

Can India's "literate" read?

Brij Kothari · Tathagata Bandyopadhyay

Published online: 4 February 2011
© Springer Science+Business Media B.V. 2011

Abstract This paper takes a close look at India's literacy rate by exploring whether the officially "literate" can read and at what level. In a large sample, aged 7+, drawn from four Hindi-speaking states, two methods were used to measure literacy. One was the standard Census Method (CM) which relies on self-reporting and the other was a Reading Method (RM) which required the same individuals to actually read a simple text at grade 2 level. The findings revealed a substantial difference between the reading literacy rates obtained by CM and RM. CM over-reported RM by 16%. The overestimation was higher for males. Decoding skills were found to erode in most cases after completion of primary schooling, assuming no further education. A minimum grade 8–9 education was required for decoding skills to not deteriorate after schooling.

Keywords India · Literacy · Reading · Decoding · Skill · Census · Measurement · Method

Résumé Les « lettrés » de l'Inde savent-ils lire ? – Les auteurs ont analysé les taux d'alphabétisation de l'Inde, en vérifiant si les personnes officiellement « lettrées » savent réellement lire, et dans l'affirmative à quel niveau. Ils ont utilisé deux méthodes pour mesurer la littératie auprès d'un vaste échantillon de personnes âgées de plus de sept ans, sélectionnées dans les États de langue hindi. L'une est la méthode standard de recensement qui repose sur l'auto-évaluation, la seconde une

B. Kothari (✉) · T. Bandyopadhyay
Indian Institute of Management, Vastrapur, Ahmedabad 380 015, Gujarat, India
e-mail: brij@iimahd.ernet.in

méthode de lecture qui requiert les mêmes participants de lire effectivement un texte simple du niveau de deuxième année de l'enseignement primaire. Les résultats révèlent une différence importante entre les taux de compétence de lecture obtenus par les deux méthodes. La première dépasse en effet de 16 pour cent les résultats de la seconde. La surestimation est plus élevée chez les hommes. Les auteurs constatent que les compétences de décodage disparaissent dans la majorité des cas après l'achèvement de l'enseignement primaire, à moins que les apprenants poursuivent leur apprentissage. Un enseignement minimum jusqu'à la huitième ou neuvième année serait nécessaire pour que les compétences de décodage ne se dégradent pas après la scolarisation.

Zusammenfassung Können die „Lesekundigen“ Indiens lesen? – In dieser Arbeit unterziehen wir die indische Alphabetisierungsrate einem kritischen Blick und untersuchen, ob Menschen, die offiziell als „alphabetisiert“ gelten, lesen können, und wenn ja, auf welchem Niveau. Anhand einer großen Stichprobe – Altersgruppe ab 7 Jahre aufwärts – aus vier Bundesstaaten, in denen Hindi gesprochen wird, haben wir die Alphabetisierungsrate mithilfe zweier Methoden gemessen. Die erste war die übliche Erhebung (Census Method = CM) auf der Basis von Selbstauskünften. Bei der anderen wurden dieselben Personen gebeten, einen einfachen Text auf dem Schwierigkeitsniveau der zweiten Klasse zu lesen (Reading Method = RM). Es stellte sich heraus, dass sich die Raten für Lesefertigkeiten erheblich unterschieden, je nachdem, ob diese nach der CM oder der RM erhoben wurden. Die CM-Rate lag um 16% höher als die RM-Rate. Männliche Befragte wurden stärker überschätzt als weibliche. Wenn der Bildungsweg nicht fortgesetzt wurde, verlor sich die Fertigkeit, Schrift zu entziffern (Dekodierung), in den meisten Fällen nach Beendigung der Primarschule wieder. Erst nach mindestens 8 bis 9 Schuljahren blieb die Dekodierungsfertigkeit nach Beendigung der Schule erhalten.

Resumen ¿Saben leer las "personas alfabetizadas" de la India? – Con este trabajo, los autores estudian muy de cerca la cuota de alfabetización de la India, explorando si las personas oficialmente denominadas "alfabetizadas" saben leer, y a qué nivel. En una amplia muestra tomada con personas de 7 años y mayores, en cuatro estados de habla hindi, se utilizaron dos métodos para medir la alfabetización: el CM (Census Method), o sea el método de censo estándar basado en la autoevaluación, y el RM (Reading Method), el método de lectura que requería de las mismas personas que efectivamente leyeran un texto simple del nivel de 2.^{do} grado. Los resultados revelaron una diferencia sustancial entre CM y RM en cuanto a las cuotas de alfabetización obtenidas, puesto que el informe basado en CM presentaba un excedente del 16%. La sobreestimación era mayor en cuanto a los varones. Se comprobó que las competencias de decodificación se iban perdiendo en la mayoría de los casos una vez completada la enseñanza primaria, partiendo de la base de que no había educación subsiguiente. Se necesita como mínimo una educación de 8.^o o 9.^o grado para que las competencias de decodificación no se deterioren después del aprendizaje escolar.

Резюме Умеют ли читать «грамотные» в Индии? – В данной статье тщательно исследуется уровень грамотности в Индии, и рассматривается вопрос о том, умеют ли читать официально признанные "грамотные" и на каком уровне. На примере большой выборки в четырех штатах, где говорят на языке хинди, использовались два метода для определения уровня грамотности в возрастной группе 7+. Один из методов – это стандартный метод сбора данных (СМ), который основывается на самооценке, а другой метод - метод чтения (РМ), в котором предлагается тем же участникам на самом деле прочитать простой текст уровня второго класса. Согласно этим двум методам, полученные данные выявляют существенную разницу между показателями уровня грамотности относительно чтения. Показатели первого метода (СМ) на 16% превосходили показатели второго (РМ). Завышенная оценка преобладала среди мальчиков. Обнаружилось, что в большинстве случаев способность декодировать информацию после окончания начальной школы при отсутствии дальнейшего обучения постепенно уменьшалась. 8-9 классов – это минимальный срок обучения, необходимый для того, чтобы способность декодировать информацию сохранялась и после окончания школы.

Introduction

If there is one fundamental skill that we expect all students to acquire through education, and early enough, it is reading. Reading literacy is the building block which mediates almost all learning and teaching in modern educational systems. It is, arguably, the most important skill for a student to acquire in order to survive and thrive in schooling and life. Youth and adults who have not acquired basic reading literacy as children almost certainly find themselves at a disadvantage. Their aspirations are clipped. Their self-esteem is tarnished. Their functioning in everyday life is compromised.

Reading literacy is so elemental to an individual's and society's progress that the International Association for the Evaluation of Educational Achievement (IEA), starting in 2001, has been conducting a multi-national comparative study every 5 years, the Progress in International Reading Literacy Study (PIRLS). Forty-five countries participated in PIRLS 2006 and 55 are expected to participate in PIRLS 2011. 2011 is also the year when the next Indian census will be conducted. It will include a measure of India's literacy rate, a number which will provide an important indicator of progress – relative to India a decade ago. But, unlike PIRLS, it will say little about what it means to be counted among the "literate" in India.

Literacy in the Indian census

In theory, the Indian census follows UNESCO's definition of literacy, which expects at minimum "an ability to read and write, with understanding, a short simple

statement relevant to one's everyday life". If the census were to do justice to this definition, even with a focus on reading only, it would have to assess in some form or other both decoding ability and comprehension of a simple written statement taken from everyday life. Clearly, that would pose a daunting task in a population of over one billion people.

In practice, it is therefore understandable that the Indian census operation checks for neither decoding ability nor understanding. But what is troubling is that it does not employ a proxy method which could, with a reasonable degree of simplicity and accuracy, determine whether a person can decode and comprehend a simple text taken from his/her life. For instance, asking a person if he/she can read the bus board (or a letter) would provide a more accurate measure of literacy as defined by UNESCO than asking whether he/she is literate. Implicit in the question is a test of decoding and comprehension. At present, the census relies on more than 2.2 million enumerators, mostly teachers, to visit every household in India. The standard procedure simply asks the head of the household (or a senior household respondent) to report for each member whether the individual is "literate" or "illiterate," leaving the interpretation of what the terms mean to the respondent. Besides being at odds with the definition of literacy which it claims to measure, this approach "suggests a dichotomy between 'literate' and 'illiterate', although, inarguably, literacy lies on a continuum of proficiency or competence" (UNESCO 2008, p. 62). It also tends to inflate the literacy rate since most people would want to report themselves as literate.

