Effects of mother tongue education and multilingualism on reading skills in the regional language and English in India

tesol QUARTERLY
tesol international association

Effects of Mother Tongue Education and Multilingualism on Reading Skills in the Regional Language and English in India

MARGREET VOGELZANG , IANTHI MARIA TSIMPLI , AND ANUSHA BALASUBRAMANIAN
Theoretical and Applied Linguistics Section, Faculty of Modern and Medieval Languages & Linguistics, University of Cambridge
Cambridge, UK

MINATI PANDA
School of Social Sciences, Jawaharlal Nehru University
Delhi, India

SUVARNA ALLADI AND ABHIGNA REDDY
Department of Neurology, National Institute of Mental health and Neuro Sciences (NIMHANS)
Bengaluru, India

LINA MUKHOPADHYAY
Department of Training and Development, School of English Language Education, The English & Foreign Languages University (EFL-U)
Hyderabad, India

JEANINE TREFFERS-DALLER
Department of English Language and Applied Linguistics, University of Reading
Reading, UK

THEODOROS MARINIS
Department of Linguistics, University of Konstanz
Konstanz, Germany; and
School of Psychology and Clinical Language Sciences, University of Reading
Reading, UK

Abstract

In a highly multilingual country like India, challenges and opportunities arise in education and language policy. Although multilingualism is often associated with developmental advantages, Indian primary school children generally show low learning outcomes, specifically on literacy. Here we examine the influence of mother tongue education and multilingualism on the reading skills and reading comprehension of 1272 Indian primary school children from low SES homes. The children performed the ASER literacy task in both the regional, majority language and in English, which was followed by newly developed reading comprehension questions. The results show that minority language speakers from monolingual households—who do not receive mother tongue education—underperform compared to majority language speakers when reading in the majority language. When reading in English, growing up in a multilingual household improves children's performance. Finally, in sites which have a larger proportion of mother tongue-educated children, children perform better in literacy in the regional language and worse in English. Overall, these results provide insight into the influence of mother tongue education and multilingualism on reading abilities and show that more support is needed for minority language speakers to develop literacy in the majority language, and for all children to develop literacy in English.
doi: 10.1002/tesq.3326

INTRODUCTION

India is one of the most multilingual and linguistically diverse countries in the world, with 22 recognized languages potentially used as mediums of instruction and a total of around 462 different languages spoken in the country (Simons & Fennig, 2018). It is thus inevitable that children are not always educated in their mother tongue or home language. Crucially, learning outcomes are reportedly low in Indian primary schools, in particular with regard to literacy skills (Pratham, 2019; UNESCO, 2020/2021). However, it is largely unknown how mother tongue education relates to literacy development in the Indian population. It is important to learn more about this in order to develop adequate language policies and advise educators. In this article, we focus on literacy skills of Indian primary school children from low socio-economic backgrounds and examine the effect that mother tongue education has in such a multilingual society. We consider the

advantages as well as the challenges that are associated with a multilingual society and aim to experimentally identify factors related to language and education that impact children's literacy skills in both the regional language (Hindi or Telugu) and in English as an additional language.

Literacy Development: Issues and Influences

A specially developed widely used tool to assess basic reading skills in India is the Annual Status of Education Report literacy task (ASER task; Pratham, 2019). Bi-annual national surveys between 2012 and 2018, which used the ASER task with over 500,000 children across India, consistently found low learning outcomes in letter, word, and sentence reading, with more than half of children in Year 5 (age 9–10) being unable to read a Year 2 level text in the regional language fluently (Pratham, 2019). Moreover, a large longitudinal project called LiRIL examined literacy in Year 1 to 3 children in India from two socio-economically disadvantaged regions—Yadgir and Palghar—in the children's first language, which was also the dominant regional language (Kannada or Marathi). They found that even when children were able to read the script, they were often unable to understand what they were reading in the regional language, even in Year 3 (Menon et al., 2017). This same study found that at the end of Year 3, most children could read a simple list of words, but less than a quarter could read a list of words at their grade level. Improving these early reading skills in children is important, as they are associated with greater academic achievement at a later age (Blackenbecker, 2020; Duncan et al., 2007; Patel, Torppa, Aro, Richardson, & Lyytinen, 2022; Pluck, 2019; Rabiner, Godwin, & Dodge, 2016) and form the basis for children's learning in school and everyday life.

Importantly, reading is a skill that needs to be taught, usually at school. Reading for comprehension is a higher-level skill which develops as a result of good decoding in combination with listening comprehension skills (the "Simple View of Reading"; Gough & Tunmer, 1986; Scarborough, 2001). This idea has been used mostly to explain reading in alphabetic languages but there is some research on alpha-syllabaries as well (Nag & Snowling, 2012; Nakamura, Koda, & Joshi, 2014; Nishanimut, Johnston, Joshi, Thomas, & Padakannaya, 2013). Although some reading skills can transfer between different languages (Bialystok, 2001; Dressler & Kamil, 2006; Durgunoğlu, 2002; Koda, 2008), such as comprehension strategies (Jiménez, 1997; van Gelderen, Schoonen, Stoel, de Glopper, & Hulstijn, 2007), an independent contribution of vocabulary knowledge to reading comprehension is specific to the language

used (Cobo-Lewis, Pearson, Eilers, & Umbel, 2002; Nakamoto, Lindsey, & Manis, 2008).

Thus, building vocabulary in a certain language is a prerequisite for listening and reading comprehension. This is particularly relevant for speakers of minority languages who do not speak the regional language at home, and multilinguals, who generally have a less extensive vocabulary compared to monolinguals in one specific language (Bialystok, 2009; although they have larger vocabulary when considering all of their languages, Oh, Mancilla-Martinez, & Hwang, 2023). For example, a child in Delhi could speak Bhojpuri (a minority language) as their mother tongue and home language, Hindi as the regionally dominant, majority language and the main medium of instruction at school, and English as an additional language at school. Since this type of multilingualism is the everyday reality in India and multilingual students are in the majority (Mohanty, 2010, 2019), many children will have a mother tongue or home language that is different from the language(s) used in school[1] (Jhingran, 2019). The mixing of cultures and languages is further increased in larger urban contexts, which often see migrants from other regions (Lightfoot et al., 2021). Although in recent years more inclusive and multilingual learning practices in Indian schools have actively been advocated (Mohanty, 2019; Nag, 2017) and evidence shows that supporting both the regional language and the language of minority speakers leads to developmental advantages (Cummins, 2009; Leikin, Schwartz, & Share, 2009), this is currently not common practice in India.[2] In addition, educators themselves are not always equipped with the same bi-/multilingualism as their pupils (Blair, Haneda, & Bose, 2018). This could cause multilinguals or minority language-speaking children to have problems acquiring literacy skills in the majority language or in English, languages which they are expected to learn in school. Indeed, bilingualism has been associated with weaker reading development in immigrant minority language-speaking children with English as a second language in the UK (Babayiğit & Shapiro, 2020). More generally, it has been argued that mother tongue education is valuable for

[1] Lightfoot, Balasubramanian, Tsimpli, Mukhopadhyay, and Treffers-Daller (2021) found a predominant use of the regional language and of language mixing in Indian primary school classrooms—by both teachers and students. Similarly, self-reports of children from the Indian Barwani district indicate that nearly half of them sometimes mix their language with other languages in school, and that their teachers do the same (Panda, 2019).

[2] The National Education Policy (Ministry of Human Resource Development, 2020) aims to change this practice by promoting mother-tongue/home language/regional language education for all children at least up to Year 5 and preferably up to Year 8 (p. 14). For a comparative analysis of language-in-education policies and mother tongue-based multilingual education (MTB-MLE) in India, see Mohanty (2019).

learning and learning outcomes, especially in primary school-aged children (Ball, 2011; Ouane & Glanz, 2010; Romaine, 2013). Thus, children who have a home language different from the regional language, which is predominantly used in schools, may be disadvantaged in their academic development. As these are often members of minority groups that already experience social disadvantages that negatively affect learning outcomes (e.g., low socio-economic status; Alcott & Rose, 2017), learning in a language different from their home language may be an additional disadvantage (Mohanty, 2010, 2019). As a consequence, these children may experience difficulties understanding the subject matter being taught (Bhattacharya, 2013; Endow, 2018). To address this, more research into reading development in minority language speakers and multilinguals from disadvantaged socio-economic backgrounds is paramount.

Aside from potential difficulties that multilinguals face in their development because the educational system does not provide sufficient support, multilingualism is sometimes also associated with developmental advantages. Specifically, multilinguals may have better phonological awareness (Canbay, 2011; Souza & Leite, 2014, but see Bialystok, Majumder, & Martin, 2003), potentially aiding the decoding process in reading. Additionally, linguistic interdependence theory (Cummins, 1979) argues for positive relations between a bilingual's different languages, as evidenced in, for example, children's (receptive) vocabulary size in their first language predicting vocabulary size in a second language (Sierens, Slembrouck, Van Gorp, Agirdag, & Van Avermaet, 2019). Conversely, the complexities of learning to read in a language when another language also exists in the mind have also been emphasized (Bernhardt, 2003). Finally, cognitive benefits of bi-/multilingualism (e.g., Bialystok, 2009; Costa, Hernández, & Sebastián-Gallés, 2008; but see Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Paap & Greenberg, 2013; Samuel, Roehr-Brackin, Pak, & Kim, 2018) may improve learning outcomes and have been found to interact with literacy levels (Kaufman, DeYoung, Gray, Brown, & Mackintosh, 2009; Kosmidis, Zafiri, & Politimou, 2011; Ziegler & Goswami, 2005). However, it should be noted that pupils' cognitive benefits from bilingualism are related to their perceptions of teacher appreciation of their home language, and thus that there is a sociolinguistic component to it (Goriot, Denessen, Bakker, & Droop, 2016).

