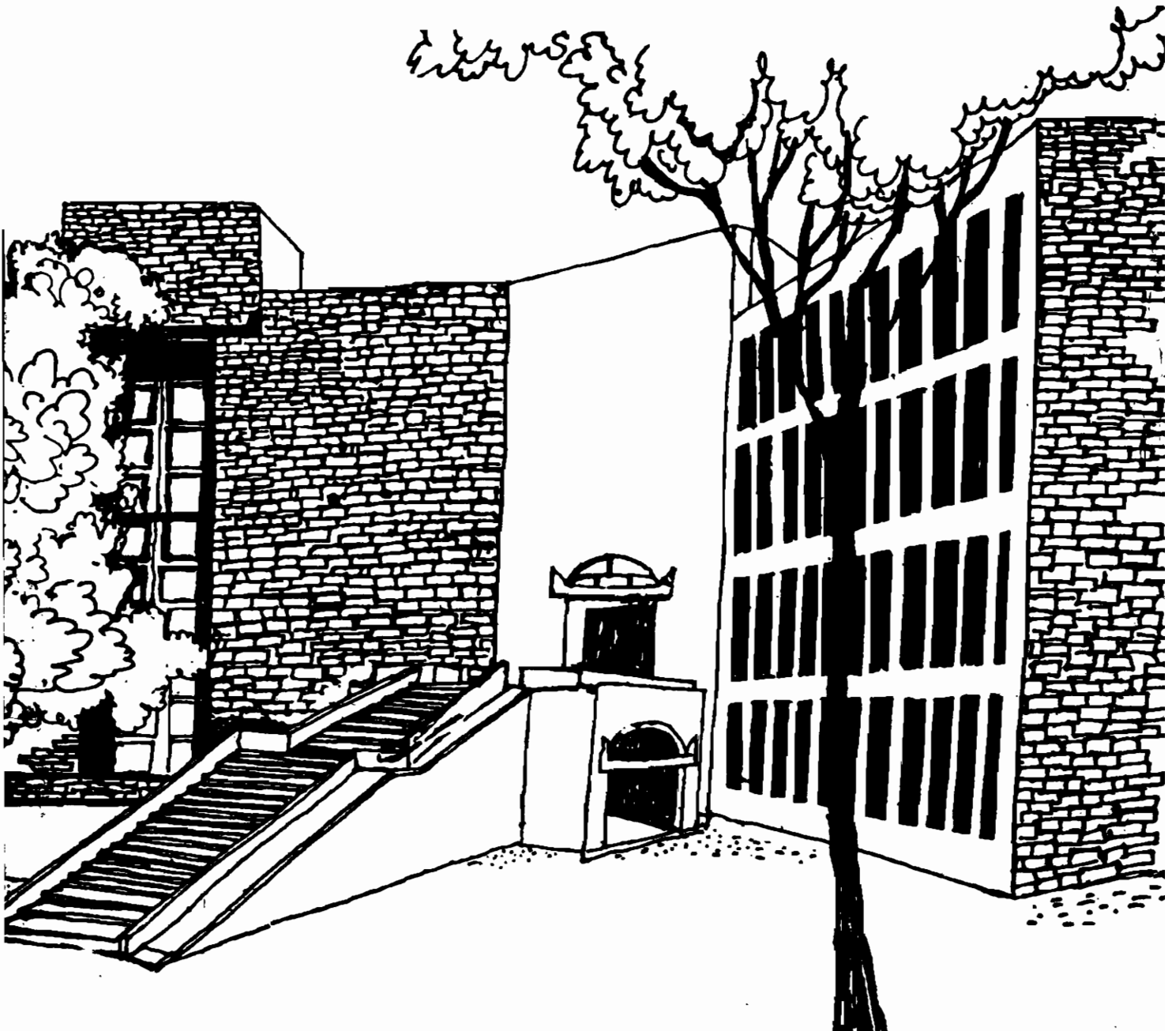




# Working Paper



**A SIMPLE TOOL FOR BENCHMARKING EARLY  
LITERACY SKILLS IN INDIAN LANGUAGES**

**By**

*Brij Kothari*

**W.P.No. 2001-07-03**

**July 2001**

*/1661*

**The main objective of the working paper series of the IIMA is to help faculty members to test out their research findings at the pre-publication stage.**