The literacy rate in India is thus a perceptual figure, merely giving the percentage of people perceived to be literate by the head of the household. At what level of literacy ability do people start being reported as literate? As soon as they can write their name, or draw it even without alphabetic knowledge? After they have acquired partial or full alphabet decoding ability? Or are people considered literate in their families only after they can at least decode simple texts they come across in their everyday lives? In other words, is a basic ability to decode necessary for someone to be counted as literate in the census? Ideally, the bar to be counted as literate would be set minimally as an ability to decode a basic text with understanding. But in the context of a census in a developing country, it may be acceptable to further relax it to a demonstration of decoding skills at grade 2 level.

In the community of nations, every nation would like to report a high literacy rate. Thus, the shortest path to achieving higher literacy rates is to employ self-reporting as the method of measurement. This results in a major problem. National literacy efforts then tend to go for the "low-hanging fruit" by imparting only as much instruction as is necessary before someone refers to himself/herself as being literate. For an illiterate person, the transition from "illiteracy" to "literacy", in self-perception, happens very early on in the process of reading skill acquisition, long before someone can even decode the simplest imaginable text. A literacy policy focused on increasing the self-reported literacy rate, but not equating literacy with an ability to decode, much less comprehend, creates conditions in which a large number of so-called "literate" remains or returns to being lifelong non-readers. India, unfortunately, is home to millions of lifelong literates who are also, somewhat confusingly, lifelong non-readers (who cannot read and not because they do not want to).

According to the 2001 Census, India's literacy rate for the population aged 7 and above was 65.4%. Can 65.4% of the 7+ population read—and by that we simply mean decode, with or without understanding, a simple text that a second grader would be expected to read? Indeed, that is a minimum expectation of anyone who is considered to be literate, regardless of one's definition of literacy. Most definitions of literacy expect far more than an ability to decode simple texts, but no definition excludes the ability to decode at a very basic level (see UNESCO (2008, pp. 236–243) for a comprehensive list of literacy definitions used in different countries).

For the purpose of this article, we define a person as Reading-Literate (RL) if he/she can decode a grade 2 level text, with the requirement that every word is sounded out as a whole word and not put together bit by bit. We do not require comprehension, although we posit that an ability to decode whole words is a precursor to reading comprehension. If a person could decode a simple grade 2 level text, we leniently considered that as a marker of reading literacy, recognising that this is clearly insufficient in an information society. If the person can decode but needs to sound out word components first before putting the words together, he/she will be called Early-Reading-Literate (ERL). It would be interesting to see how the rate of RLs would compare to the self-reported rate of Census-Literates (CLs). Another question would be: Among CLs, what is the proportion of RLs, ERLs and Non-Literates (NLs)?

Clearly, the literacy rate is inextricably tied to the method used to measure it. Therefore, all we can say for sure is that, in 2001, 65.4% of the 7+ population in India was reported to be "literate" by household respondents. This figure is projected to cross 80% by the next census in 2011 (Chand 2007). The powerful, if misleading, perception it creates is that eight out of every 10 Indians aged above 7 years are readers. In the Caribbean, Jennings (2000) argues that an increasing or high literacy rate can actually mask the problem of low literacy achievement. Many countries are moving beyond the measurement of literacy as a dichotomy.

International comparative measurements of literacy

Internationally, there has been a number of appreciable efforts to measure the quality of literacy in a manner that allows comparison across nations. The International Adult Literacy Survey (IALS) was the path-breaker, conceptualised as a tool to measure functional literacy in three domains – prose literacy, document literacy and quantitative literacy. It was first conducted in 1994 in seven OECD countries. Functional literacy was defined as "the ability to understand and employ printed information in daily activities at home, at work and the community – to achieve one's goals and to develop one's knowledge and potential." The two core innovations of IALS were: (a) the "measurement of *varying degrees* of literacy skill" (not simply a measurement of "literate" versus "illiterate") and (b) the implementation of this tool across nations and, importantly, different languages, within a comparative framework (Darcovich 2000). The IALS instrument of 1994 was repeated in 1996 with five more countries, and in 1998 with eight more, or a

total of 20 countries. After that, IALS went into hibernation and reappeared in 2003, as the Adult Literacy and Lifeskills Survey (ALL).

The hiatus is indicative of concerns with the construction and application of IALS across countries. IALS prioritised the processing of data-based information over experiential and contextualised information (Gomez 2000). Hamilton and Barton (2000) argued that IALS “provides a partial picture of literacy but claims definitively to represent all of literacy.” Despite the limitations, IALS inarguably gave impetus to significantly more complex approaches to a determination of literacy levels than a simple categorisation of people as “literate” or “illiterate” based on self-reports.

By the time IALS was recast as ALL in 2003, three major international efforts had gathered steam, with greater precision in definition, measurement and customisation to context. The Programme for International Student Assessment (PISA) was launched in 2000 with a sharper focus on 15-year-olds’ mathematics and science reading literacy. Starting with 43 countries, PISA has been repeated every 3 years. The latest PISA 2009 featured an impressive 65 countries. The other two international comparative initiatives were: (i) the Progress in International Reading Literacy Study (PIRLS) among fourth graders, starting with 35 countries in 2001 and 55 countries expected in 2011 (Mullis et al. 2007), and (ii) Trends in International Mathematics and Science Study (TIMSS), first conducted in 1995 among third, fourth, seventh and eighth graders from 40 countries but more focused now on fourth and eighth graders. Sixty-four countries are expected to participate in TIMSS, 2011. ALL cast a wider net to include 16–65-year-olds. As a result, it could only muster participation by seven countries in its first and up to now only round in 2003.

PISA, PIRLS and TIMSS are gaining international acceptance (Naumann 2005). However, this acceptance is mainly among countries reporting a high literacy rate. The average basic literacy rates reported for the countries participating in PISA, PIRLS and TIMSS were, respectively, 96.7, 94.5 and 92.5%. In PIRLS, which measured reading literacy among fourth graders and is directly relevant to the present study, the few participating countries with a literacy rate below 90% were: Libya (84.2%), Saudi Arabia (82.9%), South Africa (82.4%), Iran (82.4%), Oman (81.4%), Botswana (81.2%), Honduras (80%), Egypt (71.4%) and Morocco (52.3%). Most low-literacy countries, it would seem, are not ready yet for their quality of literacy to be measured independently.

To our knowledge, there are no well-researched estimates of reading levels in the out-of-school population in India. Pratham’s Annual Status of Education Report (ASER) is the first national attempt, starting in 2005, at understanding reading levels among school children. Five editions of ASER are now out and in the latest round of this laudable effort, one of the key findings is that 47.2% of rural children in fifth grade could not read a simple story pitched at grade 2 level (ASER 2009).

Our research takes a close look at actual reading literacy ability among CLs, not just children in school. There is no dearth of informal accounts of low reading literacy skills among many CLs, but, to our knowledge there are no systematic studies on reading literacy levels among all those considered to be officially “literate” in India.

A Bangladesh study looked at the discrepancy between literacy rates resulting from self-reporting and those resulting from actual testing (Nath 2007). Self-

reporting over-reported the literacy rate by 5.6% when compared to testing. Over-reporting was found to be higher among females and in rural areas. Over-reporting in Bangladesh, it would seem, was not as high as in some of the other countries cited, such as Laos (29%) and Bhutan (38%). But unlike PIRLS, TIMSS and PISA, the use of different methodologies precludes cross-country comparisons. It also highlights the need for an international comparative study to assess the quality of literacy, especially in low-literacy countries.