A factor additionally affecting reading skills in Indian children may be the style of teaching and of school assessments. These often focus on rote learning (Smith, Hardman, & Tooley, 2005) and piecemeal memorization of knowledge expected to be used in summative assessments (Clarke, 2001). In fact, the most commonly used practice in the Indian classroom is a monologic, teacher-led approach which limits

the child's interaction to answering questions (Lyle, 2008). Children's ability to comprehend what they read receives little attention in the classroom (at least for English reading; Endow, 2018). Although learning and teaching styles are not the main focus of the current article, they might cause discrepancies between decoding skills in reading and reading comprehension (such as found in Menon et al., 2017), the former of which is a necessary but not sufficient condition to achieve the latter.

In the current study, we aim to further explore the relation between mother tongue education, multilingualism, and literacy skills. Before moving on to the experimental part of the study, it is important to consider literacy in the Indian context in more detail.

Literacy in the Indian Context: Languages, Scripts, and Learning

The participants in our study had one of the alpha-syllabic Indian languages Hindi or Telugu as the societal, majority language and were exposed to alphabetic English in addition at school. Despite their differences in both origin and script, Hindi and Telugu use similar basic units of writing called "Akshara," which have properties of both alphabets and syllables (Nag, 2011; Share & Daniels, 2016). Aksharas have consonant clusters (CV, CCV combinations) with vowel markers which can occur in different places (Prakash & Joshi, 1995). For example, in Hindi, the syllable sound /ka/ is written as "का" where the vowel marker occurs "after" the consonant symbol. As Aksharas have around 1000 symbols (made up of combinations of consonants/consonant clusters and vowel markers), these scripts are relatively transparent compared to English (with 26 symbols) (Nag, 2007). To illustrate, "कि" always sounds as /ki/ but "c" in English can correspond to /k/ or /s/.

The regularity of reading rules in Hindi/Telugu as well as the one-to-one grapheme/syllable to phoneme correspondence may seem to make them easier to learn than English. However, alpha-syllabaries like Telugu/Hindi require both syllabic and phonemic awareness (Reddy & Koda, 2013; Share & Daniels, 2016; Vaid & Gupta, 2002), making the process of decoding more taxing than in alphabetic languages. Therefore, mastery of decoding skills actually takes longer to learn in alpha-syllabic languages and is not achieved until Year 4 or 5 (Joshi, Nakamura, & Singh, 2017; Nag, 2007, 2011; Nakamura et al., 2014; Tiwari, Nair, & Krishnan, 2011) contrary to alphabetic languages (Grade 1 to 3) (Liberman, Shankweiler, William-Fisher, & Carter, 1974).

Differences can also be found in the way in which the different languages are taught. English is usually taught through the phonics approach where sounds of individual phonemes (Adams, 1990; Goswami, 2005) or letters (Gupta, 2014) are taught and/or through sight-word/whole-word reading (Rao, Vaid, Srinivasan, & Chen, 2011). That is, children are taught letter names and word spelling without orienting them to the sounds of the grapheme. Alpha-syllabaries, in contrast, are learnt through mastering a set of design principles which help to choose not just the vowel marker but also where to place it (Nag, 2022; Rao et al., 2011); children are expected to master these design principles by Year 2 (Joshi et al., 2017). A study by Shenoy, Wagner, and Rao (2020) in Bangalore found that children from low-cost Government schools had difficulty in English decoding even in Year 5, which was attributed to them being taught through the alphabet-spelling method (Gupta, 2014) rather than the phonics method. The type of activities the teachers used in the classroom such as copying from board, choral reading etc. did not facilitate reading development. Additionally, poor English proficiency of teachers and lack of L1 support to scaffold L2 development could have played a role.

The aforementioned transfer that can occur between languages is also dependent on orthographies (Dworin, 2003; Gort, 2006, 2012), as transfer involves applying the graph-phonemic, morphological, semantic, and orthographic rules of one language to the other (Gort, 2006, 2019). Nishanimut et al. (2013) found that use of L1 to teach English (L2) can be beneficial. Mishra and Stainthorp (2007) examined the association between phonological awareness and reading in Oriya-English Year 5 and found that cross-linguistic transfer could be asymmetrical across languages, which could be dependent on characteristics of the writing systems; Oriya phonological awareness transferred to English, but English phonological awareness did not always transfer back. Similar results have been found for transfer between Kannada and English (Reddy and Koda (2013), children aged 10–14) and between Hindi and English (Patel, Chatterjee Singh, and Torppa (2022), Year 1 and 2 children). In a longitudinal study (Years 3–5 to Years 6–8), Nakamura et al. (2014) found that decoding skills and listening comprehension predicted reading comprehension in both Kannada and English, extending the applicability of the Simple View of Reading model to alpha-syllabaries. There was a strong cross-linguistic relationship of reading sub-skills in both languages, suggesting that there is transfer and sharing of resources between languages for linguistic comprehension.

Blackenbecker (2020) emphasizes that early literacy programs often do not consider L1-based multilingual approaches; they often focus only on short-term results and tend to ignore the psycholinguistic,

sociopolitical and sociolinguistic aspects of a multilingual society. This could be detrimental to children's learning when they transition from L1 to L2 (Alidou, Boly, Brock-Utne, & Satina, 2006; Baker, 2011). In contrast, multilingual early education programs with inputs of literacy from both L1 and L2 could have long-lasting benefits (Blackenbecker, 2020; Nakamura et al., 2023; Nakamura, 2023). In India, the absence of well-developed and culturally appropriate early literacy programs could have severe consequences. In a large-scale study, Nakamura, De Hoop, and Holla (2018) investigated reading sub-skills in 550 children (Years 1–5) speaking either alpha-syllabic Kannada or Telugu as their L1 and alphabetic English as an L2. The authors demonstrated that children performed better on Kannada and Telugu decoding compared to English decoding. Their findings also indicate that if children are taught in English before reaching a certain threshold, this may result in difficulties in acquiring English literacy—they concluded that 20% of the children in Year 5 were not ready to be taught in English.

Finally, socio-economic status could also influence literacy development. A recent study by Shenoy, Wagner, Overton, and Rao (2023) examined the effects of socio-economic status – comparing low, middle, and high cost schools—and gender on English reading sub-skills in Year 3 and 5 children in Bangalore. Socio-economic status was a significant predictor of all reading measures in both years, with children from low-cost schools underperforming. In addition, girls outperformed boys on reading comprehension. For this reason, socio-economic status will be kept constant in the sample of children examined in the current study, and the potential influence of gender will be taken into account.

The Current Study

As stated above, this study aims to explore the relation between mother tongue education, multilingualism, and literacy skills. We used the ASER literacy tool to assess literacy skills in both the regional, majority, language and in English as an additional language in a large cohort of children from underprivileged contexts in India. We examined whether (1) decoding ability in reading is related to reading comprehension in the regional language and in English, (2) being schooled in a language different from the primary home language (thus not receiving mother tongue education) disadvantages children in literacy skills in the regional language and in English, (3) growing up in a multilingual household impacts children's literacy skills in the regional language and in English.

The study took place across three sites in India, namely Delhi, Hyderabad, and Patna. In Delhi and Patna, the official/majority/regional language is Hindi, whereas in Hyderabad, it is Telugu. Note that especially in the case of Hindi this is an oversimplification (as it is not merely a regional language, but rather a national language), but for the purposes of the current study we will refer to it as such. Hindi being the national language means that children will have been exposed to it in a variety of situations, for example, in society as well as on TV and radio. However, Telugu is a dominant language in Hyderabad, to which local children will have similarly been exposed in their neighborhoods and through the media (Lightfoot et al., 2021). English, in contrast, is a language to which the children are much less likely to be exposed in their neighborhoods, although English-language media is available far and wide in India.

An important note is that we will limit our definition of multilingualism to children growing up with multiple languages *in the home*. This is different from overall multilingualism, which will be higher when taking into account the home, school, and potentially societal languages. This more constrained definition was chosen for two reasons. Firstly, all children in the Indian school system and society are to some extent multilingual, as they learn multiple languages in school and are almost certainly exposed to multiple languages in the society. Secondly, since it was very difficult to obtain reliable information about children's language use and exposure outside of the home, the more reliable information collected in this project was on language use in the home. Although we are aware of the skepticism against viewing mono/multilingualism as a binary state (rather, it is a continuum, Kašćelan et al., 2022), we chose to proceed with this distinction for the current study.