**INDIAN INSTITUTE OF MANAGEMENT  
AHMEDABAD-380 015  
INDIA**

IIMA WP-2001-07-03

<b>PURCHASED</b>
<b>APPROVAL</b>
<b>GRATE/EXCHANGE</b>
<b>PRICE</b>
<b>ACC NO. 250407</b>
<b>VIKRAM SARABHAI LIBRARY</b>
<b>1 W, AHMEDABAD.</b>

## **A Simple Tool for Benchmarking Early Literacy Skills in Indian Languages**

Brij Kothari and Ashok Joshi

### **Abstract**

India's literacy rate has been rising rapidly. Unfortunately we are in the dark about the literacy levels of the so-called literates. A simple, quick, and easy to administer tool for measuring literacy skills is described. The tool has been found to be useful in literacy and primary education contexts. Use of the tool on a sample population can throw light on the distribution of literates along a range of skill levels. Knowledge of this distribution has direct policy implications for national literacy initiatives.

## A Simple Tool for Benchmarking Early Literacy Skills in Indian Languages

Brij Kothari and Ashok Joshi

### Introduction

In percentage points, India has taken giant strides on the road to literacy. As highlighted in an analysis of the National Literacy Mission (NLM, 2000, p. 13), "India could be expected to touch the sustainable threshold level of 75 % sometime between 2005 and 2006. This would be far earlier than the previous expectation of around 2011." Table 1, based on the National Sample Survey Organisation's findings of decadal literacy growth (as reported in NLM, 2000), summarises the reasons for this optimism.

**Table 1: Decadal literacy growth in India (7+ age group)**

Year	Literacy Percentage	Decadal Growth	Non-literates (millions)
1951	18.3	--	--
1961	28.3	10.0	249.4
1971	34.4	6.1	283.03
<b>1981</b>	<b>43.5</b>	<b>9.1</b>	<b>305.31</b>
1991	52.5	8.7	328.88
1997	62	9.8	294.46
2001	--	--	268.42*

\* Extrapolation by NLM

While the decadal growth in the literacy rate for the 7+ age group has averaged 8.7 % since 1961, the actual number of non-literates is projected to be on a downward trend. The figures, although very encouraging, call for cautious optimism. The caution is primarily due to the fact that our understanding of the literacy levels of the millions of people whom we consider to be literate, is severely limited. The only method that is used in practice by the census machinery is to ask a household respondent information on all the members – whether they are literate or non-literate, thus, reducing the state of literacy to a black or white response.

Absolute non-literacy is a state of being but literacy itself lies on a continuum. Yet in reporting national statistics, literacy is reported as a static state, whether one can barely write one's name or one is an accomplished writer. What percentage of the so-called literate population is 'literate' at what level, is not a question policy has grappled with so far in the drive for higher literacy rates. Since literacy, most narrowly defined includes reading, writing, and numeracy skills, the question can be refined further to ask: What percentage of the literate population is, to coin a few terms, readerate, writerate, and numerate at what level? The answers to this question, it is obvious, have serious policy implications in terms of the effort and focused interventions that are still required on the road to literacy. To say the least, we cannot and should not be satisfied with mere milestones of an advancing literacy rate.

Hence, there is a need for tools that could simply and precisely determine the levels of the 3Rs. These tools could be used to bring some clarity on the fuzzy halo surrounding the literacy rate. They could also benchmark progress in a variety of literacy projects, programmes, and campaigns around the country, at the early, intermediate, and final stages of intervention. Tools that make comparisons on certain basic parameters possible within (at different times) and across districts are not easy to find despite the enormous need.