Methodology

Sampling

In September 2002, we commissioned Nielsen's ORG-Centre for Social Research (ORG-CSR) to independently collect data on the literacy status of entire households, in four Hindi-speaking states, namely Rajasthan, Uttar Pradesh, Madhya Pradesh and Bihar. The data were originally collected by ORG-CSR as part of a baseline to select experimental and control groups of weak readers for a study that would eventually go on to measure the impact of Same Language Subtitling (SLS) on reading skills at different points in time.¹ We used the baseline data from the SLS study for the analysis presented in this paper. The baseline data had both the elements for all the sampled households: the reported literacy for all household members and their reading ability as measured by their ability to decode a simple paragraph at grade 2 level. The data, therefore, are well-suited to explore the extent of early reading literacy among the officially "literate".

ORG-CSR stratified all the districts in the four states into high, medium and low literacy districts, based on the rural literacy rate reported in the 2001 census.² One medium literacy district was selected randomly from each of the states. They happened to be Dausa (Rajasthan), Lucknow (Uttar Pradesh), Umariya (Madhya Pradesh) and Muzzafarpur (Bihar). One block was selected randomly from each of the selected districts. For each block, a list of large (population > 5,000), medium (population between 1,000 and 5,000) and small (population < 1,000) villages was prepared. One large, two medium and two small villages were selected randomly from every selected block. Eventually, 20 villages were covered in the survey, five from each state. Finally, a representative sample of households was chosen from the socio-economic clusters in the selected villages.

Measurement approach

ORG-CSR undertook two rounds of measurement in their baseline. The first round mimicked the Census Method (CM) by collecting data only from the head of the household, or a suitable representative, on all the household members. This included

¹ For more on the SLS work, see Kothari et al. (2004).

² Administratively, a state is divided into districts, blocks, *panchayats* (village councils) and finally villages.

socio-demographic data on all household members, and for those aged 7 years and above, whether the person was thought to be “literate” or “illiterate”. All sampled households in the village were surveyed in this manner, before a second round of visits to the same village.

In the second round, ORG-CSR followed the Reading Method (RM). The same households were revisited and every member present was asked to read a simple paragraph at grade 2 level, printed in reasonably large-sized letters (see Appendix 1). To minimise the possibility of rote repetition within a household, three versions of the same paragraph were administered, all slightly different from each other but at the same level of simplicity. A person demonstrating an ability to decode smoothly, at any pace, was marked as reading literate (RL). Anyone who could decode a bit, but took recourse to sounding out syllables before uttering the whole word, or could not read all the words correctly, or read at a laboured and broken pace, was marked as reading early-literate (REL). Finally, those who were unable to decode at all were classified as reading non-literate (RNL).

Data collection

In the first round, ORG-CSR collected data on all the individuals, in 3,179 households, including 17,782 individuals aged 7 and above (see Table 1).

The 7+ sample included 47.1% females, probably due to the skewed sex ratio in the population. In the second round, ORG-CSR was able to administer the reading test to a greater proportion of females (54.2%), because they were more likely to be at home. This pattern was consistent and comparable across all states. Overall, RM could be administered to 64.5% of the 7+ individuals in the sampled households (present, therefore tested sample); others were not present at the time of the visit (absent, therefore untested sample). Between the four states, there was considerable variation in the proportion of household members who were present (Table 2).

Accounting for the absent sample that could not be tested

The literacy rate (CM), as reported by heads of household for those present at home at the time of visit was 66.3% (Table 3). This contrasts sharply with the 73.1% literacy rate (CM) reported for those not present at home at the time of visit. Clearly, the absent sample had a higher proportion of reported literates than the present sample, explained by the fact that literate people were less likely to be found at

Table 1 Sample size by state

State	Households	Percent	Individuals (7+)	Male (%)	Female (%)
Rajasthan	764	25.9	4,605	53.1	46.9
Uttar Pradesh	630	20.5	3,639	53.0	47.0
Madhya Pradesh	1,194	34.6	6,144	52.3	47.7
Bihar	591	19.1	3,394	53.5	46.5
Total	3,179	100.0	17,782	52.9	47.1

Table 2 Percentage tested

	Sampled	Present	Present %
Rajasthan	4,605	3,096	67.2
Uttar Pradesh	3,639	2,621	72.0
Madhya Pradesh	6,144	3,252	52.9
Bihar	3,394	2,493	73.5
All States	17,782	11,462	64.5

Table 3 Literacy rate, census method in present versus absent samples

Literacy rate (CM)	Present sample % (n)	Absent sample % (n)	Total % (n)
Total	66.3 (11,462)	73.1 (6,320)	68.7 (17,782)
Male	78.5 (5,248)	82.7 (4,156)	80.4 (9,404)
Female	56.0 (6,214)	54.7 (2,164)	55.7 (8,378)

home during the day. For males, there were 82.7% reported literates in the absent sample as compared to 78.5% reported literates in the present sample. The reported literacy rate among females was comparable in the present and absent samples and even slightly higher in the present sample.

It was therefore not justifiable to draw conclusions on literacy achievement based on a testing of the present sample alone, nor could we impute patterns found in the present sample to the absent sample. In fact, if we had done that, this would have painted a gloomier picture of literacy achievement than what we have found and reported in this article.

With the goal of assigning RM-based reading levels to individuals who were absent and therefore could not be tested, we used the following logic. First, anyone reported as illiterate by CM was marked as such by RM. This is justifiable because, in the sample which was present, we found that 97% of those reported as illiterate by CM were also found to be completely illiterate when tested, i.e., by RM. Next, anyone who had completed grade 8 or above was considered to be a good reader, even though, in the tested sample which had completed grade 8 or above, we found only 50.7% to be good readers, 47.4% early readers and 2% non-readers. This latter step could only enhance the literacy profile of the untested sample by according a "good reader" status to a significant proportion of early readers. These two steps essentially took care of 66% of the absent or untested sample. It allowed us to confidently determine the reading literacy level – RL, REL, or RNL – of 88% of the total sample.

As a final step, we developed a multinomial logit regression model based on 88% of the total sample for which we could be sure of the literacy level. We then used the fitted regression model to impute the missing reading literacy levels of the remaining 12% of the total of the absent sample. The dependent variable was the reading literacy level as determined by the test: RL, REL or RNL. The model was significant at $p < 0.000$ (pseudo $R^2 = 0.75$, Cox and Snell). The model predicted 83.8% correct for RL, 75.6% correct for REL and 90.1% correct for RNL.

The factors found to be significant at $p < 0.05$ were: in-school or out-of-school, reported as literate/illiterate by CM, sex, head of household reported as literate/illiterate by CM, spouse of head reported as literate/illiterate by CM, availability of electricity, caste, TV ownership (none, colour or b&w), head of household's reading literacy level as determined by the test and spouse of head's reading literacy level as determined by RM. The covariates found to be significant at $p < 0.05$ were: grade completed, age, grade completed by head of household, grade completed by spouse of head and size of land.

We used the fitted model to find the probabilities of absent individuals being RL, REL, or RNL in the absent/untested sample. Every individual was then assigned the reading literacy level for which the probability of being RL, REL or RNL was the highest.

Results

Reading literacy: census method versus reading method

As already reported in Table 3, the CM applied to a sample of 17,782 people indicated a literacy rate of 68.7%. Female literacy (CM) was found to be 55.7% and male literacy (CM) 80.4%. This is what the national census operation would have found in the same sample and thus offers a good point of comparison for the reading literacy levels we actually found after testing (Table 4).