MATERIALS AND METHODS

Participants

A total of 1272 children from Delhi, Hyderabad, and Patna participated in the study (see Table 1). These three sites were selected to reflect differences in geographical and linguistic context, and thus, to make a wider generalization over different Indian contexts possible. Specifically, Delhi (in the Northwest), is the capital of India and together with Hyderabad (in the North of southern India), it is among the six metropolitan Indian cities with better educational infrastructure. Delhi and Hyderabad have different regionally dominant languages (resp. Hindi and Telugu). Patna (eastern India), in contrast, is a smaller municipal area surrounded by rural blocks and is located in

TABLE 1

Overview of the Children's Demographic Information (f = female; m = male), Including the Proportion of Children Speaking a Minority Language in the Home and of Children Coming from Multilingual Households

Site	Number	Gender	Mean age (SD)	Age range	Language in the home	Percentage of multilinguals[b]
Delhi	387	193 f, 194 m	8.78 (0.64)	8–12	81% majority (Hindi)	21%
					19% minority	97%
Hyderabad	461[a]	256 f, 205 m	9.57 (1.19)	7–15	68% majority (Telugu)	25%
					32% minority	76%
Patna	424	253 f, 171 m	9.35 (1.15)	7–15	90% majority (Hindi)	23%
					10% minority	91%
Total	1272	702 f, 570 m	9.26 (1.09)	7–15	79% majority	23%
					21% minority	91%

[a]All children in Hyderabad performed the task in English; only the 286 children who officially attended Telugu-medium schools also performed the task in Telugu.
[b]The percentage of multilinguals is relative to the percentage of majority/minority language speakers, so, for example, of the 81% Hindi speakers in Delhi, 21% spoke another language at home in addition to Hindi.

a more remote, relatively deprived area of the country with lesser educational infrastructure.

The participating children all attended Year 4 and the vast majority (1108 children, 87%) was aged 8–10 years, but some younger (age 7, $n = 7$) and older (age 11, $n = 98$; age 12, $n = 42$, age 13–15, $n = 12$) children took part. The children were all from low socio-economic slum or non-slum neighborhoods and can be considered disadvantaged with respect to housing facilities, water facilities, and lack of parental education and educational support at home. They were all attending free government primary schools, as opposed to paid private schools.

The children's literacy was assessed in both the regionally dominant language and in English. The participating schools had either the regional language (Hindi or Telugu) or English as the official language of instruction. However, the regional language is often used in oral communication in government primary schools irrespective of the official language of instruction (Lightfoot et al., 2021). Regardless of the official school language, all schools taught both the regional language and English as a subject.

The children were tested on both decoding in reading and reading comprehension, with the exception of the children in Hyderabad, who were not tested on reading comprehension in Telugu, as the majority of children found it impossible to respond to the comprehension questions. All children attended Year 4, although it should be noted

that ASER task tests Year 2 level literacy (Pratham, 2019). According to teachers' reports, none of the participating children had any known or suspected reading disability. The participating teachers ($n = 61$) all indicated speaking multiple languages, although only 18% ($n = 11$) spoke multiple languages at home. None of them reported speaking English at home. Although the teachers' (English) language and literacy background is undeniably important, we will focus on the investigation of the children's literacy in this article. All government school teachers have to have obtained university-level teaching qualifications.

With the help of a research assistant, the children answered a questionnaire about their language background, which asked the children how many and which languages they speak with their family members. Children who spoke the regional language as their primary language in the home were considered mother tongue speakers of the majority language. Most of the children reported speaking the regional, majority language in the home (79% on average). A total of 265 (21%) of the children had a mother tongue in the home different from the regional language and were therefore considered minority language speakers. Minority language speakers do not receive mother tongue education, whereas majority language speakers do. Patna was the site with the lowest proportion of minority language-speaking children (10%), Hyderabad had the highest proportion (32%). These minority language-speaking children spoke a wide variety of languages, including languages from the state of Bihar, Bhojpuri, Haryanvi, Nepali, Pahari, Punjabi, and Rajasthani. None of the children reported speaking English in the home, so for all children English was an additional language, in which they only received input in school. We categorized children as growing up in a multilingual household when they reported being spoken to in more than one language in their home. Note that this is a different measure from overall multilingualism, which will be higher when taking into account both the home and school languages. The proportion of children growing up in multilingual households varied slightly between the three sites with the highest proportion in Hyderabad (45%) and the lowest in Patna (30%). In addition, it can be seen that 91% of the minority language-speaking children compared with 23% of the majority language-speaking children were growing up in a multilingual household.

Informed consent was obtained from all children as well as the principal or head teacher of the school. The study was conducted in accordance with the Declaration of Helsinki, the ESRC's Framework for Research Ethics (ESRC, 2010), and the guidelines of the Indian Council for Medical Research (ICMR, 2006). The protocol was approved by the Ethics Committees of the University of Cambridge (RG83665), the Jawaharlal Nehru University, and the National Institute of Mental

Health and Neurosciences. The study is part of a larger project on multilingualism and multiliteracy in primary education in India (ESRC-DfID MultiLila ES/N010345/1, see Tsimpli et al., 2019).

Methods

Literacy task. We used the ASER literacy tool (Pratham, 2019) to assess children's decoding skills. This is a concise test in which children are asked to read aloud letters, words, sentences in a paragraph and a short story, which are decoding measures of increasing difficulty. The words used in the word task are common, familiar, and relatively short. The paragraphs consist of four sentences of up to six words. These words can have multiple syllables and are based on the children's Year 1 level textbooks. The stories consist of 8–10 sentences of a higher difficulty level, using words based on the children's Year 2 level textbooks. Importantly, however, these tasks do not fully reflect general literacy, which would arguably entail the ability to decode as well as comprehend the word or text at hand (Gough & Tunmer, 1986; in fact, decoding is only one of five pillars of reading comprehension, Pressley, 2000). Therefore, novel comprehension questions were developed to assess children's understanding of the content of the story. These questions were developed in collaboration with local Indian scholars to make sure that they were culturally appropriate. An example story (shortened version) and an example question are shown below:

(Story section):
A big tree stood in a garden. It was alone and lonely. One day, a bird came and sat on it. The bird held a seed in its beak. It dropped the seed near the tree. A small plant grew there.

(Question):
How did the small plant grow near the tree?

The task was presented in English and in Hindi or Telugu. For each correctly read letter in the letter reading subtask, word in the word reading subtask, and sentence in the paragraph or story, children obtained 1 point.[3] If one word was read incorrectly in a sentence, children still obtained 1 point. If two or more words were read incorrectly, children

[3] When the letter was pronounced like it would sound in the alphabet, and the word was read as it would typically sound in English (variations in accent were accepted). Note that this scoring system differs from the classical ASER scoring, in which children are categorized in one of five reading skill levels based on the number of correctly and incorrectly read words and sentences. With our scoring method, we obtain more information, namely information about a child's performance on every ASER subtask separately rather than only a general skill level (e.g., this child is reading at "word level").

did not obtain any points for that sentence. The numbers of letters and sentences were not exactly equal across languages. There were 10 letters, 10 words, 4 sentences in a paragraph, and 8 sentences in a story in English; 10 letters, 10 words, 4 sentences in a paragraph, and 9 sentences in a story in Hindi; 8 letters, 10 words, 4 sentences in a paragraph, and 10 sentences in a story in Telugu. For comparability, the children's scores are reported as percentage correct.

Two comprehension questions followed the story in each of the languages, which children could answer orally; children obtained one point for each correctly answered question. The children were instructed to either point or read the words in the passage corresponding to the answer. In case, the children pointed or read different words in the passage not corresponding to the answer, it was scored as an incorrect response. When the children could not read the questions themselves, the research assistant read the questions aloud to them. However, the letters, words, and sentences in the ASER were read only by the children. Note that because the comprehension questions were about the story, for which decoding was separately scored, these two measures may be expected to correlate.

Analysis. Data processing and analysis were carried out using R (version 3.6.2; R Core Team, 2019). First, Pearson correlation coefficients and their significance were calculated with the package 'Hmisc' (Harrell Jr, 2019) to examine the relation between scores on the different ASER subtasks and reading comprehension (using the full data set of children from monolingual and multilingual households at all three sites). Correlations are reported with their strength (≤0.19 very weak; 0.2–0.39 weak; 0.4–0.59 moderate; 0.6–0.79 strong; ≥ 0.8 very strong, based on Evans, 1996). Linear models were used to examine the proportion of variance in comprehension explained by the four decoding measures (multiple *R*-squared).

Then, linear models were used to examine the effect of *Primary language in the home* (majority/minority) and *Household type* (monolingual/multilingual) as well as their interaction (independent variables) on performance in the ASER subtasks and the reading comprehension questions (one model per outcome variable) in the regional language and in English (dependent variables). *Site*, *Age*, and *Gender* were added as covariates; therefore, models took the form of lm(Dependent_variable ~ Home_language * Household_type + Age + Gender + Site). Treatment coding was used for all categorical variables, with speakers from monolingual (vs. multilingual) households, majority (vs. minority) language speakers, Delhi (vs. Hyderabad and Patna), and females (vs. males) as the reference levels. Models were rotated to additionally investigate the differences between Hyderabad

and Patna. The results are reported as significant when they are so at the 5% ($p < .05$)-level. Note that this means that no corrections were performed for multiple comparisons, and the p-values will have to be interpreted with caution. For this reason, a partial eta-squared was calculated for each factor with the package "heplots" (Friendly, Fox, & Monette, 2022) as an indication of effect size (with $p\eta^2$ of 0.0099, 0.0588, and 0.1379 reflecting small, medium, and large effects, respectively; Cohen, 1969; Richardson, 2011).