The present article describes one such tool in Gujarati language that was developed as part of a literacy experiment in Gujarat State. A version of it in Hindi was used later in the evaluation of primary education in Madhya Pradesh State (Kothari et al., 2000a) and in the state's Padhna-Badhna Andolan (literacy movement).

### **A tool for benchmarking early literacy skills<sup>1</sup>**

#### *Genesis*

The need for a tool to benchmark early literacy skills emerged out of a project for **strengthening the skills of neo- and semi-literate**. The project idea was evaluate the impact of Same Language Subtitling (SLS) on the reading ability of early literates. SLS refers to the idea of subtitling film-song-based entertainment on television, in the same language as the audio. Two experiments were conducted. In the first one, conducted in a government primary school, one group of children was exposed to SLS song programmes, the other group saw the videos without subtitling, and the third control group saw nothing (Kothari and Takeda, 2000; Kothari et al., 2001). In the second experiment, SLS was added to a popular programme of Gujarati film songs, Chitrageet, telecast weekly on Doordarshan Kendra, Ahmedabad (State TV for Gujarat). The impact of SLS on Chitrageet viewers' reading ability was compared with non-viewers (Kothari et al., 2000b). In both the experiments, **precise benchmarking of skills was required before and after intervention to be able to comment on the effectiveness of SLS.**

#### **Objectives**

Several objectives guided the development of the tool. Primarily, the tool needed to have the qualities given below:

- i) capacity to ~~of~~ discern small differences in levels among early literates so that incremental improvements resulting from interventions could be measured accurately;
- ii) ability to determine precisely the areas of weakness in reading and numeracy (e.g., this information is helpful to the testee for targeted effort at improvement);
- iii) possibility of reuse with or without simple modifications;
- iv) there should be no or minimum residual learning with reuse;
- v) application in all Indian languages that have a script; and
- vi) short, simple and quick administration process.

---

<sup>1</sup> Editor: if necessary, a sample tool in Gujarati (or Hindi) language may be included as an appendix.

The tool described presently, first developed in Gujarati and subsequently adapted for Hindi, professes the above qualities. While the discussion below is based on the tools developed in these two languages only, it applies with minor variations to all Indian languages based on the Devanagari script and easily extendable to South Indian languages.

## Tool

The alphabet in Gujarati and Hindi does not have letters as in English but, rather, syllables and a set of 11 vowels to modify the syllables. Unlike English, Indian languages are phonetic (consistent pronunciation), hence, the ability to recognise syllables and their modifications with vowels is a major concrete step in reading. In English, recognition of letters and their modification with vowels is an insufficient building block of reading since exceptions in pronunciation abound (e.g., put and but) and these exceptions are important to learn before one can read correctly.

The tool has six small sections. Sections 1-3 focus on reading ability, 4-5 on arithmetic, and 6 on writing. Recognition ability for syllables (Section 1), and syllable + vowel pairs and conjunct syllables (Section 2) are good ways to determine reading ability in Indian languages. In addition, simple sentence reading ability was measured in Section 3.

### Section 1

Meaningful two syllable words were created to cover all the syllables, except those that are rarely used (Table 1; 'S#' stands for Syllable # and 'S#S#' is a two-syllable word). Only simple syllables, without the use of vowels and/or conjunct syllables, were included in Section 1. An attempt was made to not repeat any syllable, unless absolutely necessary to form words with syllables not yet included. For the Gujarati tool developed, all the syllables were covered by S47S48 without repetition of syllables.

