Understanding *Anupad*: Preliminary Findings from the Analysis of a Mathematics Diagnostic Assessment Tool in India

Nayanatara Gautham, Jioo Nimkar

Quality Education Support Trust (Quest)

Persistent gaps in foundational numeracy remain a significant challenge in Indian classrooms, with many students progressing through grades without mastering basic mathematical concepts. As a result, learning levels often vary widely within the same classroom, making grade-level assessments insufficient for informing instruction. Diagnostic assessments can help teachers identify students' conceptual understanding and learning gaps to support targeted instruction. This paper presents preliminary findings from the standardisation process and analysis of Anupad, a mathematics diagnostic assessment developed as part of a level-based learning programme implemented by the Quality Education Support Trust (QUEST) in Maharashtra, India. Data from 2,323 students in Grades 6-8 across 27 project sites were analysed. Results indicated high internal consistency (Cronbach's $\alpha = 0.88$) and strong content validity. Factor analysis examining the underlying structure of the data identified a two-factor model representing procedural and conceptual understanding, accounting for 32% of the variance. These findings provide preliminary evidence supporting the reliability of the assessment.

Keywords: diagnostic assessment; learning gap; mathematics education; Anupad; India

Introduction and Background

Evidence from across India indicates persistent gaps in foundational numeracy among school students (Sthapak, 2025). National surveys and research studies consistently show that many children struggle with basic mathematical concepts even after several years of schooling. For example, the ASER Centre through the Annual Status of Education Report has repeatedly demonstrated that a substantial proportion of children in upper primary grades are unable to perform basic arithmetic operations such as subtraction or division (ASER Centre, 2019). In Maharashtra, only 23% of Grade 3 students were able to solve subtraction problems, while many Grade 5 students struggled with basic reading and numeracy skills.

These learning gaps are compounded by the structure of classroom instruction, which typically follow a grade-level curriculum irrespective of students' current learning levels. Consequently, many students progress through grades without mastering foundational concepts. Within a single classroom, learners often exhibit wide variation in mathematical understanding, making it difficult for teachers to provide instruction that effectively supports all students. Research has shown that such mismatches between instruction and students' learning levels can contribute significantly to low learning outcomes (Banerjee et al., 2007; World Bank, 2017).

In response, level-based learning approaches have emerged as a potential strategy; however, contextual factors must also be taken into account (Am et al., 2023). These approaches group students according to their current learning levels rather than grade level and provide targeted instruction to address specific gaps.

Diagnostic assessments play a critical role to such an approach as they help teachers identify students' existing knowledge and skills (Das & Singhal, 2021).

Existing Mathematics Assessment Tools

Globally, a range of assessment tools have been developed to measure student learning, including Programme for International Student Assessment and International Development and Early Learning Assessment (Pisani et al., 2018). However, these assessments originate in contexts that differ substantially from those in India in terms of socio-cultural conditions and access to education. As a result, their design and item content may be perceived as misaligned or potentially inequitable (Banerji et al., 2013). This concern is reflected in India's performance in 2009, when it ranked among the lowest-performing participants.

Within India, large-scale assessments such as the Annual Status of Education Report (ASER) have been administered to over 600,000 students annually since 2005 (Banerji et al., 2013). While ASER provides valuable insights, it primarily measures children's ability to perform basic arithmetic. In contrast, the National Education Policy 2020 conceptualises foundational numeracy more broadly, encompassing counting, arithmetic, and wider mathematical thinking (NEP, 2020, p. 9). This underscores the need for assessments that capture multiple domains of mathematical understanding rather than focusing solely on basic arithmetic.

To better identify and address specific learning gaps, it is therefore important to examine distinct dimensions of mathematical learning. In this context, the present study focuses on the Anupad tool, developed in Maharashtra, India, by the Quality Education Support Trust, which is designed to assess students' mathematical competency.

Anupad Program

The Anupad program, developed by the Quality Education Support Trust (QUEST), is one such initiative aimed at addressing learning gaps in mathematics among middle school students. The program typically runs for approximately three years and involves working with clusters of 30–35 teachers. It focuses on students in Grades 5-8 who are performing below grade level in mathematics.