RESULTS

Children's performance is reported as the percentage of correct responses on each of the ASER subtasks (letters, words, sentences in a paragraph, and a short story) and the reading comprehension questions. 25 children (2%) were unable to perform any of the tasks in the regional language, and 42 children (3%) were unable to perform any of the tasks in English. The children who could not perform the tasks consisted of both minority and majority language speakers from both monolingual and multilingual households. Nevertheless, a relatively large proportion of these children were minority language speakers (children unable to read in the regional language, percentage minority language speakers: 33% in Delhi, 25% in Hyderabad, 67% in Patna; children unable to read in English, percentage minority language speakers: 0% in Delhi, 38% in Hyderabad, 39% in Patna). For example, in Patna 67% of the children unable to read in the regional language were minority language speakers, whereas only 10% of the total set of children at this site were minority language speakers.

Figure 1 presents the overall performance on the different subtasks by language at each site. The children showed relatively high performance on letter decoding in both the regional language and in

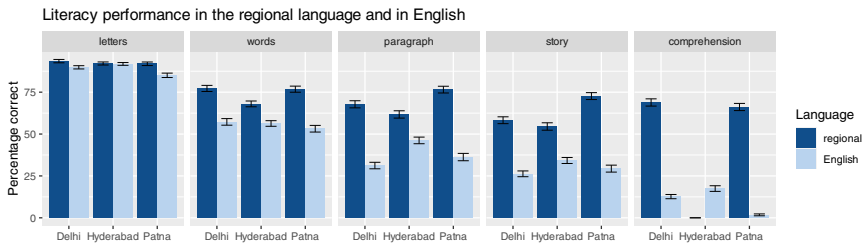

FIGURE 1. Average percentage of correct responses on the ASER subtasks and comprehension questions in the regional language (Hindi or Telugu) and in English at each site. The empty bar for regional language comprehension in Hyderabad is because these data were not collected. Error bars represent the 95% confidence interval.

English, with performance decreasing as the complexity of the task increased (i.e., lowest for story reading). On story reading in the regional language, the children scored 63% correct on average, and children answered more than two thirds of the reading comprehension questions (67%) correctly. In English, children were very accurate in decoding letters, but were considerably less accurate in decoding words and sentences, scoring less than one-third correct on story reading (30%). On the comprehension questions in English, the children only scored an average of 11% correct.

Relation Between Decoding in Reading and Reading Comprehension

The first research question asked whether decoding ability in reading is related to reading comprehension. This was examined separately in the regional language and in English.

Regional language. As can be seen in Table 2 for the regional language, all correlations between scores on the ASER subtasks were positive. There were moderate to strong correlations between the letter subtask and the other decoding subtasks, with correlation coefficients varying from 0.48 to 0.62. The correlations between the other subtasks (words, paragraph, and story) were very strong, with correlation coefficients varying from 0.83 to 0.89. These effects seem consistent between Hindi and Telugu (see Appendix A, Tables A1 and A2). These results seem to reflect the fact that the words, paragraph, and story subtasks all required lexical (word-level) decoding, whereas letter reading depends on sub-lexical (letter-level) decoding, and may thus be a distinct process. The correlations between scores on the ASER subtasks and reading comprehension were moderate to very strong, with the

TABLE 2

Correlation Coefficients for the Relation Between Performance on the ASER Subtasks and Reading Comprehension in the Regional Language (Data Collapsed Over Sites; $n = 1097$ Children for the Decoding Subtasks, $n = 811$ for Reading Comprehension)

Factor	Letters	Words	Paragraph	Story
Letters				
Words	0.62			
Paragraph	0.52	0.89		
Story	0.48	0.83	0.89	
Comprehension	0.48	0.80	0.77	0.76

Note. All correlations are significant at the $p < .001$ level.

correlation with letters being moderate and the correlations with words, paragraph, and story being strong or very strong. The proportion of variance in comprehension explained by the four decoding measures (multiple R^2) is 0.67. These results are in line with previous findings from a meta-analysis by García and Cain (2014) and indicate that decoding ability in reading is related to reading comprehension in the regional language.

English. For English, the correlations between scores on the ASER subtasks were positive as well (Table 3). There were moderate correlations between the letter subtask and the other decoding subtasks, with correlation coefficients varying from 0.35 to 0.55. The correlations between the other subtasks (words, paragraph, and story) were strong or very strong, with correlation coefficients varying from 0.70 to 0.84. Thus, like in the regional language, letter reading seems to be relatively dissimilar to word and sentence reading. In terms of decoding, there seems to be little difference between the paragraph and the story task, as these show very strong correlations with each other in both the regional language and in English. Crucially, the correlations between the ASER subtasks and reading comprehension are either very weak or weak, indicating that in English, decoding skills in reading do not ensure comprehension. This is confirmed when looking at the proportion of variance in comprehension explained by the four decoding measures (multiple R^2), which is only 0.13. Instead, oral vocabulary knowledge is likely needed to ensure comprehension, also reading comprehension (Treffers-Daller, Mukhopadhyay, Balasubramanian, Tamboli, & Tsimpli, 2022). An overview of the relation between English and the regional languages when it comes to decoding and reading comprehension can be found in Appendix A (Tables A3 andA4); although some relation in decoding is found, reading comprehension does not seem to transfer.

TABLE 3

Correlation Coefficients for the Relation Between Performance on the ASER Subtasks and Reading Comprehension in English (Data Collapsed Over Sites; n = 1272 Children for All Tasks).

Factor	Letters	Words	Paragraph	Story
Letters				
Words	0.55			
Paragraph	0.39	0.76		
Story	0.35	0.70	0.84	
Comprehension	0.17	0.32	0.36	0.30

Note. All correlations are significant at the $p < .001$ level.

Relation Between Minority Languages, Multilingualism, and Literacy Skills

Next, we addressed whether children who are being schooled in a language different from the primary home language (and thus do not receive mother tongue education) or who speak multiple languages in the home are disadvantaged in literacy skills in the regional language and/or in English. This section thus investigates both the second research question (*Does being schooled in a language different from the primary home language disadvantage children in literacy skills in the regional language and in English?*) and the third research question (*Does growing up in a multilingual household impact children's literacy skills in the regional language and in English?*).

Regional language. An overview of the results of the statistical analyses for the regional language are presented in Appendix A (Table A5). When looking at the decoding measures (letter, word, paragraph, and story reading) in Hindi or Telugu ($n = 1097$ children), minority language speakers performed less well than majority language speakers (letter reading mean 94% (SD = 18) versus 86% (SD = 29); $B = -30.61$; $t = -5.80$; $p < .001$; $p\eta^2 = 0.027$; word reading mean 75% (SD = 36) versus 68% (SD = 40); $B = -32.31$; $t = -3.38$; $p < .001$; $p\eta^2 = 0.009$; paragraph reading mean 71% (SD = 43) versus 63% (SD = 46); $B = -35.60$; $t = -3.18$; $p < .01$; $p\eta^2 = 0.007$; story reading mean 65% (SD = 43) versus 53% (SD = 44); $B = -34.40$; $t = -3.07$; $p < .01$; $p\eta^2 = 0.010$). These effects are visually presented in Figure 2. No effects of household type (monolingual/multilingual) were found (Figure 3), but interactions between speaking a minority language and household type were found in letter reading, word reading, and

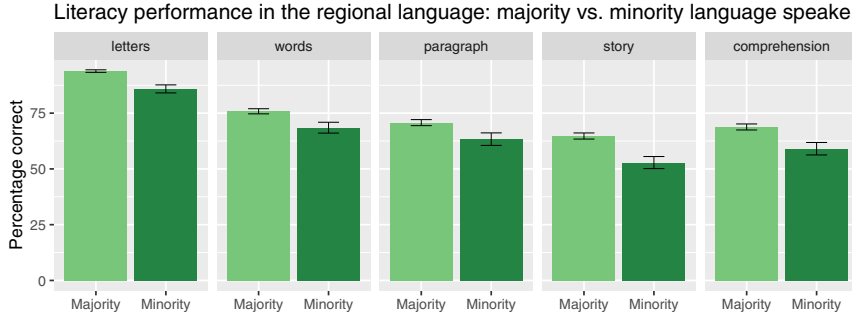

FIGURE 2. Literacy performance of majority and minority language speakers in the regional language (Hindi/Telugu). Significant differences between the two groups were found in all five subtasks. Error bars represent the 95% confidence interval.