A sample in the Hindi belt which, when tested for basic reading ability, the census would have found to be 68.7% literate, was found at best to be 52.6% reading-literate, even if one combined the RL and REL groups and considered them to be reading-literate. At the lowest possible bar for reading literacy, such as an ability to read something, anything, from a paragraph at grade 2 level, only a little over half the 7+ population proved able to read. The census method could be said to overestimate the literacy rate by at least 16.1%. If the definition of literacy is tightened to a point where a person is expected to demonstrate a minimum reading

Table 4 Literacy levels found after testing (RM)

	Present/tested sample only		Present/ tested + absent/ untested sample for which we are near-certain		Entire sample after regression imputation to those absent/untested		Best case reading literacy scenario ^a	
	n	%	n	%	n	%	n	%
Literate (RL)	1,716	15.0	4,245	27.0	4,528	25.8	4,536	25.9
Early-literate (REL)	4,471	39.0	4,471	28.5	4,703	26.8	4,848	27.7
Non-literate (RNL)	5,275	46.0	6,975	44.5	8,295	47.4	8,142	46.4
Total	11,462	100.0	15,691	100.0	17,526	100.0	17,526	100.0

^a Accounts for possible error from mal-attribution of RL to REL/RNL groups and of REL to RNL group

Table 5 Literacy rate by different definitions

Sex	Literacy rate, census (%)	Reading literacy rate at basic alphabetic familiarity or better (%)	Reading literacy rate at ability to read grade 2 text or better (%)	Overestimation (%)	
				At alphabetic familiarity	At grade 2 ability
Male	80.4	63.4	36.1	17.0	44.3
Female	55.7	40.8	14.4	14.9	41.3
Total	68.7	52.6	25.8	16.1	42.9

ability at grade 2 level – not an unreasonable expectation – the reading literacy rate drops to 25.8% (Table 5). The CM, in that case, could be said to overestimate the literacy by an astounding 42.9%.

At a minimum expectation of basic alphabetic familiarity, the census overestimates the literacy rate for males by 17.0% and for females by 14.9%. This pattern of relatively greater overestimation for males holds true even for second grade reading ability. This is possibly because of the higher social expectation for males to be reported as literate.

In our sample which was found to be 68.7% literate by CM, a sample that is not too different from the national literacy rate of 65.4% in the actual census conducted a year before data collection for this study, it is instructive to look at the proportion of RL, REL and RNL only among the officially "literate". We found only 37.7% of the census literates to actually be able to read at grade 2 level and almost one out of every four "CM literates" could not demonstrate any reading ability (Table 6). Thus, 62.3% of the "CM literates" are in need of reading skill improvement before they can read simple texts comfortably. While one cannot extrapolate this nationally, especially given the fact that literacy skills tend to be more fragile in Hindi states, this begs the question even in high literacy states: How reading-literate are the census literates?

The 2001 Census points to a male–female differential of 21.7% in the literacy rate. Unfortunately, this gap is further widened with an understanding that only 26.0% female "CM literates" can read at grade 2 level as compared to 45.0% for males.

Having found 26.8% of the total sample to be REL, we looked at the variation in this group and determined the proportion of RELs who were likely to experience reading skill loss. Abadzi (2003) points out that learners need to acquire a minimum reading speed of 1 to 1.5 s/word, at 95% accuracy, to become lifelong readers. The paragraph used in our reading test had 37 words, setting up an expectation that it should have been read within 56 s with no more than two errors. We relaxed this to

Table 6 Census method literates who can or cannot read at grade 2 level

Sex	Reading literate (%)	Reading early-literate (%)	Reading non-literate (%)
Male	45.0	33.5	21.6
Female	26.0	46.3	27.7
Total	37.7	38.4	23.9

a maximum of 60 s and four errors and found that 36% REL qualified. Thus, 64% of the sample we found to be REL have a high probability of losing their already weak reading skills further, unless they are able to reach higher levels of automaticity in their reading. As a practical question, one might ask at what minimum grade level do school children subjected to the present quality of education in most rural schools reach a state of lifelong and irreversible reading ability at a minimum of at least grade 2 level?

Literacy and schooling

Anyone who has gone through five years of primary schooling is expected to become a lifelong reader, i.e., to be able to read a paragraph at grade 2 level. We looked at this issue separately for active students at the time of data collection and for non-students, including children, youth and adults (Appendix 2). The in-school and out-of-school sets were analysed separately because of the possibility of reading skill loss, over time, after a person has completed schooling. Figures 1 and 2 plot the proportion of good readers (RLs), weak readers (RELs) and non-readers (RNLs) by grade completed, among students and non-students. While it is clear that the completely non-reading group shrinks rather rapidly with increasing education, we also see, rather surprisingly, a large proportion of weak-readers all the way up to grade 10.

Comparing students and non-students, we find:

- (1) **Students:** 90% of first-grade completers were reported as literate by CM. Yet among those who had completed first grade, only 0.6% were found to be good readers and, generously, only 27% could be thought of as able to read at least a bit. Educationalists and policy-makers often argue that if we could only get everyone into primary school, the literacy rate would be near 100%. That is true; 90% would be reported as literate simply because they spent a year in school, not because they can read.

Non-students: 80.2% of first-grade completers were reported as literate by CM.

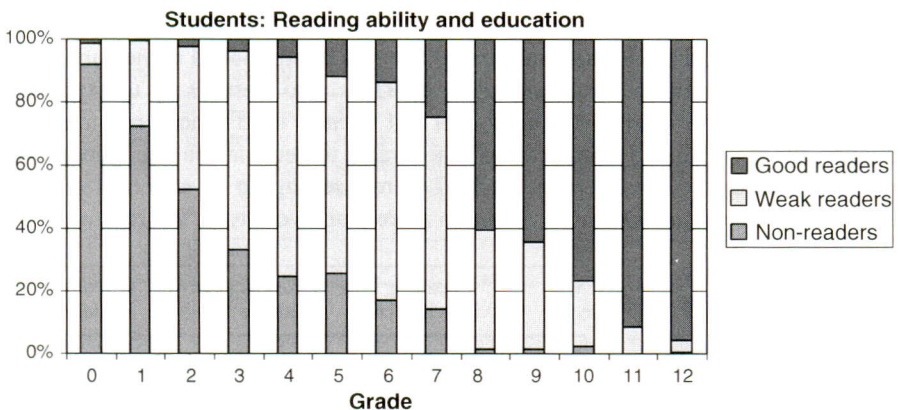


Fig. 1 Students, reading ability by education

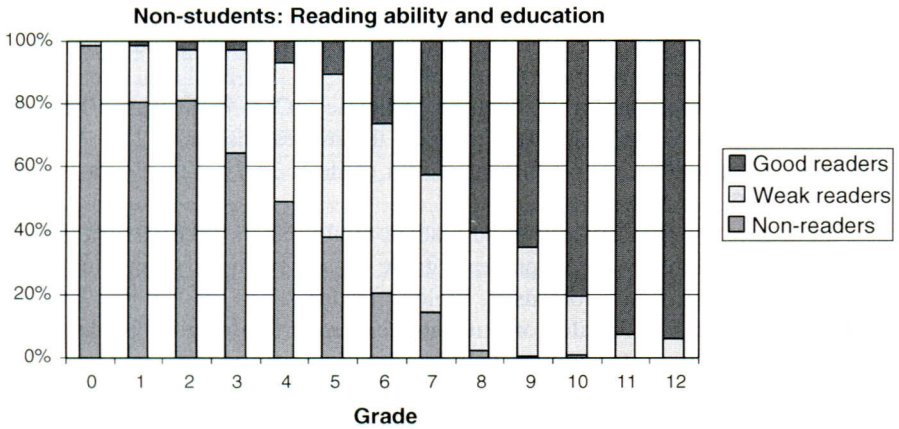


Fig. 2 Non-students, reading ability by education

Although this is a drop from 90% for students who have completed first grade, it is still high considering that 80.5% could not read at all. This points to the fact that once people are reported as literate in the census, no matter what their actual reading ability may be in the future they will be reported as being literate for the rest of their lives. Lifelong census literates are easily created. Lifelong readers, at grade 2 level or better, are not.

- (2) Students: Even after completing grade 5, we found that a quarter of the students could not read at all and only about 12% became good and presumably lifelong readers (13.9% boys and 9.3% girls). Interestingly, girls who persevered to grade 8 actually included a markedly greater proportion of good readers than boys – 67.6% as compared to 55.8%.