**Table 1: Simple two-syllable words**

S1S2	S3S4	S5S6	S7S8	S9S10
S11S12	S13S14	S15S16	S17S18	S19S20
S21S22	S23S24	S25S26	S27S28	S29S30
S31S32	S33S34	S35S36	S37S38	S39S40
S41S42	S43S44	S45S46	S47S48	S49S50

During administration, the reader was asked to read the two-syllable words in order. In case of inability to read any syllable, help was provided for that syllable only and the reader had to try and complete the word. Individual syllables not read correctly or not read at all were circled on a separate sheet. Thus, part credit for words was possible. This provided an accurate picture of syllable-level reading difficulties. Repeated syllables in Section 1, if any, were not considered for marking. Whether the reader read most words in a smooth flow or with breaks, was recorded. The time it took to read the complete set of simple two syllable words was noted.

## *Section 2*

Recognition ability for all syllables having been covered, the aim of Section 2 was to test the reading ability of vowel modifications and conjunct syllables. The necessary number of syllable + vowel pairs in meaningful words and words with conjunct syllables were created. Help with simple syllables was provided to the reader wherever necessary and inability to read simple syllables was not recorded since this was undertaken earlier in Section 1. Only inability to read vowel modifications and conjunct syllables was marked on a separate recording sheet, resulting in the identification of weak spots as in Section 1. Time taken to read the total set of words. The general flow of reading was recorded, this time, in three ways: smooth, broken, or with vocal combination of elements. In the latter, very early literates often voice each element before combining them. This was not necessary with simple syllables in Section 1 because there were no elements to combine. For the Gujarati tool developed, 24 words were required to cover all the vowel modifications and conjunct syllables.

## *Section 3*

Sections 1-2 primarily focused on the ability to recognise all the symbols that appear in a language, through meaningful words. Sentence reading ability is the subject of Section 3. Three short sentences, increasing in difficulty, had to be read. For each sentence the flow was noted in four options: i) vocal combination of elements, ii) broken, iii) smooth, and iv) could not read.

## *Section 4*

Section 4 ascertains familiarity with numbers. The subject is asked to read single digit numbers from zero to 9, written in a random order. Numbers read wrongly are circled on a separate marking sheet. This is followed by a reading of nine two-digit numbers, each chosen randomly from 11-19, 21-29, 31-39, ..., 91-99. The ordering in the tool is random. Numbers not read directly but as a run-up from a lower number, are underlined. This problem is commonly noted in early learners.

Oral administration of questions ends with Section 4. Hereafter, the respondent solves basic arithmetic problems in Section 5.

## *Section 5*

The following types of problems are given:

Addition:

- i) two single-digit numbers without carry over;
- ii) two single-digit numbers with carry over;
- iii) one single-digit number with zero; and
- iv) two three-digit numbers with carry over.



**Subtraction:**

- i) smaller single digit from larger single digit;
- ii) of zero from a single digit; and
- iii) smaller three-digit from larger three-digit, involving carry forward.

**Multiplication:**

- i) single digit with single digit;
- ii) two-digit with single digit; and
- iii) two-digit with two-digit numbers.

**Division:**

- i) larger single-digit with smaller single-digit; and
- ii) three-digit with single-digit.

The arithmetic problems are posed at the neo- to semi-literate level, from simple to gradually more complex problems. The guiding principle in choosing the problems has been to uncover specific areas of weakness rather than just give a score.

**Section 6**

Respondents are asked to write 3-4 lines from any song or poem. If no song occurs to them, then a popular song is suggested. If they cannot write even the dictated lines, they are simply asked to write their name. (It is well-known that many people have learnt to read and write their name as patterns but are unable to read or write other words constructed from the syllables that make up their name.) A score is given based on spelling mistakes and readability in the writing sample.

**Use of tool**

The tool is a two-page test that can be administered to someone in less than 15 minutes per person. For reading and numeracy the responses precisely reveal the strengths and weakness in an early literate. Admittedly, the tool's weak spot is its ability to evaluate writing skills. This, however, is unavoidable in a tool that also aims for simplicity and quick administration.

