Development of the ASER 2022 reading tool

Introduction

The Annual Status of Education Report (ASER) survey has been conducted since 2005. Driven by concerns about low levels of attainment in reading and mathematics across the elementary school years and beyond, ASER assessments were designed to capture basic learning outcomes using a common assessment for all children irrespective of age or grade.

A key aspect of the ASER survey is the availability of reliable longitudinal data to explore trends over time. It is therefore imperative to have comparable tools for all years of the ASER survey, while also controlling for overexposure of the test content. Given that the survey assesses children in their homes, alternate assessment forms (four test samples) are used with different children. The alternate test forms are designed such that each test form comprises a different set of items drawn from the same domain and based on the same content specifications (see below). Such a design is called a matrix sampling design and is appropriate when group-level estimates are reported rather than results for individual children, as is the case for the ASER survey. ASER reports estimates at the district and state level; it does not report child-level or village-level results.

Alignment with NCERT language textbooks

Since its inception, the ASER reading tool has been aligned with the state mandated curriculum and language textbook of each state in India. In the process of tool development, language textbooks for Std I, II and III in all states are analysed using various quantitative and qualitative metrics. This is done for several reasons. First, it helps to determine the type of text and level of difficulty children are exposed to in primary grades. Second, it helps to align the ASER assessment with the learning outcomes mandated by the National Council of Educational Research and Training (NCERT), which is important given that the reading tool is developed based on the learning outcomes a child is expected to achieve by the beginning of Std III. Lastly, given that teaching-learning activities in India are primarily based on textbooks, these books are used as the main source of guidance while developing the ASER reading tool.

Identification of curriculum changes since 2018

The 'basic' ASER survey was conducted in 2022 after a gap of 4 years. During this time, the National Education Policy (NEP) was also released. Therefore, although the process of 2022 reading tool development was broadly similar to the process followed in previous years, a preliminary nationwide exercise was conducted to track changes in expected learning outcomes, content of primary grade language textbooks, and teaching learning methods since the last ASER tool was developed in 2018.

Since this process was being undertaken the year after the release of the NEP, changes to textbooks, curriculum frameworks and/or learning outcome specifications in individual states were tracked by analysing government circulars/notifications, examining news articles, and speaking to government officials. Results of this exercise were documented for each state.

This exercise identified changes in the specification of expected learning outcomes in Madhya Pradesh, Uttar Pradesh, and Himachal Pradesh, as well as changes in Std I, II and III textbooks in Tamil Nadu, Andhra Pradesh, and Chhattisgarh. However, the nature of these changes did not require altering the assessment framework and tool development process for ASER 2022.

Textbook analysis

After tracking curriculum changes, a quantitative analysis of language textbooks was conducted for all states. The analysis consisted of several steps.

First, a list of all chapters was created for each textbook. Details such as language, textbook name, publication year, and total number of pages were recorded. Next, individual units/chapters were categorised as poetry, picture stories, fictional text in narrative format, and non-fictional text such as essays. This was a way to map different text genres and generate an overview of the kinds of texts children engage with in Std I, II and III. All units categorised as fictional text were further tagged under a sub-category, such as traditional fiction, contemporary fiction, and drama, among others.

The third step was to identify fictional texts written in narrative format, which is a written account of connected events that tell a story. This is the format used by the stories and paragraphs in the ASER reading assessment as they present an easy flow and syntax through a connected set of sentences, and are closer to 'real-world' purposes and audiences.

For this exercise a national training was conducted to orient the ASER state teams on the formats and definitions of each type of text available in language books for Std I, II and III.

Once this process was completed, a software-based analysis of all identified narrative texts was conducted to supplement and strengthen the earlier analysis of narrative texts identified in textbooks. The analysis provided information such as number of sentences, number of words, and average words per sentence. The frequency of occurrence of individual letters and words was also calculated. The words were then sorted into high, medium, and low frequency based on the number of times they occurred in the textbook. This was done to create a repository of letters and words to use for the development of the four tasks of the ASER reading tool – Letters, Words, Std I level text (Paragraph), and Std II level text (Story).