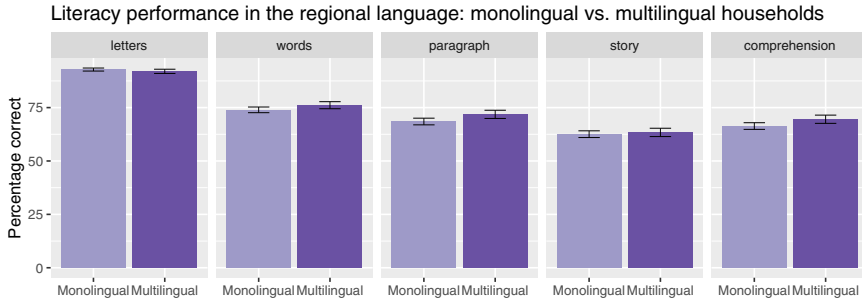

FIGURE 3. Literacy performance of children from monolingual and multilingual households in the regional language (Hindi/Telugu). No significant differences between the two groups were found. Error bars represent the 95% confidence interval.

paragraph reading (letter reading $B = 22.94$; $t = 4.00$; $p < .001$; $p\eta^2 = 0.014$; word reading $B = 24.12$; $t = 2.32$; $p < .05$; $p\eta^2 = 0.005$; paragraph reading $B = 28.01$; $t = 2.30$; $p < .05$; $p\eta^2 = 0.005$). These interactions are illustrated in Figure 4. The interactions were all positive—as opposed to the main effects of speaking a minority language—and can be taken to mean that minority language speakers growing up in a multilingual household outperformed minority language speakers growing up in a monolingual household. Note, however, that the confidence intervals in Figure 4 for the monolingual minority language-speaking group are larger than for the other groups, as there are fewer children who only speak the minority language in the home than children who speak the majority language (as their first home language or as an additional home language).

In addition, significant effects of site and gender were found for word reading, paragraph reading, and story reading (performance is

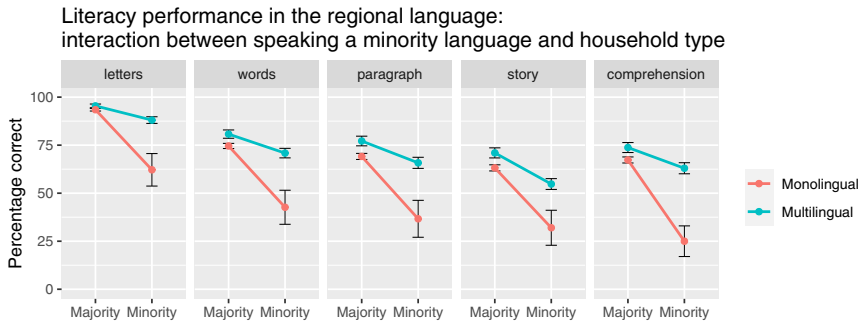

FIGURE 4. Literacy performance of majority language speakers and minority language speakers from both monolingual and multilingual households in the regional language (Hindi/Telugu). Significant interactions between speaking a minority language and household type were found in letter reading, word reading, paragraph reading, and reading comprehension. Error bars represent the 95% confidence interval.

reported in Appendix A, Table A6). On word reading, children from Delhi and Patna outperformed children from Hyderabad (resp. $B = -8.75$; $t = -2.98$; $p < .01$ and $B = -7.57$; $t = -2.65$; $p < .01$; $p\eta^2 = 0.009$) and girls outperformed boys ($B = -6.28$; $t = -2.81$; $p < .01$; $p\eta^2 = 0.007$). On paragraph reading and on story reading, children from Patna outperformed children from Delhi and Hyderabad (resp. $B = 7.79$; $t = 2.46$; $p < .05$ and $B = 13.67$; $t = 4.09$; $p < .001$; $p\eta^2 = 0.016$; story reading resp. $B = 14.03$; $t = 4.44$; $p < .001$ and $B = 16.49$; $t = 4.93$; $p < .001$; $p\eta^2 = 0.027$; see Figure 1) and girls outperformed boys (paragraph reading $B = -6.96$; $t = -2.67$; $p < .01$; $p\eta^2 = 0.006$; story reading $B = -7.56$; $t = -2.90$; $p < .01$; $p\eta^2 = 0.008$). No effects of age were found.

On reading comprehension ($n = 811$ children), which was only measured in Hindi (in both Delhi and Patna), minority language speakers also underperformed compared to majority language speakers (mean 69% (SD = 43) vs. 59% (SD = 45); $B = -42.09$; $t = -3.33$; $p < .001$; $p\eta^2 = 0.014$), speaking a majority language interacted with household type ($B = 29.91$; $t = 2.17$; $p < .05$; $p\eta^2 = 0.006$), and girls outperformed boys (see Table A6 in Appendix A; $B = -7.60$; $t = -2.49$; $p < .05$; $p\eta^2 = 0.008$).

In sum, minority language-speaking children performed less well than their majority language-speaking peers on decoding and reading comprehension in the regional language, especially when they only spoke the minority language in the home, and girls outperformed boys. In addition, regional differences were found in decoding in reading, potentially because of differences in the languages used in schools and the dominance of the regional language in the local society. These points will be discussed further in the Discussion section.

English. An overview of the results of the statistical analyses for English are presented in Table A7 in Appendix A. Of all decoding tasks in English (n = 1272 children), minority language speakers performed less well than majority language speakers only on letter reading (mean 89% (SD = 21) vs. 88% (SD = 25); $B = -10.85$; $t = -2.47$; $p < .05$; $p\eta^2 = 0.014$). This effect is visually presented in Figure 5. On all four decoding subtasks, children from multilingual households outperformed children from monolingual households (letter reading mean 91% (SD = 20) vs. 88% (SD = 23); $B = 6.19$; $t = 3.85$; $p < .001$; $p\eta^2 = 0.015$; word reading mean 61% (SD = 36) vs. 52% (SD = 39); $B = 11.58$; $t = 4.05$; $p < .001$; $p\eta^2 = 0.015$; paragraph reading mean 45% (SD = 43) vs. 34% (SD = 42); $B = 13.63$; $t = 4.36$; $p < .001$; $p\eta^2 = 0.016$; story reading mean 36% (SD = 40) vs. 27% (SD = 38); $B = 13.70$; $t = 4.78$; $p < .001$; $p\eta^2 = 0.021$). These effects are visually presented in Figure 6. No interactions between speaking a minority

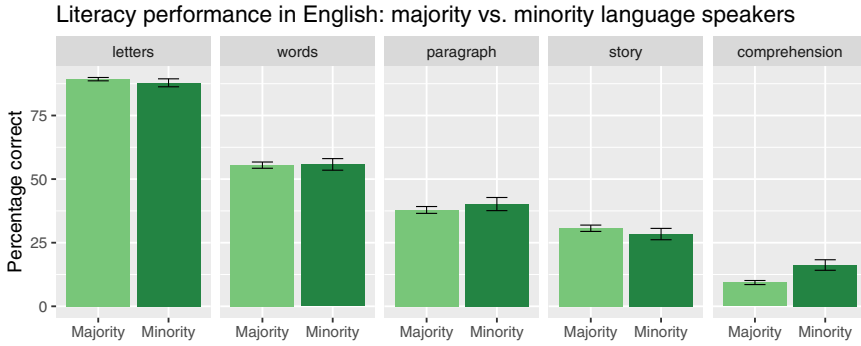

FIGURE 5. Literacy performance of majority and minority language speakers in English. A significant difference between the two groups was found in letter reading. Error bars represent the 95% confidence interval.

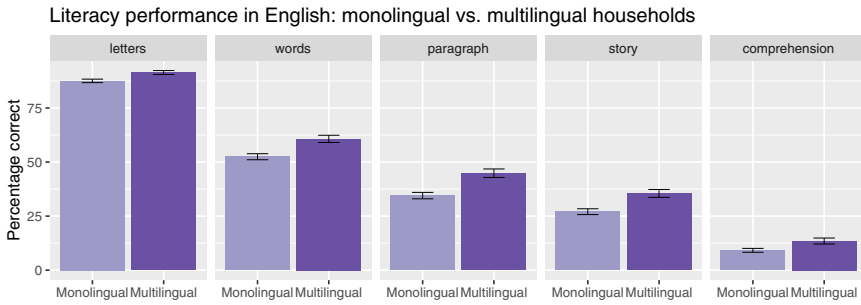

FIGURE 6. Literacy performance of children from monolingual and multilingual households in English. Significant differences between the two groups were found in letter reading, word reading, paragraph reading, and story reading. Error bars represent the 95% confidence interval.

language and household type were found (Figure 7), indicating that growing up in a multilingual household provides an advantage for decoding in English regardless of whether the child speaks a minority language or the majority language as their mother tongue.

In addition, significant effects of site were found (performance is reported in Table A8 in Appendix A), with children from Delhi and Hyderabad outperforming children from Patna on letter reading (resp. $B = -5.22$; $t = -3.32$; $p < .001$ and $B = -7.43$; $t = -4.97$; $p < .001$; $p\eta^2 = 0.020$), and children from Hyderabad outperforming children from Delhi and Patna on paragraph reading and story reading (paragraph reading resp. $B = 13.68$; $t = 4.51$; $p < .001$ and $B = 9.86$; $t = 3.39$; $p < .001$; $p\eta^2 = 0.017$; story reading resp. $B = 8.10$; $t = 2.91$; $p < .01$ and $B = 5.57$; $t = 2.09$; $p < .05$; $p\eta^2 = 0.007$; see Figure 1; Table A8). No effects of gender or age were found.