Non-students: Evidence of reading skill loss comes from the finding that 37.9% of non-students who had completed grade 5, could not read at all, as compared to 25.5% for students at the same education level. The loss is substantial among female eighth-grade completers – only 43.3% remained good readers later in life, as compared to 67.6% among female students with the same education. Males reverse this trend, with 69.5% remaining good readers as non-students, compared to only 55.8% being good readers as students.

- (3) Students: After grade 5, 74.5% acquire good or early-reading skills. Most are early readers (62.7%). Among female students, 9.3% are good readers after a grade five education as compared to 13.9% for males.

Non-students: Among fifth-grade completers, only 62.1% (down from 74.5% for students) maintain some reading ability – another sign of skill loss after schooling. Among female non-students, 13.7% are good readers after a grade 5 education as compared to 8.3% for males. Curiously, this is the opposite trend to that we found for males and females with a grade 8 education.

This suggests that for those who have completed a grade 8 education, reading improvement and practice is more likely to become a part of everyday life for

males than for females. However, among those with a grade 5 education, a greater proportion of females continues to engage in reading and continues to progress whereas men experience skill loss.

- (4) **Students:** There is an overestimation of literacy (CM) by at least 10% up to seventh-grade completion. At fifth grade the overestimation is 24%.
Non-students: The overestimation at fifth grade is 35.7% (21.3% for females and 48.3% for males).
- (5) **Students:** The proportion of good readers to all readers exceeds 50% only in grade 8 and beyond. At present, grade 8 achievement seems to be the minimum expectation for raising lifelong readers.
Non-students: For female non-students, the proportion of good readers to all readers surpasses 50% only at tenth grade.

Skill loss after schooling

Based on the figures reported in Appendix 2, we plotted good reading proportions (Fig. 3) and non-reading proportions (Fig. 4) by grade completion separately for in-school and out-of-school samples. Figure 3 shows that a person who becomes a good reader in school tends to remain a good reader later in life. Some who may have been weak readers even after grade 6 and 7 education in school, may still go on to become good readers later on, probably through continued reading engagement. However, Fig. 4 shows that the proportion of non-readers in the out-of-school sample at every grade achievement up to grade 6 is substantially higher than in the in-school sample. Many people who may have acquired weak reading skills in school end up as non-readers later in life. Taken together, Figs. 3 and 4 present a strong indication of reading skill loss among millions of weak readers after they leave schooling.

To understand skill loss, ideally one would follow cohorts through schooling and later in life. Our methodology did not allow for this, so a proxy for skill loss after schooling is to look at the proportions of good, weak and non-readers in in-school and out-of-school samples, comparing them at the same grade level. In the analysis, we included those who had been out-of-school for at least 3 years at the time of data

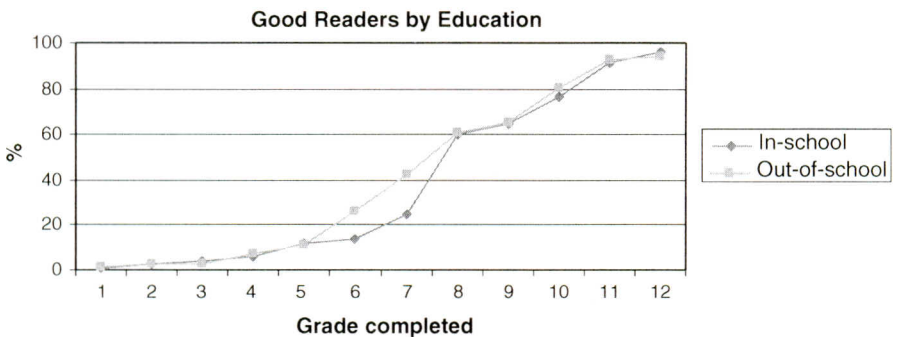


Fig. 3 Good readers in school remain good readers in life

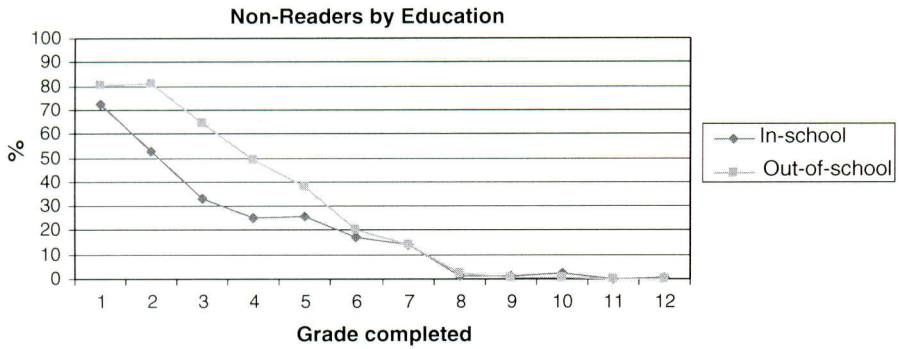


Fig. 4 Proportion of non-readers by education

collection (beyond-school group). The change in the percentage of good, weak and non-readers for the three reading levels is presented by grade in Appendix 3 and plotted in Fig. 5. For a given grade completed, a zero percent change implies that the proportion of, say, good readers (or weak or non-readers) is the same in the in-school and beyond-school groups. A positive “% Change” means there is a greater proportion in the beyond-school group, which would be desirable in the case of good readers and not so encouraging in the case of non-readers.

We infer from Fig. 5 that anyone who has become a good reader by grade 5 is likely to remain a good reader later in life. Most persons acquiring weak reading skills up to grade 3 will relapse into non-reading. Some of those who persevere to grade 5 will probably remain weak readers, and are unlikely to become good readers in the absence of any intervention. Grade 6 is the turning point, where some who may have acquired weak skills in school may go on to become good readers.

We compared the mean education of students found to be good, weak and non-readers by running an analysis of variance test. We repeated the same with those out

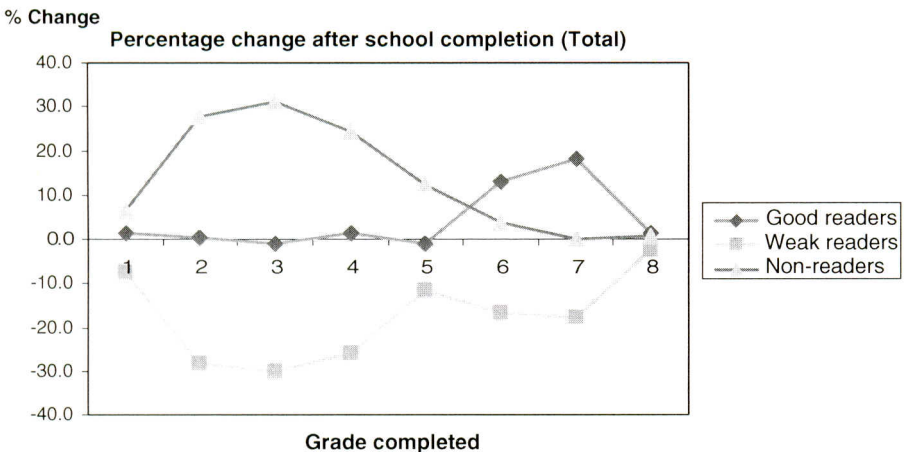


Fig. 5 Reading skill loss beyond schooling

Table 7 Mean education at different literacy levels

Reading level found	Students: Mean education (st. dev.)*	Out-of-school but schooled: mean education (st. dev.)*
Good readers	8.8 (2.6)	9.8 (2.4)
Weak readers	4.6 (2.3)	6.6 (2.3)
Non-readers	2.8 (1.9)	3.9 (1.8)

* $p < 0.000$

of school but who had gone through at least some schooling. In both cases, we found the difference in mean education between good, weak and non-reading groups to be highly significant (Table 7).