Results from the two experiments to which the tool owes its genesis, confirm that the tool is short, simple, and quick to administer. It is capable of discerning small improvements in reading ability. For instance, the tool was able to measure incremental improvements in reading ability in the groups that saw film songs with Same Language Subtitling (Kothari et al., 2001; Kothari et al. 2000b). Benchmarking was done before and after intervention with the exact same tool, without any modifications. In the control group, no residual learning was observed merely with the tool's reuse (Kothari et al., 2000b). In Madhya Pradesh State, a Hindi version of the tool was used to measure reading ability differences between children studying in different systems of primary education. The tool is easily translatable into other Indian languages and permits reuse with the same

subjects with just a modification of the words and sentences constructed in Sections 1-3 for reading and with fresh arithmetic problems posed in Sections 4-5.

### **Policy implications**

One potential use of the tool is to answer larger questions hidden under the literacy rate and discussed at the beginning of the article. What percentage of the literate population is readerate, numerate, and writerate, and at what level? Census operations can hardly afford to use even a simple tool such as this on a large scale. The 15 minutes of administration per person makes the tool prohibitive. However, census operations or others can use the tool on a very small scale and based on the sample, project a fairly accurate picture of literacy levels on a larger scale. For example, the tool can be administered to every one in selected villages where the literacy rate is known from the census. Analysing the results of school-going children separately, one can build a scale of averages at different grade levels for the 3Rs. One can then arrive at percentage figures in the non-school-going population that are literate at different grade levels. Another possibility is to look at the distribution of scores based on syllables read correctly in Section 1 and make judgements on what percentage of the literate population can be classified as non-, neo-, semi-, and completely literate. Finally, one could work toward an approach, whereby, given a literacy rate of any geographical locality, one could broadly translate this into projected percentages for neo-, semi- and total literacy. Literacy levels come in all shades of gray. Literacy rates, including the 'sustainable' threshold of 75 % we are slated to reach by 2005, may hide the fact that far too many in India lie in the darker shades of gray.

Manuscript length: 2318 words.

## References

- Kothari, Brij and Joe Takeda, 2000. "Same Language Subtitling for Literacy: Small Change for Colossal Gains." In *Information and Communication Technology in Development*, Bhatnagar, Subhash C. and Robert Schware (editors), Pp: 176 – 186, New Delhi: Sage Publications.
- Kothari, Brij, Joe Takeda, Ashok Joshi, and Avinash Pandey (2001). Same Language Subtitling: A Butterfly for Literacy? *International Journal of Lifelong Education*, forthcoming.
- Kothari, Brij, P.G. Vijaya Sherry Chand, and Rajeev Sharma November, 2000a. Review of Primary Education Packages in Madhya Pradesh. Indian Institute of Management, Ahmedabad, unpublished report, IIM Ahmedabad.
- Kothari, Brij, Ashok Joshi, Avinash Pandey, Pawan Bharadwaj, Darshana Patel, Kalpesh Patel, Hemlata Jadwani, Nirali Joshi, and S.R. Joshi, 2000b. Same Language Subtitling on Chitrageet: Post Literacy Through Television in Gujarat, unpublished report, IIM Ahmedabad.
- National Literacy Mission, 2000, Literacy Rates: An Analysis Based on NSSO Survey 1998, New Delhi. Also available at: <http://www.infoindia.net/nlm/nssso.htm>, checked on Nov. 18, 2000.

### **Authors' Address and Contact Details**

Brij Kothari is an Assistant Professor at the Indian Institute of Management, Ahmedabad (IIMA), in the Ravi J. Matthai Centre for Educational Innovation. Ashok Joshi is a research assistant at IIMA. Corresponding author: Brij Kothari, Wing 14, Indian Institute of Management, Vastrapur, Ahmedabad - 380015, Gujarat, India. E-mail: [brij@iimahd.ernet.in](mailto:brij@iimahd.ernet.in); 91-79-6307241; Fax: 91-79-6306896.

**PURCHASED**

**APPROVAL**

**DATE/INITIALS**

**PRICE**

**ACC NO.**

**VIRAM SARABHAI LIBRARY**

**L. I. N., AHMEDABAD.**