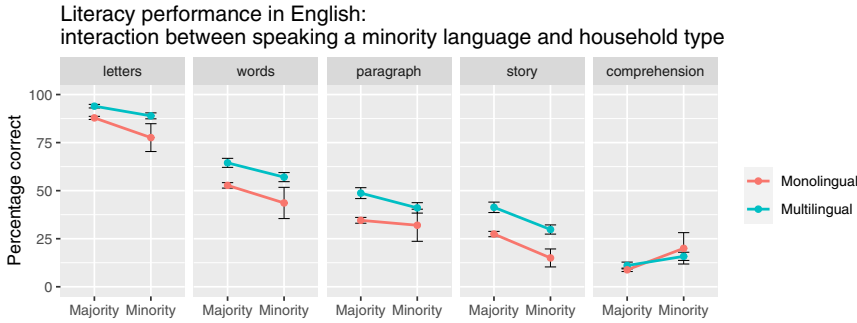

FIGURE 7. Literacy performance of majority language speakers and minority language speakers from both monolingual and multilingual households in English. Error bars represent the 95% confidence interval.

On reading comprehension ($n = 1272$ children), no significant effects of household type or speaking a minority language were found, but children from Hyderabad outperformed children from Delhi ($B = 4.15$; $t = 2.16$; $p < .05$), who in turn outperformed children from Patna ($B = -11.14$; $t = -5.76$; $p < .001$; $p\eta^2 = 0.055$). In addition, girls outperformed boys ($B = -3.73$; $t = -2.48$; $p < .05$; $p\eta^2 = 0.005$; see Table A8).

In sum, children from multilingual households outperformed children from monolingual households on decoding in English. Disadvantages of speaking a minority language were only found in letter decoding. Like in regional language decoding, regional differences were found in decoding and reading comprehension in English.

DISCUSSION

In this study, we investigated the effect of speaking a minority language in the home (and thus not receiving mother tongue education) and multilingualism in the household on literacy skills in a large group of Indian primary school children from low socio-economic backgrounds. The data show that around two thirds of the children were able to read and comprehend a short story in their regional language, but nearly all children had difficulties reading in English. Note that the ASER is developed for Year 2, whereas the participating children were in Year 4. For reading in the regional language, these results are nevertheless more encouraging than those of previous investigations, in which more than half of Year 5 children could not read a Year 2 text fluently (Pratham, 2019). It is alarming, however, that a small percentage of the children in our study (2–3%) could not

read any letters in a given language. Note that the current data set was collected before the COVID-19 pandemic; children's reading difficulties are likely to have increased during the pandemic due to school closures (UNESCO, 2020/2021). There is little support for such children who are far behind, and little support for non-typically developing children in general in the Indian school system (Sharma, Moore, & Sonawane, 2009; Singal, 2009).

Three research questions were investigated. The first question asked whether decoding ability in reading is related to reading comprehension in the regional language and in English. The results show strong correlations between decoding (words, paragraphs, stories) and reading comprehension in the regional language (contra Menon et al., 2017), but weak correlations between decoding and reading comprehension in English. Note that word, paragraph, and story decoding were strongly correlated to each other in both languages, even though they were of different difficulty levels. Letter reading was not always strongly correlated with the other measures, especially in English, possibly because of the opaqueness of English orthography (in line with the contrast between alphabetic and alpha-syllabary writing systems as presented in Section "Literacy Development: Issues and Influences") and because of the contrast between a sub-lexical and a lexical task. Some additional reasons why the tested children struggled with English literacy could be (a) the input they receive and (b) their (disadvantaged) background (in line with Shenoy et al., 2020, 2023). The input they receive in school is predominantly in the regional language or mixed between the regional language and English or minority languages, even in English-language classes (see Lightfoot et al., 2021 on language use in the same classrooms in Delhi and Hyderabad as used in the current study). With regard to their background, not only are these children not exposed to English at home (or outside of school in general), they likely also receive little to no literacy support in the home environment (Tsimpli et al., 2019). Since the vocabulary knowledge required for reading comprehension is specific to the language of the text (Cobo-Lewis et al., 2002; Nakamoto et al., 2008), children may have learned to decode English letters and words at school (through rote learning; Smith et al., 2005), but have not acquired word meaning—Oral vocabulary knowledge is a key predictor of reading comprehension (Treffers-Daller et al., 2022; in line with the Simple View of Reading by Gough & Tunmer, 1986). In addition, it is possible that children have acquired word meaning but are unable to synthesize sentence meaning when words that they partially know are presented in a text. Importantly, not understanding English texts could aggravate these children's disadvantage and hinder their

academic development, as many children will be unprepared to be taught in English (Nakamura et al., 2018) and the level of English presented in school textbooks is far higher than the children's knowledge (Treffers-Daller et al., 2022). In contrast, most children speak the regional language in the home and are in contact with it at school, making it likely that they have acquired the appropriate vocabulary through oral language use; these oral skills will then contribute to reading comprehension (e.g., Gough & Tunmer, 1986).

The second research question asked whether children who are being schooled in a language different from the primary home language, thus not receiving mother tongue education, are disadvantaged in literacy skills in the regional language and in English. In the regional language this was indeed found to be the case, with majority language speakers outperforming minority language speakers on decoding and reading comprehension. In addition, interactions between speaking a minority language and growing up in a multilingual household suggest that children who are only exposed to the minority language in the home underperform compared to their peers growing up in multilingual households. This in turn can be explained by most (i.e. 84%) of these minority language-speaking children being exposed to the majority language as a second or third language in the home. So, even though these children speak a minority language as their mother tongue, they are orally exposed to the majority language in the home as well (which would help with decoding in the Simple View of Reading, Gough & Tunmer, 1986). Together, these findings indicate that minority language-speaking children, especially those who are not exposed to the majority language in their home environment, lag behind in their reading abilities. This is in line with previous research showing lower learning outcomes for children who are not educated in a familiar language (Babayiğit & Shapiro, 2020; Ball, 2011; Ouane & Glanz, 2010; Romaine, 2013). Thus, minority language-speaking children from low socio-economic backgrounds face further disadvantages when educated in a language they do not speak in the home (in line with Mohanty, 2010, 2019). In English, majority language-speaking children also outperformed minority language-speaking peers, but only on letter decoding. Since none of the children spoke English in the home, this may be a consequence of the learning process, in which minority language-speaking children can experience difficulties understanding the subject matter (in this case English) being taught primarily in the regional, majority language (Bhattacharya, 2013; Endow, 2018). These findings further emphasize the need for more inclusive and multilingual learning practices in Indian schools, in which both the regional language and the mother tongues of minority speakers are supported (Leikin et al., 2009;

Mohanty, 2019; Nag, 2017). First steps in this direction are already being taken in some Indian regions (Odisha; Mohanty, 2015; Panda, 2020), but the development of inclusive and multilingual learning practices is necessary across all regions to provide equal opportunities to all children in India, especially those coming from low socio-economic backgrounds.

The third and final research question asked whether growing up in a multilingual household impacts children's literacy skills in the regional language and in English. We found no negative effects of multilingualism on literacy in the regional language or in English. This is good news for a multilingual society, such as India, in which multilingualism is the norm. Although multilinguals generally have a less extensive vocabulary compared to monolinguals in one specific language (Bialystok, 2009), this did not negatively affect their literacy performance, potentially due to bilinguals' advantages in phonological awareness (e.g., Canbay, 2011; Souza & Leite, 2014) or the familiarity with more than one writing system in school (alphabetic, English and alpha-syllabic, Hindi/Telugu) for children from monolingual and multilingual households alike. In contrast, growing up in a multilingual household was found to have a positive effect on decoding in the regional language for children who spoke a minority language as their mother tongue, as the additional language spoken was in many cases the regional language. However, multilingualism in the home was also found to have a positive effect on decoding skills in English, which cannot be explained by it being spoken in the home as this was not the case for any of the participating children. Thus, the explanation in this case seems to be multilingualism itself (due to improved phonological awareness, for instance). Based on these results, it seems that growing up in a multilingual household provides advantages regardless of whether this includes the majority language. This result is in line with studies indicating that it is easier for multilinguals to learn additional languages (see Hirosh & Degani, 2018, for a review).