An average education of grade 9 is necessary to become a good reader in school, but to become a lifelong good reader, a grade 10 education is required at the present quality of education in rural schools. A grade 4 to 7 education is more likely to result in weak reading skills in school and later in life. Those who do not complete primary education to grade 5 are very likely to be non-readers later in life, assuming no additional intervention.

Discussion

The Indian census accepts a definition of literacy provided by UNESCO as: “An ability to read and write, with understanding, a short simple statement relevant to one’s everyday life.” Conceptually it is a perfectly acceptable definition, it is however not easy to operationalise for a census and moreover a census as immense as India’s. The Indian census has simplistically operationalised the UNESCO definition by asking a senior member of every household to report for every member aged seven and above, whether he/she is “literate” or “illiterate”. As we have argued, this approach is flawed because it can result in an overestimation of the literacy rate by 16%, even at a very basic definition of reading literacy, and as high as 43% if one tests for minimum reading ability at grade 2 level before counting someone as literate.

Understandably, many policy-makers are sharply focused on raising the census literacy rate. That is the benchmark by which progress in literacy is assessed, nationally and internationally. So the shortcut to raising the national literacy rate is to follow strategies which merely succeed in getting more people to report themselves as literate. The literacy rate has thus become a relative measure of progress from one census to another, but it is not a measure of the proportion of people who are actually able to read a simple text. And yet that is the underlying impression it creates.

We do not disregard the importance of estimating relative progress in the literacy rate, especially since it has been measured using the same approach for several decades. But we caution that an increase in the literacy rate, as measured by the census, can give policy-making an exaggerated sense of accomplishment, so much so that it is ultimately detrimental to the cause of genuine literacy. A person

reported as "literate" in the census cannot be assumed to stay on course to becoming a good reader. As we found, the minimum threshold for acquiring lifelong reading ability at grade 2 level, at present quality of formal education in rural India, is a grade 8–9 education. We noted that a majority of people prematurely report themselves as "literate" upon completion of first grade.

In the 1990s, the National Literacy Mission (NLM), in partnership with civil society institutions and volunteers, is credited to have started an estimated 100 million illiterate people on the path to literacy through mass literacy campaigns (Chandran 1994; Yagi 2006). At the heart of NLM's strategy was the Total Literacy Campaign (TLC) unleashed in every Indian district. It broadly involved the following stages: (a) create a culturally-charged environment to drum up a district-wide interest and enthusiasm for mass literacy, (b) conduct a survey to identify non-literates, (c) mobilise an army of Volunteer Teachers (VTs) and Master Trainers, (d) open Learning Centres (LCs) keeping in view learners' local needs and constraints, (e) monitoring and evaluation and finally (f) handing the learners over to post-literacy. Thus, within a short time frame of typically less than one year, many youth and adult learners went from complete non-literacy to "literacy" to "post-literacy". As Ghose (2007) concludes, "In the absence of a literate environment the investments made in making people literate is as good as providing water in leaking glasses."

The first few steps towards literacy were sufficient for the census to add 100 million more people to the ranks of the literate. Srivastava and Patel (2006) capture both the achievements and unmet challenges of NLM's efforts. Based on our findings, it is our contention that most of those who became "literate" via the NLM's literacy campaigns did not acquire the necessary foundation to become lifelong readers. It is an open question how reading-literate they are today, but if the present findings are any indication, they do not seem to have progressed too far from the starting line. A critical weakness in the NLM's strategy was the inadequate creation of lifelong reading opportunities for early readers, both during and after the acquisition of rudimentary skills (Rogers 2002). Other than boutique small-scale projects to keep a few people on the path to regular reading, there is little by way of national scale that has made reading an integral part of most new literates' everyday lives.

Several studies point to the importance of the Home Literacy Environment (HLE), conceptualised as the complex of attitudes, resources and activities related to literacy in which preschool children grow up. Burgess et al. (2002) tested the influence of HLE on four different predictors of future reading achievement: oral language ability, phonological sensitivity, letter-name knowledge and word decoding. Their conceptualisation of HLE included considerations of:

- (i) parental attitude toward literacy, their reading ability, education and socio-economic status,
- (ii) home reading practices which parents engage in, by themselves and with children, e.g. shared reading,
- (iii) oral communicative activities and interaction with children around songs, stories, riddles, rhyming games, etc. and finally,
- (iv) elders' modelling of literacy usage, i.e. when children see parents reading newspapers or books or watching certain programmes on TV.

They found that the diverse conceptualisations of HLE fostered reading and language development. There is a growing collection of studies which further confirm the import of HLE (Rodriguez et al. 2009; Evans and Shaw 2008; Hood et al. 2008; van Steensel 2006). Using PIRLS 2001 data from 25 countries, Park (2008) found that in almost all the countries, HLEs, as determined by early home literacy activities, parental attitudes toward reading and the number of books at home, are significantly associated with the subsequent development of reading abilities.

Very few studies have, however, emerged from a conceptualisation of HLE in developing countries, where in many cases the children are in the process of becoming first-generation readers, growing up in a multilingual and resource-poor milieu. The only study we could find on HLE, coming out of India, was conducted with middle-class urban children living in homes where parents are able to read and speak in English (Kalia and Reese 2009). But the question is: What constitutes an effective HLE in 70% or more households in India who cannot afford books or any other reading material, where many parents or other adults at home might be weak or non-readers and where multilingualism is the norm?

One approach to enriching HLE in low-resource contexts is to see what already enters people's households or forms a part of their lives and to infuse it with a literacy-promoting component. For instance, the Same Language Subtitling (SLS) project has, since 2002, been adding karaoke-like subtitles of lyrics to existing Bollywood film songs on national and state TV networks in 10 languages. SLS is designed to invite a "what you hear is what you read" effect. This simple introduction of SLS, if scaled up on every film song on TV, can infuse inescapable reading practice for an estimated 650 million viewers at present, of whom over 200 million are also weak readers. Importantly, the addition of SLS was preferred by over 90% viewers and led to measurable decoding skill improvement among school children, youth and adults (Kothari et al. 2004). In the developing world, interventions which can leverage mass media to deliver and strengthen any aspect of literacy are of particular interest from a cost and HLE perspective.

However, before enriched HLE initiatives can take root, policy makers would need to accept that the goal of increasing national literacy is not to be able to count more and more people as "literate," but to ensure that they achieve, at the minimum, basic lifelong reading ability which allows them to read and understand, for example, the day's newspaper headlines in any one language. Our research would suggest that only 40% of the officially "literate" people in India can, at present, read and understand newspaper headlines.

To begin addressing the national problem of low reading literacy achievement, the census needs to tighten the way it measures literacy by drawing upon the definition of literacy as used in the Multiple Indicator Cluster Surveys (MICS) mentioned in UNESCO (2008): "Literacy is the ability to read easily or with difficulty a letter or a newspaper." Unfortunately, we did not include this question in our own survey. This would have given us a good perspective on how the resulting literacy rates – from the census, MICS, and reading methods – would compare against each other. Our contention is that a simple question, "Are you able to read a newspaper or a letter?", would, on a qualitative scale – easily, with difficulty, or not

at all – have resulted in a much more accurate measure of basic reading literacy than the current census approach of having a household member report for every member whether he/she is literate. The irony is that even scholars continue to debate the definition of “literacy”. Yet we have no problem throwing the term into a census questionnaire. Why wouldn't anyone who has taken the first steps toward literacy report himself/herself or be reported as “literate” when the opposite choice is socially uncomfortable? The reading method employed in this study may not be pragmatic on the scale of the Indian census. However, sample testing of census literates can serve as a much needed corrective. A high literacy rate in the census is good for national pride. But if it is genuinely high, it is good for national progress.

Appendix 1: Reading exercise: text in Hindi at grade 2 level

Ask to read any one of the passages.

Circle any syllable read incorrectly.

