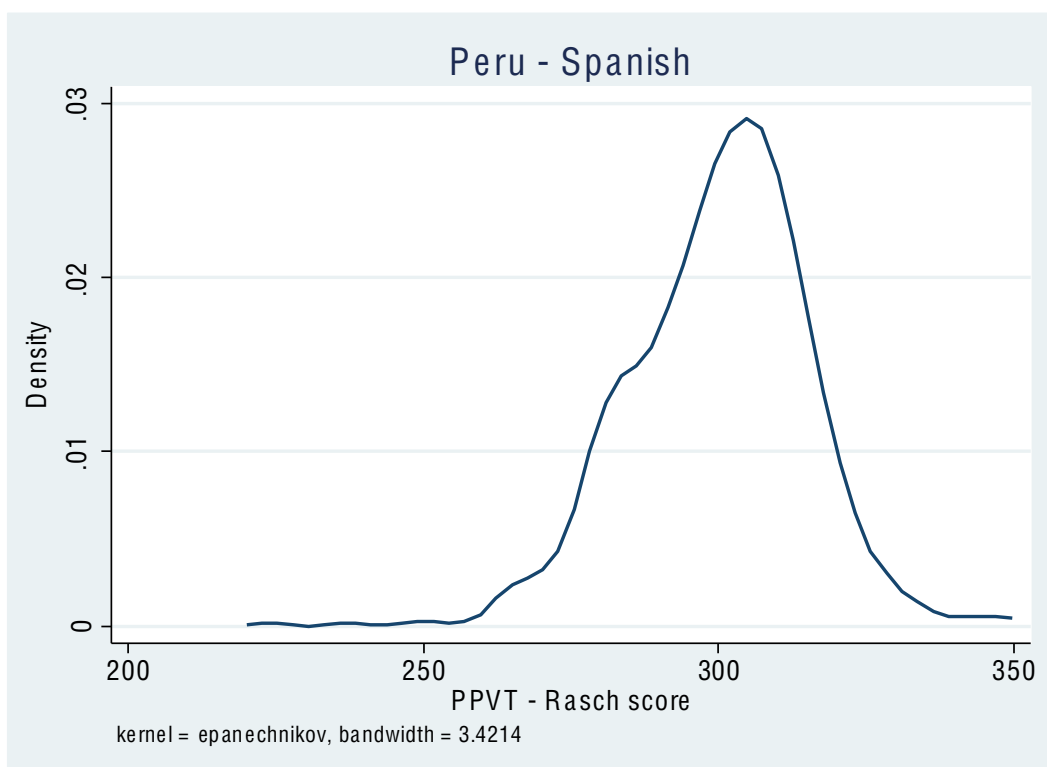


Figure 47. *Distribution graph of the Raw scores for the PPVT test of the Older Cohort – Vietnam*