Importantly, our results indicate that nearly all children had difficulties reading English. As elaborated in Section "Literacy Development: Issues and Influences", the type of orthography could have an influence on literacy development (Florit & Cain, 2011; Joshi, Padakannaya, & Nishanimath, 2010). It is possible that the children found reading alphabetic English (which has a deep orthography) to be more difficult than reading relatively transparent alpha-syllabic Hindi or Telugu (Nag, 2007, 2022); the regional language will also be more familiar to the children (Skutnabb-Kangas, 2000). Importantly, this is specific to the Indian context and thus the order of learning does not follow the prototypical order of alphabetic English first, then

alpha-syllabaries (Joshi et al., 2017; Liberman et al., 1974; Nag, 2007, 2011; Nakamura et al., 2014; Tiwari et al., 2011) but is in line with other findings from India on Kannada/Telugu versus English (e.g., Nakamura et al., 2018). Another reason could be the method employed in teaching English. Unlike the phonics method employed to teach English in the western context, the alphabet-spelling method (Gupta, 2014) is used to teach English in India where only the letter names and word spellings are taught bypassing the sounds of the alphabet resulting in poor decoding in English (Shenoy et al., 2020).

We now briefly return to an important consideration addressed in the introduction, namely our definition of multilingualism, which was restricted to language use in the home. Our expectations about the sample of children were confirmed, namely all children were exposed to English at school and since none of the children spoke English in the home, they are in a sense all multilingual, although to different extents. We therefore emphasize that this means that the children from a "monolingual household" are not truly monolingual and should in no way be taken as a monolingual baseline. Nevertheless, the fact that influences of multilingualism in the home were found in this research emphasizes the need to take home languages into account as a separate factor when investigating literacy development in highly multilingual societies.

Finally, several effects of demographic factors were found, which were not explicitly stated in the research questions but are nevertheless worthy of discussion. Firstly, effects of site were found, with children from Hyderabad generally performing less well than their peers from Delhi and Patna in regional language literacy, but better in English literacy. Conversely, children from Patna outperformed their peers on regional language literacy, but underperformed in English. There thus seems to be a trade-off between literacy in the different languages, possibly due to differences in the languages used in schools and the dominance of the languages in the local society. Indeed, primary schools in Hyderabad seem to use more English than schools in Delhi (Lightfoot et al., 2021). English may be perceived as important because of its international prominence and employment prospects (Dearden, 2014). This is more so, when the comparison with English involves Telugu and Hindi which are languages associated with differences in social and political power; Hindi is the national language and the language of the capital city of India whereas Telugu is a regional language of Andhra Pradesh and Telangana, spoken by much fewer people than Hindi (Sirsa, 2014). In addition, fewer children in Hyderabad speak the regional language, Telugu, at home than children in Delhi and Patna speak Hindi at home, indicating that the regional language is less prominent in Hyderabad. Patna is a much more rural site

(compared with the urban cities of Delhi and Hyderabad), which might explain the lesser focus on English and stronger focus on the regional language. The research team observed that children in Patna used Hindi in and out of school whereas children in Delhi spoke more English in school, partly explaining their performance in the different languages. Secondly, effects of gender were found, with girls outperforming boys on decoding in the regional language and reading comprehension in both the regional language and in English. Although it is a common perception that boys would be favored in the school system, this does not seem to be the case in some countries in general and in India specifically, where girls, especially from poorer households, have previously been found to outperform boys in language subjects (Natta, Desai, & Vanneman, 2017; Shenoy et al., 2023; UNICEF, 2012). Reasons for this could be that girls who attend school are from households with more resources, as households with fewer resources do not send their girls to school, that girls learn faster and study more diligently than boys, and that boys have more work opportunities and demands outside of school (Natta et al., 2017). Notably, no effects of age were found, even though the participating children had a wide age range (7–15). This may be because they all attended the same year in school (Year 4)—possibly because they had not continuously attended school throughout their childhood—thus, all were approximately at the same level of school skills.

To conclude, our results show that children from a monolingual household who do not receive mother tongue education underperform compared with majority language speakers when reading in the majority language. This shows that support of minority language speakers from disadvantaged backgrounds with minimal literacy support from home who do not receive mother tongue education is paramount when they are expected to develop literacy skills in the regional language. To this end, we encourage the use of multilingual practices in education in general and for vocabulary building specifically, as vocabulary is a prerequisite for reading comprehension. Performance on literacy in English as an additional language improved when the child spoke multiple languages in the home, but overall performance left much to be desired, especially in reading comprehension but also in decoding. It is thus essential for these children to be better provided in terms of educational resources and richer input in order to become proficient readers of English. We encourage educators to focus on teaching reading for comprehension rather than piecemeal memorization or reading only for decoding, as reading comprehension forms the basis for children's learning in school and everyday life.

ACKNOWLEDGMENTS

We would like to thank all the research assistants, children, and teachers for their participation and support. We are very grateful to Debanjan Chakrabarti, Ganesh Devy, Dhir Jhingran, Vasanta Duggirala, Amy Lightfoot, Rama Matthew, Ajit Mohanty, Bapi Raju, and Pallawi Sinha for their advice and support throughout the project.

FUNDING STATEMENT

This data set is part of the MultiLila project. The MultiLila project (2016–2020) was led by the University of Cambridge in collaboration with the University of Reading, Jawaharlal Nehru University, The English and Foreign Languages University, and the National Institute of Mental Health and Neurosciences (NIMHANS). This project was funded by the Department for International Development and the Economic and Social Research Council (ESRC-DFID, Grant Ref: ES/N010345/1).

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare that are relevant to the content of this article.

ETHICS STATEMENT

The protocol was approved by the Ethics Committee of the University of Cambridge (RG83665) and the Ethics Committees of Jawaharlal Nehru University and the National Institute of Mental Health and Neurosciences (NIMHANS).

THE AUTHORS

Dr. Vogelzang is a postdoctoral research associate at the Department of Theoretical and Applied Linguistics, University of Cambridge, UK. Her work combines psycholinguistic and statistical methods to investigate syntactic processing and the acquisition of literacy.

Prof. Tsimpli is Chair of English and Applied Linguistics at the University of Cambridge, UK. She works on language development in the first and second language in children and adults, language impairment, attrition, bilingualism, language processing and the interaction between language, cognitive abilities, and print exposure.

Dr. Balasubramanian is a postdoctoral research associate at the Department of Theoretical and Applied Linguistics, University of Cambridge, UK. Her work involves the investigation of Language Disorders, Language Development, and Speech and Language Processing.

Prof. Panda works the Zakir Husain Centre for Educational Studies, School of Social Sciences, at the Jawaharlal Nehru University. Her Areas of specialization include Numeracy Practices in tribal communities in India and among the Immigrant Communities in UK, Social Theories of Learning including Cultural Historical Activity Theory, and Multilingual Education.

Prof. Alladi is Professor of Neurology, National Institute of Mental Health and Neurosciences, Bangalore, India. She established one of the first memory clinics in India. Her work also focuses on investigating the influence of life-course experiences that include bilingualism, education and social engagement on cognitive resilience.

Dr. Reddy was a research associate at the National Institute of Mental Health and Neurosciences, Bangalore, India. Together with prof. Alladi, she coordinated the Hyderabad team for the MultiLila project.

Professor Mukhopadhyay is Head of the Department of Training and Development, School of English Language Education at The English and Foreign Languages University, Hyderabad, India. Her areas of research interests are bi/multilingual education, second language acquisition, second language writing and language assessment and evaluation.

Prof. Treffers-Daller is Professor of Multilingualism at the University of Reading. She is interested in a wide range of areas in the field of Multilingualism and Second Language Acquisition. She has recently worked on the relationship between cognitive control and code-switching, and on measuring language dominance in bilinguals/heritage speakers.

Prof. Marinis is Professor of Multilingualism at the University of Konstanz and at the University of Reading. His research focuses on language acquisition and processing across populations of typically and atypically developing learners and aims to uncover the nature of language processing in typical and atypical language development.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: Bradford.

Alcott, B., & Rose, P. (2017). Learning in India's primary schools: How do disparities widen across the grades? *International Journal of Educational Development*, *56*, 42–51.

Alidou, H., Boly, A., Brock-Utne, B., & Satina, Y. (2006). *Optimizing learning and education in Africa – The language factor*. Paris, France: Association for the Development of Education in Africa.