- (1) एक था हाथी । जिसका नाम था सोमा । सोमा जंगल में रहता था । वह रोज नहाने नदी जाता था । एक दिन नदी में उसे तरबूज मिला । तरबूज उसे बहुत मीठा लगा । तब से सोमा नदी किनारे रहने लगा ।
- (2) एक थी मैना । जिसका नाम था सीमा । सीमा जंगल में रहती थी । वह रोज नहाने नदी जाती थी । एक दिन नदी में उसे अमरूद मिला । अमरूद उसे बहुत मीठा लगा । तब से सीमा नदी किनारे रहने लगी ।
- (3) एक था कछुआ । जिसका नाम था भीमा । भीमा नदी किनारे रहता था । वह रोज घुमने जंगल जाता था । एक दिन जंगल में उसे टमाटर मिला । टमाटर उसे बहुत मीठा लगा । तब से भीमा जंगल में रहने लगा ।

	कहानी का नं.	कैसे पढ़ा
अटक अटक के पढ़ा 1. जोड़नी से पढ़ा 2. नहीं पढ़ा 3.	<input type="checkbox"/>	<input type="checkbox"/>

Appendix 2

See Table 8.

Table 8 Reading ability by formal education

Grade completed	Census method Literacy (%) CM	Reading method				n	Minimum overestimation by CM (%) CM - (A + B)	% Good readers among readers A/(A + B)
		Literacy at best (%) A + B	Good readers (%) A	Early-readers (%) B	Non-readers (%)			
<i>Students</i>								
Total								
0	25.3	8.0	1.6	6.4	92.0	188	17.3	20.0
1	90.0	27.6	0.6	27.0	72.4	681	62.4	2.2
2	96.2	47.4	2.3	45.1	52.6	743	48.8	4.9
3	98.4	66.6	3.9	62.7	33.4	620	31.8	5.9
4	98.7	75.1	5.7	69.4	25.0	617	23.6	7.6
5	98.5	74.5	11.8	62.7	25.5	534	24.0	15.8
6	98.9	83.0	13.9	69.1	17.0	446	15.9	16.7
7	99.1	85.7	24.7	61.0	14.3	449	13.4	28.8
8	99.2	98.7	60.3	38.4	1.4	365	0.5	61.1
9	100.0	98.8	64.4	34.4	1.3	320	1.2	65.2
10	100.0	97.6	76.7	20.9	2.4	206	2.4	78.6
11	100.0	100.0	91.3	8.7	0.0	92	0.0	91.3
12	100.0	99.4	95.8	3.6	0.6	168	0.6	96.4
<i>Female</i>								
0	28.6	6.7	1.9	4.8	93.3	105	21.9	28.4
1	90.8	22.5	0.0	22.5	77.5	289	68.3	0.0
2	95.1	49.3	2.1	47.2	50.7	335	45.8	4.3
3	98.0	64.5	4.5	60.0	35.5	290	33.5	7.0
4	100.0	75.7	5.2	70.5	24.4	271	24.3	6.9
5	99.2	73.9	9.3	64.6	26.0	246	25.3	12.6
6	98.6	79.3	14.9	64.4	20.8	202	19.3	18.8
7	99.1	88.5	25.0	63.5	11.5	208	10.6	28.2
8	99.3	99.3	67.6	31.7	0.7	139	0.0	68.1
9	100.0	100.0	67.5	32.5	0.0	114	0.0	67.5
10	100.0	96.4	74.4	22.0	3.7	82	3.6	77.2
11	100.0	100.0	93.1	6.9	0.0	29	0.0	93.1
12	100.0	100.0	97.8	2.2	0.0	45	0.0	97.8
<i>Male</i>								
0	21.2	9.6	1.2	8.4	90.4	83	11.6	12.5
1	89.4	31.4	1.0	30.4	68.6	392	58.0	3.2

Table 8 continued

Grade completed	Census method Literacy (%) CM	Reading method				n	Minimum overestimation by CM (%) CM - (A + B)	% Good readers among readers A/(A + B)
		Literacy at best (%) A + B	Good readers (%) A	Early-readers (%) B	Non-readers (%)			
2	97.1	45.9	2.5	43.4	54.2	408	51.2	5.4
3	98.8	68.5	3.3	65.2	31.5	330	30.3	4.8
4	97.7	74.6	6.1	68.5	25.4	346	23.1	8.2
5	98.0	75.0	13.9	61.1	25.0	288	23.0	18.5
6	99.2	86.1	13.1	73.0	13.9	244	13.1	15.2
7	99.2	83.4	24.5	58.9	16.6	241	15.8	29.4
8	99.1	98.3	55.8	42.5	1.8	226	0.8	56.8
9	100.0	98.0	62.6	35.4	1.9	206	2.0	63.9
10	100.0	98.4	78.2	20.2	1.6	124	1.6	79.5
11	100.0	100.0	90.5	9.5	0.0	63	0.0	90.5
12	100.0	99.2	95.1	4.1	0.8	123	0.8	95.9

Grade completed	Census method Literacy (%) CM	Reading method				n	Literacy (%) CM	Literacy at best (%) A + B
		Literacy at best (%) A + B	Good readers (%) A	Early-readers (%) B	Non-readers (%)			
<i>Non-students</i>								
Total								
0	5.2	1.6	0.1	1.5	98.4	5,383	3.6	6.3
1	80.2	19.6	1.6	18.0	80.5	128	60.6	8.2
2	79.7	18.9	2.7	16.2	81.1	334	60.8	14.3
3	95.2	35.8	2.8	33.0	64.2	282	59.4	7.8
4	96.2	50.9	7.0	43.9	49.1	399	45.3	13.8
5	97.8	62.1	10.8	51.3	37.9	951	35.7	17.4
6	98.1	79.8	26.3	53.5	20.2	312	18.3	33.0
7	99.1	85.7	42.5	43.2	14.3	537	13.4	49.6
8	98.9	97.7	60.6	37.1	2.3	1,138	1.2	62.0
9	100.0	99.5	65.2	34.3	0.5	581	0.5	65.5
10	100.0	99.2	80.7	18.5	0.8	763	0.8	81.4
11	100.0	100.0	92.7	7.3	0.0	151	0.0	92.7
12	100.0	100.0	94.2	5.8	0.0	573	0.0	94.2
Female								
0	5.2	1.5	0.1	1.4	98.5	3,672	3.7	6.7
1	85.0	11.9	0.0	11.9	88.1	59	73.1	0.0
2	71.2	18.7	3.6	15.1	81.3	166	52.5	19.3
3	94.1	37.9	4.3	33.6	62.1	116	56.2	11.3
4	95.0	61.8	7.9	53.9	38.2	191	33.2	12.8
5	98.0	76.7	13.7	63.0	23.3	446	21.3	17.9

Table 8 continued

Grade completed	Census method Literacy (%) CM	Reading method						Literacy (%) CM	Literacy at best (%) A + B
		Literacy at best (%) A + B	Good readers (%) A	Early-readers (%) B	Non-readers (%)	n			
6	97.5	84.0	17.6	66.4	16.0	119	13.5	21.0	
7	98.5	90.5	34.5	56.0	9.5	200	8.0	38.1	
8	98.2	96.4	43.3	53.1	3.6	386	1.8	44.9	
9	100.0	98.6	44.8	53.8	1.4	145	1.4	45.4	
10	100.0	98.1	65.9	32.2	1.9	208	1.9	67.2	
11	100.0	100.0	80.0	20.0	0.0	20	0.0	80.0	
12	100.0	100.0	87.6	12.4	0.0	105	0.0	87.6	
Male									
0	5.2	1.8	0.2	1.6	98.2	1,711	3.4	11.1	
1	76.1	26.1	2.9	23.2	73.9	69	50.0	11.1	
2	88.0	19.1	1.8	17.3	81.0	168	68.9	9.4	
3	96.0	34.3	1.8	32.5	65.7	166	61.7	5.2	
4	97.3	40.9	6.3	34.6	59.1	208	56.4	15.4	
5	97.6	49.3	8.3	41.0	50.7	505	48.3	16.8	
6	98.5	77.2	31.6	45.6	22.8	193	21.3	40.9	
7	99.4	82.8	47.2	35.6	17.2	337	16.6	57.0	
8	99.2	98.4	69.5	28.9	1.6	752	0.8	70.6	
9	100.0	99.8	72.0	27.8	0.2	436	0.2	72.1	
10	100.0	99.6	86.3	13.3	0.4	555	0.4	86.6	
11	100.0	100.0	94.7	5.3	0.0	131	0.0	94.7	
12	100.0	100.0	95.7	4.3	0.0	468	0.0	95.7	

Appendix 3

See Table 9.