A key component of the program is the use of diagnostic assessments to determine students' current learning levels. These assessments are not intended for grading or ranking students. Instead, they aim to identify conceptual gaps and inform instructional planning. Based on the assessment results, students are grouped according to their learning levels, and teachers provide targeted instruction for approximately 45 minutes each day. The program also includes several additional components designed to support teachers and students. Teachers participate in capacity-building workshops that focus on understanding student thinking and implementing level-based instruction. They also receive regular onsite support from program facilitators. Instructional materials such as Anupad booklets and storybooks are used to support learning activities.

Within this broader program, the Anupad Assessment serves as a diagnostic tool for identifying students' mathematical understanding. The assessment aims to capture a range of competencies, including procedural skills, conceptual understanding, and familiarity with mathematical conventions. Given the importance of diagnostic assessments in supporting instructional decision-making, it is essential to examine the psychometric properties of such tools. This study therefore analyses the reliability and factor structure of the Anupad assessment to better understand its measurement properties.

Methodology

Anupad Assessment Design

The Anupad mathematics assessment was designed as a diagnostic tool to measure students' understanding of key mathematical concepts relevant to upper primary grades. The assessment was developed in Marathi and is administered in Marathi, the medium of instruction in the government schools of Maharashtra. The assessment consisted of 64 questions covering a range of topics including number sense, arithmetic operations, fractions, measurement, and data handling. Most questions were multiple-choice items scored as correct or incorrect. Fourteen of the questions were calculation-based and were graded using a rubric and awarded a score between 0 and 2 marks.

Sample

The dataset analysed in this study consisted of responses from 2,698 students in Grades 6-8 (approximately 11–14 years old). The students were drawn from various different project sites across Maharashtra, India where the Anupad program was implemented. The dataset included tests conducted a mix of data from mid-point (n=1,055) and end-point (n=1,643). Here we used the mid and post data since the intervention was giving students input to get students to their appropriate grade level (e.g. Fabrigar & Wegener, 2012). Initially, data from 2,698 students were available. However, responses from students who were absent during one of the testing rounds or who did not complete the assessment were excluded from the analysis (valid n =2,323).

Preliminary Findings

Descriptive Statistics

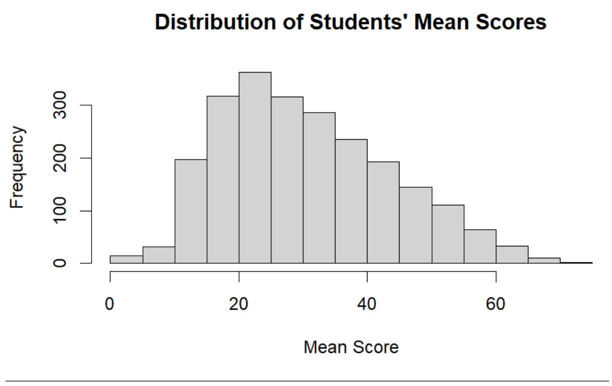


Figure 1.

The instrument consisted of 64 items. Total scores ranged from 0 to 75 (maximum possible = 80), with a mean of 30.91 (SD = 13.2) and median 29. The distribution of total scores showed slight positive skew (0.47) and was approximately mesokurtic (−0.37), indicating no substantial deviation from normality. Figure 1 shows the distribution of student scores.

Item-level analysis of standardised scores showed that Question 1 was the easiest item (M=0.82, SD=0.38), while Question 52 was the most difficult (M=0.12, SD=0.33).

This indicates a broad spread of item difficulty, suggesting that the assessment included both accessible and challenging questions.

Internal Consistency

The scale demonstrated high internal consistency, with a Cronbach's alpha of 0.88 (standardized $\alpha = 0.87$).

Content Validity

The questionnaire was developed through expert review involving two practising teachers, one teacher educator, and two education researchers. Although the instrument does not encompass every aspect of the curriculum, it was aligned with grade-level learning objectives such that students who are able to respond successfully to the items are likely to demonstrate competence in the expected content for their grade.