Development of the ASER 2022 reading tool

After completion of the textbook analysis, each state team created 8 paragraphs and 4 stories in their regional language. They were trained to incorporate the textbook analysis, and to adhere to qualitative and quantitative guidelines provided by the assessment team while creating stories and paragraphs. The word list used for story and paragraph development was generated from the software analysis conducted earlier. Quantitative factors such Type-Token Ratio (TTR), length and frequency of words, length and number of difficult words in a sentence were taken into account. Similarly, qualitative elements were specified, such as the rural context of the story, inclusion human characters, and elements such as a problem statement and solution.The software analysis was also used to verify that the texts met the guidelines for length, vocabulary and TTR indices.

These stories and paragraphs were sent to the assessment team, along with their English and Hindi translations. As a first step, the assessment team reviewed this content keeping in the mind qualitative and quantitative guidelines such as context, usage of words with conjoint sound/alphabets, usage of words from the word repository, gender-sensitive content, and ensuring no usage of first person.. After reviewing the stories and paragraphs from all states, the shortlisted ones were tweaked to align the structure, syntax, context and grammar with Std I and II level texts.

Shortlisted stories were then shared with senior ASER staff and linguistic experts to select a total of 4 stories and 8 paragraphs. These stories and paragraphs were transadapted in the 19 languages (including Hindi and English) in which the ASER 2022 survey was administered.

Transadaptation

Transadaption means the process of preparing content created in one language and culture for use in a second language and culture. It is different from literal translation where context may or may not be taken care of. A form of forward translation method, transadaptation is considered best practice for multilingual assessments. The content in the source language has to be both translated as well as adapted to fit the need and cultural or linguistic requirements of the target language.

The process of transadaptation involves:

- Evaluating how well the source language content "fits" in the target language based on linguistic and cultural factors

- Replacing items that don't make sense in the target language with appropriate items.

This ensures the linguistic quality of the tool by making it equivalent in different languages.

A process of transadaptation was followed for ASER 2022 along with other guidelines which kept the tool consistent and comparable with tools created in earlier years. To transadapt the stories and paragraphs for 2022, state teams first analysed the stories and paragraphs from the ASER 2018 reading tool on various parameters such as structure (character, sequence of events, problem, solution and ending), total number of words, unique words, and words with conjoint sound/alphabets. With 2018 stories and paragraphs as the point of reference, teams began the process of transadaptation.

After completing the transadaptation process, stories were compared using parameters like usage of grade appropriate vocabulary, length of sentences, and number of complex and simple sentences. The transadapted stories were once again translated back to English and Hindi, enabling the central assessment team to review the texts.

This exercise led to the creation of the ASER reading tool in 19 languages, 4 samples in each, for a total of 76 new stories and 152 new paragraphs.

Establishing test form comparability

The ASER surveys employ four test forms, representing content from the same domain and developed using pre-specified criteria. The expectation is that a child's performance on reading or arithmetic will not be unduly affected by the particular test form administered. In other words, children should be ranked at the same level of reading or arithmetic irrespective of the test form administered to them.

Once the 4 new samples were created for ASER 2022, test form comparability was assessed using the alternate form method, which requires administering different test forms to the same children without major lags between testing occasions, to minimise changes in performance that may stem from changes in children's ability levels.

A total of 72 children were assessed for each of the 19 languages. The sample was equally distributed between children in Std III, IV, V and VI – i.e., 18 children per grade. Assessment was conducted in schools to reduce the implementation time. Each child was tested on 5 reading tool samples – sample 1 of 2018 and samples 1, 2, 3, 4 of 2022. In order to avoid test fatigue, 3 samples were administered before lunch, and 2 after lunch. The tool administration was similar to the household testing process, and the highest level that each child could read at was recorded for each test form individually. Words that children who were fluent readers (at 'story' level) found difficult were also captured to understand the vocabulary load of the reading task.

The evaluation of alternate form reliability is estimated by assessing the agreement in ranking decisions across alternate forms; a process similar to evaluating agreement across repeated test administrations.[1] The results show that there was agreement in marking the child according to their reading proficiency across the five samples. The agreement was seen in more than 90% cases for a particular language.

[1] See Swaminathan, Hambleton, Ronald K. Hambleton, and James Algina. "Reliability of criterion-referenced tests: A decision-theoretic formulation." Journal of Educational Measurement 11.4 (1974): 263-267; and Traub, Ross E., and Glenn L. Rowley. "Reliability of test scores and decisions." Applied Psychological Measurement 4.4 (1980): 517-545.