Table 9 Percentage change of good, weak and non-readers beyond school

Grade completed at the time of survey	Total (%)			Male (%)			Female (%)		
	Good	Weak	Non	Good	Weak	Non	Good	Weak	Non
I									
In-school	0.6	27.0	72.4	1.0	30.4	68.6	0.0	22.5	77.5
Beyond-school	1.8	19.5	78.8	3.3	24.6	72.1	0.0	13.5	86.5
% Change	1.2	-7.5	6.4	2.3	-5.8	3.5	0.0	-9.0	9.0

Table 9 continued

Grade completed at the time of survey	Total (%)			Male (%)			Female (%)		
	Good	Weak	Non	Good	Weak	Non	Good	Weak	Non
II									
In-school	2.3	45.1	52.6	2.5	43.4	54.2	2.1	47.2	50.7
Beyond-school	2.8	16.8	80.4	1.9	18.0	80.1	3.7	15.5	80.7
% Change	0.5	-28.3	27.8	-0.6	-25.4	25.9	1.6	-31.7	30.0
III									
In-school	3.9	62.7	33.4	3.3	65.2	31.5	4.5	60.0	35.5
Beyond-school	3.0	32.6	64.4	1.9	31.4	66.7	4.5	34.2	61.3
% Change	-0.9	-30.1	31.0	-1.4	-33.8	35.2	0.0	-25.8	25.8
IV									
In-school	5.7	69.4	25.0	6.1	68.5	25.4	5.2	70.5	24.4
Beyond-school	7.2	43.4	49.4	6.5	33.8	59.7	8.1	53.8	38.2
% Change	1.5	-26.0	24.4	0.4	-34.7	34.3	2.9	-16.7	13.8
V									
In-school	11.8	62.7	25.5	13.9	61.1	25.0	9.3	64.6	26.0
Beyond-school	10.8	51.2	37.9	6.5	33.8	59.7	8.1	53.8	38.2
% Change	-1.0	-11.5	12.4	-7.4	-27.3	34.7	-1.2	-10.8	12.2
VI									
In-school	13.9	69.1	17	13.1	73	13.9	14.9	64.4	20.8
Beyond-school	26.9	52.5	20.6	31.7	45	23.3	18.8	65.2	16.1
% Change	13.0	-16.6	3.6	18.6	-28.0	9.4	3.9	0.8	-4.7
VII									
In-school	24.7	61	14.3	24.5	58.9	16.6	25	63.5	11.5
Beyond-school	42.7	43.1	14.2	47.2	35.9	16.9	34.9	55.6	9.5
% Change	18.0	-17.9	-0.1	22.7	-23.0	0.3	9.9	-7.9	-2.0
VIII									
In-school	60.3	38.4	1.4	55.8	42.5	1.8	67.6	31.7	0.7
Beyond-school	61.7	36.1	2.2	70.1	28.6	1.4	44.9	51.2	3.9
% Change	1.4	-2.3	0.8	14.3	-13.9	-0.4	-22.7	19.5	3.2

Beyond-school = At least 3+ years after school completion

References

- Abadzi, H. (2003). Improving adult literacy outcomes: Lessons from cognitive research for developing countries. World Bank. Eric Document (ED) 478961.
- Annual Survey of Education Report (ASER). (2009). <http://www.asercentre.org/asersurvey/aser09/pdfdata/aser09.pdf>, checked on April 20, 2010.
- Burgess, S. R., Hecht, S. A., & Lonigan, C. J. (2002). Relations of the Home Literacy Environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading Research Quarterly*, 37(4), 408–426.
- Chand, P. (2007). *Status of adult literacy in India: A database for literacy programmes*. National Literacy Mission, Directorate of Adult Education.
- Chandran, K. N. (1994). Literacy in India and the example of Kerala. *Journal of Reading*, 37(6), 514–517.

- Darcovich, N. (2000). The measurement of adult literacy in theory and in practice. *International Review of Education*, 46(5), 367–376.
- Evans, M. A., & Shaw, D. (2008). Home grown for reading: Parental contributions to young children's emergent literacy and word recognition. *Canadian Psychology*, 49(2), 89–95.
- Ghose, M. (2007). Gender, literacy and women's empowerment in India: Some issues. *Convergence*, 40(3–4), 193–201.
- Gomez, S. V. (2000). The collective that didn't quite collect: Reflections on the IALS. *International Review of Education*, 46(5), 419–431.
- Hamilton, M., & Barton, D. (2000). The international adult literacy survey: What does it really measure? *International Review of Education*, 46(5), 377–389.
- Hood, M., Conlon, E., & Andrews, G. (2008). Preschool home literacy practices and children's literacy development: A longitudinal analysis. *Journal of Educational Psychology*, 100(2), 252–271.
- Jennings, Z. (2000). Adult literacy in the Commonwealth Caribbean with special reference to a study of the functional literacy of young Guyanese adults. *International Journal of Lifelong Education*, 19(5), 385–406.
- Kalia, V. & Reese, E. (2009). Relations between Indian children's home literacy environment and their English oral language and literacy skills. *Scientific Studies of Reading*, 13(2), 122–145.
- Kothari, B., Pandey, A., & Chudgar, A. (2004). Reading out of the "idiot box": Same-language subtitling on television in India. *Information Technologies and International Development*, 2(1), 23–44.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). PIRLS 2006 international report. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Nath, S. R. (2007). Self-reporting and test discrepancy: Evidence from a National Literacy Survey in Bangladesh. *Review of Education*, 53, 119–133.
- Naumann, J. (2005). TIMSS, PISA, PIRLS and low educational achievement in world society. *Prospects*, 35(2), 229–248.
- Park, H. (2008). Home literacy environments and children's reading performance: A comparative study of 25 countries. *Educational Research and Evaluation*, 14(6), 489–505.
- Rodriguez, E. T., Tamis-LeMonda, C. S., Spellmann, M. E., Pan, B. A., Raikes, H., Lugo-Gil, J., et al. (2009). The formative role of home literacy experiences across the first three years of life in children from low-income families. *Journal of Applied Developmental Psychology*, 30, 677–694.
- Rogers, A. (2002). Post-literacy and second-stage adult learning in India. *Compare*, 32(2), 149–179.
- Srivastava, K., & Patel, I. (2006). Community mobilisation, gender equality and resource mobilisation in adult education. *International Journal of Educational Development*, 26, 153–165.
- UNESCO. (2008). Education for all by 2015: Will we make it? Education for All (EFA) global monitoring report. Oxford University Press. Downloadable at: <http://unesdoc.unesco.org/images/0015/001547/154743e.pdf>; checked on April 20, 2010.
- van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading*, 29(4), 367–382.
- Yagi, R. (2006). Training and supporting the 'volunteers': Revisiting total literacy campaigns in India. *International Journal of Educational Development*, 26, 67–87.

The authors

Brij Kothari is an adjunct professor at the Indian Institute of Management, Ahmedabad (IIMA), the President of PlanetRead, and an Ashoka and Schwab Fellow. He and his team have innovated, researched, and nationally implemented Same Language Subtitling (SLS) in India, for mass literacy.

Tathagata Bandyopadhyay is a professor of production and quantitative methods at IIMA. His research interests include theoretical statistics and its applications, especially in the fields of health and education. He has taught statistics at the University of Calcutta for more than two decades and has served as visiting faculty at a number of universities in the US and Canada.