In designing the instrument, high levels of literacy and numeracy were not assumed. Consequently, the language of the items was deliberately kept concise and accessible. At the same time, careful attention was given to preserving the role of linguistic comprehension in mathematical reasoning, recognising that understanding the language of mathematical problems is integral to mathematical thinking itself.

Taken together, these considerations support a strong degree of content validity for the questionnaire.

Factor Analysis

Preliminary analyses were conducted to examine the suitability of the data for factor analysis. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was 0.94, exceeding recommended thresholds and indicating high sampling adequacy. Bartlett's test of sphericity was significant, $\chi^2(2016) = 27995.97$, $p < .001$, confirming that the variables were sufficiently intercorrelated. These findings support the appropriateness of proceeding with factor analysis.

Parallel analysis and an inspection of the screen plot indicated that 2 factors were sufficient (first excluded eigen value was 2.27), (Henson & Roberts, 2006). See Figure 2

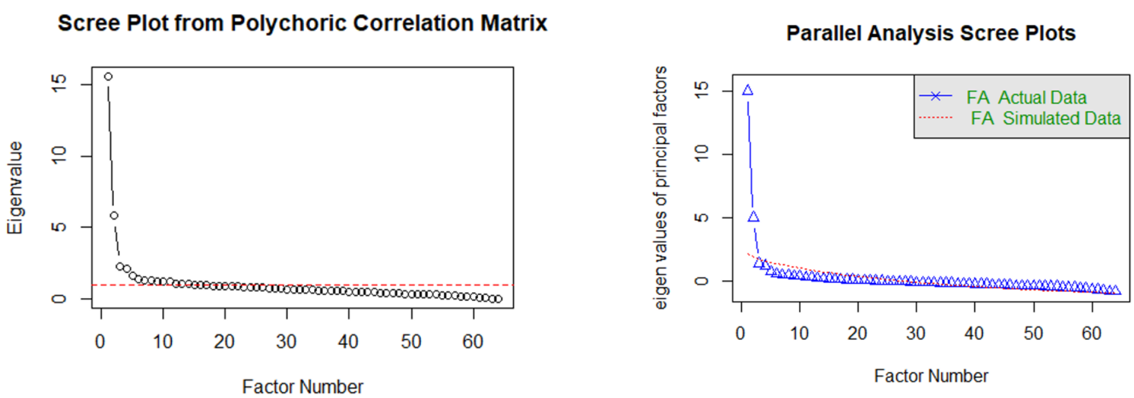


Figure 2

Initially, an oblimin rotation was applied. However, the correlation between factors was negligible (r = 0.06), indicating limited shared variance. Moreover, both oblimin and varimax rotations accounted for the same proportion of explained variance (32%). Given that the varimax solution also demonstrated stronger internal consistency, it was retained for subsequent analysis. A loading cut-off of 0.30 was applied.

The first factor was labelled *procedural understanding and conventions in mathematics*. It comprised of 36 items (Items 1–3, 5–9, 11, 12, 14, 16–22, 25, 27–32, 35, 37–39, 41–43, 47, 53, 55, and 56) and accounted for 21% of the total variance. Below are two examples of such questions from the questionnaire (See Fig 3). The internal consistency of this factor was high, with a Cronbach's alpha of 0.88.

Write the correct symbol from **>, <, =** in the blank between the given numbers.

4321 ☐ 4231

Kanchan bought a shawl for ₹130 and a handkerchief for ₹14. She paid with a ₹500 note. How much change should she receive?