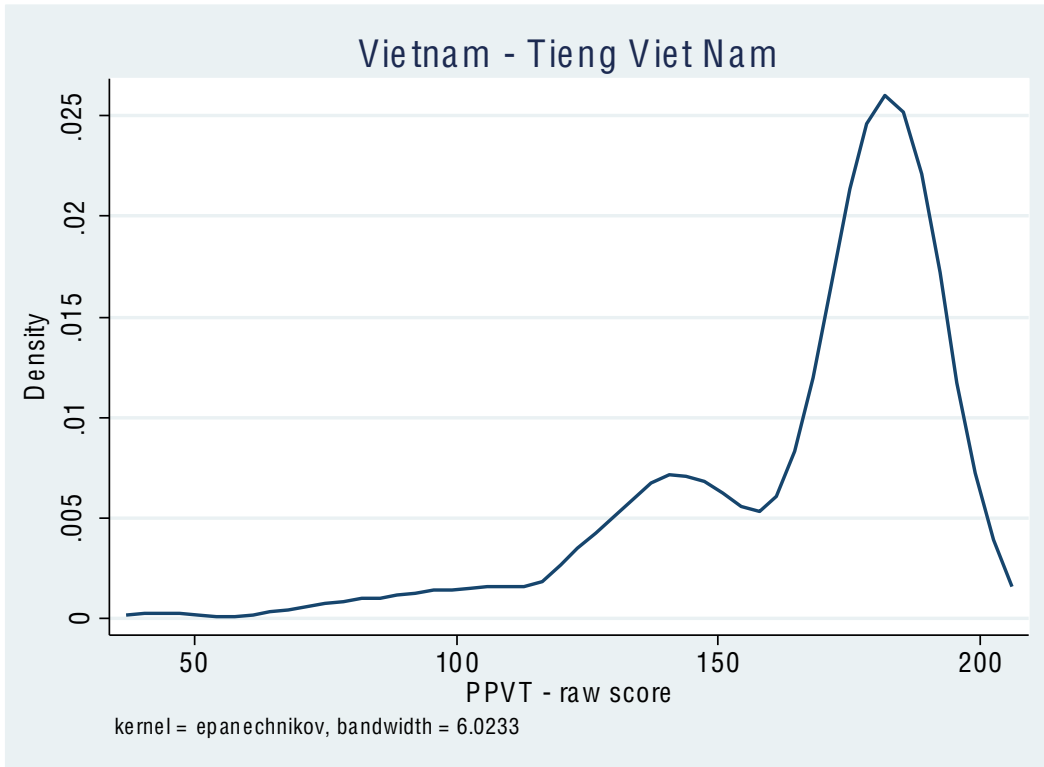
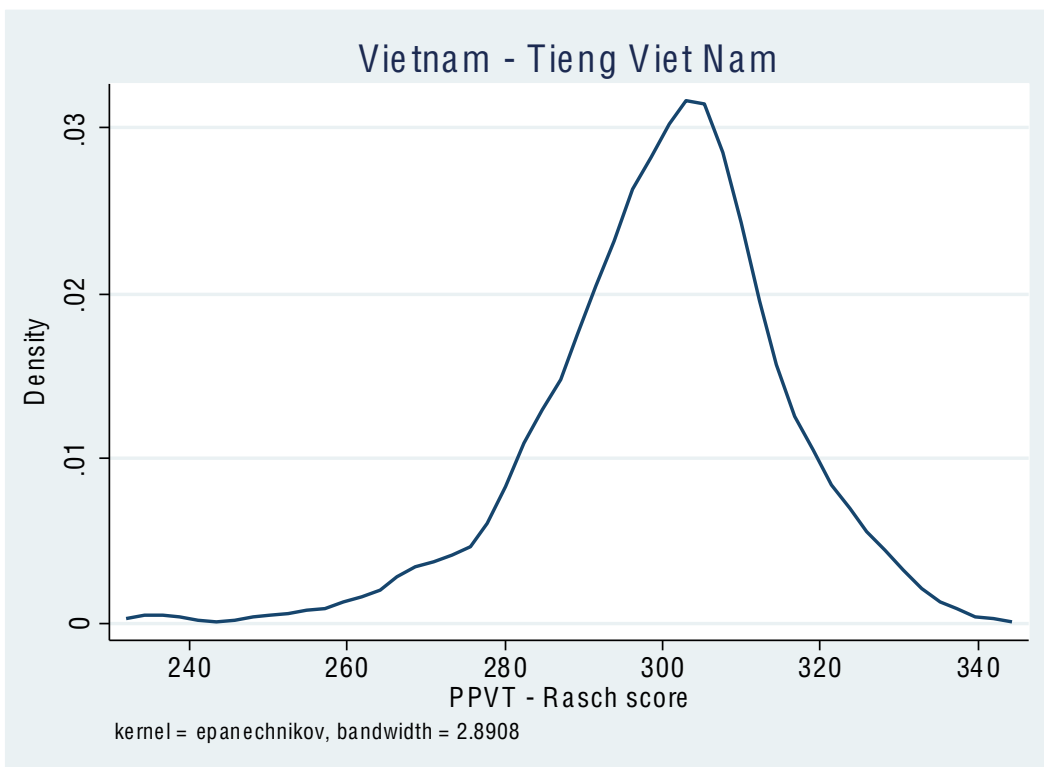


Figure 48. *Distribution graph of the Rasch scores for the PPVT test of the Older Cohort – Vietnam*



Funded by



Ministry of Foreign Affairs of the
Netherlands

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

Young Lives is core-funded from 2001 to 2017 by UK aid from the Department for International Development (DFID), and co-funded by the Netherlands Ministry of Foreign Affairs from 2010 to 2014. Sub-studies are funded by the Bernard van Leer Foundation and the Oak Foundation. 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.