Options

a. **₹356**

b. **₹144**

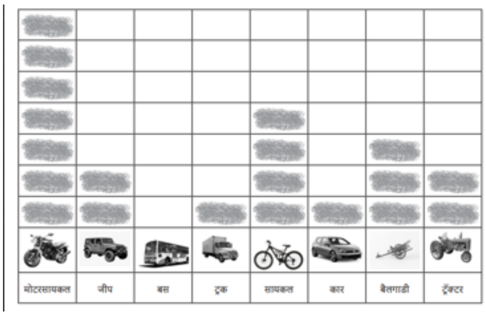
c. **₹370**

d. **₹354**

Figure 3

The second factor was labelled *conceptual understanding*. It comprised of 13 items (Items 4, 10, 13, 15, 23, 24, 26, 33, 34, 40, 44, and 48) and accounted for 11% of the total variance. See Figure 4 for an example. The internal consistency of this factor was acceptable, with a Cronbach's alpha of 0.72.

The pictograph showing the vehicles that passed in front of your school in 5 minutes. Based on this graph, what conclusions can we draw?



a. This school is likely to be located in the middle of a big city.

b. Students at this school like to ride motorcycles.

c. Buses never pass in front of the school.

d. It is likely that the school is located in a village.

Figure 4

Conclusion and Future Plans

This paper has examined the mathematics component of the Anupad assessment developed by the Quality Education Support Trust (QUEST), analysing its internal consistency, content validity, and underlying factor structure. Both internal consistency and content validity were found to be high.

Exploratory factor analysis indicated that a two-factor structure best explained the data, with the factors labelled as *procedural understanding and conventions*, and *conceptual understanding*. The two-factor solution with varimax rotation accounted for 31.8% of the variance. Model fit indices were mixed: RMSEA (0.087) and RMSR (0.049) suggested an acceptable to marginal fit, while the TLI (0.578) indicated a weaker fit relative to conventional thresholds. Taken together, these indices suggest that although the two-factor model provides an economical and interpretable structure, however it does not fully capture the complexity of the data. Overall, these findings provide initial evidence supporting the robustness of the tool as a measure of students' mathematical understanding.

Future research could extend this work by examining additional aspects of reliability and validity, including test-retest reliability. Furthermore, the application of error taxonomy analysis may offer deeper insights into students' conceptual understanding and help identify specific learning gaps.

References

Am, Muh. A., Hadi, S., Istiyono, E., & Retnawati, H. (2023). Does differentiated instruction affect learning outcome Systematic review and meta-analysis. *Journal of Pedagogical Research*, 5. https://doi.org/10.33902/JPR.202322021

ASER Centre. (2019). *Annual Status of Education Report (Rural) 2018.* ASER Centre.https://img.asercentre.org/docs/ASER%202018/Release%20Material/aserreport2018.pdf

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The quarterly journal of economics*, *122*(3), 1235-1264. https://doi.org/10.1162/qjec.122.3.1235

Banerji, R., Bhattacharjea, S., & Wadhwa, W. (2013). The Annual Status of Education Report (ASER). *Research in Comparative and International Education*, *8*(3), 387–396. https://doi.org/10.2304/rcie.2013.8.3.387

Das, U., & Singhal, K. (2023). Solving it correctly: Prevalence and persistence of gender gap in basic mathematics in rural India. *International Journal of Educational Development*, *96*, 102703. https://doi.org/10.1016/j.ijedudev.2022.102703

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory Factor Analysis*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199734177.001.0001

Henson, R. K., & Roberts, J. K. (2006). Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement*, *66*(3), 393–416. https://doi.org/10.1177/0013164405282485

National Council of Educational Research and Training. (2019). *National Achievement Survey (NAS) 2017: National report to inform policy, practices and teachin learning.* Ministry of Human Resource Development, Government of India. https://ncert.nic.in/pdf/NAS/National_Report_NAS_2017.pdf

Pisani, L., Borisova, I., & Dowd, A. J. (2018). Developing and validating the International Development and Early Learning Assessment (IDELA). *International Journal of Educational Research*, *91*, 1–15. https://doi.org/10.1016/j.ijer.2018.06.007

Sthapak, S. (2025). Foundational Literacy and Numeracy: A Systematic Review in the Context of Primary Education in India. *Educational Research Review*, *13*(1), 55-67.S

World Bank Group. (2017). *World development report 2018: Learning to realize education's promise*. World Bank Publications. World Development Report 2018: Learning to Realize Education's Promise thapak.