Babayiğit, S., & Shapiro, L. (2020). Component skills that underpin listening comprehension and reading comprehension in learners with English as first and

additional language. *Journal of Research in Reading*, *43*(1), 78–97. https://doi.org/10.1111/1467-9817.12291
Baker, C. (2011). *Foundations of bilingual education and bilingualism* (Vol. *79*). New York, NY: Multilingual matters.
Ball, J. (2011). *Enhancing learning of children from diverse language backgrounds*. Paper commissioned for UNESCO. Retrieved from: http://unesdoc.unesco.org/images/0021/002122/212270e.pdf
Bernhardt, E. (2003). Challenges to reading research from a multilingual world. *Reading Research Quarterly*, *38*(1), 112–117.
Bhattacharya, U. (2013). Mediating inequalities: Exploring English-medium instruction in a suburban Indian village school. *Current Issues in Language Planning*, *14*(1), 164–184. https://doi.org/10.1080/14664208.2013.791236
Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511605963
Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11.
Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied PsychoLinguistics*, *24*, 27–44. https://doi.org/10.1017/S014271640300002X
Blackenbecker, C. (2020). Designing for complexity in mother tongue or first language (L1)-based multilingual education programs. *Global Education Review*, *7*(1), 41–56.
Blair, A., Haneda, M., & Bose, F. N. (2018). Reimagining English-medium instructional settings as sites of multilingual and multimodal meaning making. *TESOL Quarterly*, *52*, 516–539. https://doi.org/10.1002/tesq.449
Canbay, O. (2011). Comparing the phonological awareness of bilingual and monolingual pre-school children. *Procedia-Social and Behavioral Sciences*, *15*, 976–980. https://doi.org/10.1016/j.sbspro.2011.03.224
Clarke, S. (2001). *Unlocking formative assessment: Practical strategies for enhancing pupils' learning in the primary classroom*. Abingdon: Hodder Education.
Cobo-Lewis, A. B., Pearson, P. Z., Eilers, R. E., & Umbel, V. C. (2002). Effects of bilingualism and bilingual education on oral and written Spanish skills: A multifactor study of standardized test out-comes. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 98–117). Clevedon, UK: Multilingual Matters.
Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149.
Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59–86.
Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 222–251.
Cummins, J. (2009). Multilingualism in the English-language classroom: Pedagogical considerations. *TESOL Quarterly*, *43*(2), 317–321.
Dearden, J. (2014). English as a medium of instruction – A growing global phenomenon. British council. Retrieved from: https://www.britishcouncil.org/sites/default/files/e484_emi_-_cover_option_3_final_web.pdf
Dressler, C., & Kamil, M. L. (2006). First- and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report*

of the National Literacy Panel on language-minority children and youth (pp. 197–238). Mahwah, NJ: Lawrence Erlbaum Associates.
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428
Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia*, *52*, 189–204. https://doi.org/10.1007/s11881-002-0012-y
Dworin, J. E. (2003). Examining children's biliteracy in the classroom. In A. I. Willis, G. E. García, R. Barrera, & V. J. Harris (Eds.), *Multicultural issues in literacy and practice* (pp. 29–48). Mahwah, NJ: Lawrence Erlbaum.
Endow, T. (2018). Inferior outcomes: Learning in low-cost English-medium private schools—A survey in Delhi and National Capital Region. *Indian Journal of Human Development*, *12*(1), 1–19. https://doi.org/10.1177/0973703018779725
ESRC. (2010). ESRC Framework for Research Ethics (FRE) 2010. Retrieved from: https://esrc.ukri.org/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2010/
Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/ColePublishing.
Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*, 553–576.
Friendly, M., Fox, J., & Monette, G. (2022). *heplots: Visualizing tests in multivariate linear models*. R package version 1.4-2. Retrieved from: https://CRAN.R-project.org/package=heplots
García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74–111.
Goriot, C., Denessen, E., Bakker, J., & Droop, M. (2016). Benefits of being bilingual? The relationship between pupils' perceptions of teachers' appreciation of their home language and executive functioning. *International Journal of Bilingualism*, *20*(6), 700–713.
Gort, M. (2006). Strategic codeswitching, interliteracy, and other phenomena of emergent bilingual writing: Lessons from first grade dual language classrooms. *Journal of Early Childhood Literacy*, *6*(3), 323–354. https://doi.org/10.1177/1468798406069796
Gort, M. (2012). Code-switching patterns in the writing-related talk of young emergent bilinguals. *Journal of Literacy Research*, *44*(1), 45–75. https://doi.org/10.1177/1086296X11431626
Gort, M. (2019). Developing bilingualism and biliteracy in early and middle childhood. *Language Arts*, *96*(4), 229–243.
Goswami, U. (2005). Synthetic phonics and learning to read: A cross-language perspective. *Educational Psychology in Practice*, *21*(4), 273–282. https://doi.org/10.1080/02667360500344823
Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://doi.org/10.1177/074193258600700104
Gupta, R. (2014). Change in teaching practices: Case of phonics instruction in India. *Procedia-Social and Behavioral Sciences*, *116*, 3911–3915. https://doi.org/10.1016/j.sbspro.2014.01.865
Harrell, F. E., Jr., with contributions from Charles Dupont and many others. (2019). *Hmisc: Harrell Miscellaneous*. R package version 4.3-0. Retrieved from: https://CRAN.R-project.org/package=Hmisc

Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review*, *25*, 892–916. https://doi.org/10.3758/s13423-017-1315-7

ICMR. (2006). *Ethical guidelines for biomedical research on human participants*. Ethical Review Procedures 11. Retrieved from: http://icmr.nic.in/ethical_guidelines.pdf

Jhingran, D. (2019). *Early literacy and multilingual education in South Asia*. Kathmandu: United Nations Children's Fund Regional Office for South Asia.

Jiménez, R. T. (1997). The strategic reading abilities and potential of five low-literacy latina/o readers in middle school. *Reading Research Quarterly*, *32*(3), 224–243. https://doi.org/10.1598/rrq.32.3.1

Joshi, R. M., Nakamura, P. R., & Singh, N. C. (2017). Reading research and practice: Indian perspective. *New Directions for Child and Adolescent Development*, *2017* (158), 43–53. https://doi.org/10.1002/cad.20222

Joshi, R. M., Padakannaya, P., & Nishanimath, S. (2010). Dyslexia and hyperlexia in bilinguals. *Dyslexia*, *16*, 99–118. https://doi.org/10.1002/dys.402

Kašćelan, D., Prévost, P., Serratrice, L., Tuller, L., Unsworth, S., & De Cat, C. (2022). A review of questionnaires quantifying bilingual experience in children: Do they document the same constructs? *Bilingualism: Language and Cognition*, *25*(1), 29–41. https://doi.org/10.1017/S1366728921000390

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, *37*(4), 374–382.

Koda, K. (2008). Impacts of prior literacy experience on second language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second language literacy development* (pp. 68–96). New York, NY: Routledge.

Kosmidis, M. H., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence on working memory. *Archives of Clinical Neuropsychology*, *26*(7), 575–582.

Leikin, M., Schwartz, M., & Share, D. (2009). General and specific benefits of biliterate bilingualism: a Russian–Hebrew study of beginning literacy. *Written Language and Literacy*, *8*(2), 103–130.

Liberman, I., Shankweiler, D., William-Fisher, F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, *18*(2), 201–212. https://doi.org/10.1016/0022-0965(74)90101-5

Lightfoot, A., Balasubramanian, A., Tsimpli, I. M., Mukhopadhyay, L., & Treffers-Daller, J. (2021). Measuring the multilingual reality: Lessons from classrooms in Delhi and Hyderabad. *International Journal of Bilingual Education and Bilingualism*, *25*(6), 2208–2228. https://doi.org/10.1080/13670050.2021.1899123

Lyle, S. (2008). Dialogic teaching: Discussing theoretical contexts and reviewing evidence from classroom practice. *Language and Education*, *22*, 222–240.

Menon, S., Krishnamurthy, R., Sajitha, S., Apte, N., Basargekar, A., Subramaniam, S., … Modugala, M. (2017). *Literacy research in Indian languages (LiRiL): Report of a three-year longitudinal study on early reading and writing in Marathi and Kannada*. Bangalore, New Delhi: Azim Premji University, Tata Trusts.

Ministry of Human Resource Development. (2020). National Education Policy 2020. Retrieved from: https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

Mishra, R., & Stainthorp, R. (2007). The relationship between phonological awareness and word reading accuracy in Oriya and English: A study of Oriya-speaking fifth-graders. *Journal of Research in Reading*, *30*(1), 23–37.

Mohanty, A. K. (2010). Languages, inequality and marginalisation: Implications of the double divide in Indian multilingualism. *International Journal of the Sociology of Language*, *205*, 131–154. https://doi.org/10.1515/ijsl.2010.042
Mohanty, A. K. (2019). *The multilingual reality: Living with languages*. Bristol: Multilingual Matters.
Mohanty, S. K. (2015). *Mother-tongue-based multilingual education (MTB MLE) in the Indian state of Odisha: Issues, challenges and impact*. 11th Language and Development Conference. New Delhi, India.
Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading*, *30*(1), 7–22. https://doi.org/10.1111/j.1467-9817.2006.00329.x
Nag, S. (2011). The akshara languages: What do they tell us about children's literacy learning. In R. Mishra & N. Srinivasan (Eds.), *Language-cognition: State of the art* (pp. 291–310). Germany: Lincom Publishers.
Nag, S. (2017). Theoretical assumptions regarding the mind-culture-language relationship underlying MLE models in India and their impact on resulting practices. In H. Coleman (Ed.), *Multilingualisms and development: Selected proceedings of the 11th language and development conference, New Delhi* (pp. 133–150). British Council: New Delhi.
Nag, S. (2022). Reading the akshara writing system. In J. Snowling, C. Hulme, & K. Nation (Eds.), *The science of reading: A handbook* (2nd ed., pp. 372–389). Wiley-Blackwell. https://doi.org/10.1002/9781119705116
Nag, S., & Snowling, M. J. (2012). Reading in an alphasyllabary: Implications for a language universal theory of learning to read. *Scientific Studies of Reading*, *16*(5), 404–423. https://doi.org/10.1080/10888438.2011.576352
Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading*, *12*(4), 351–371. https://doi.org/10.1080/10888430802378526
Nakamura, P., Molotsky, A., Castro Zarzur, R., Ranjit, V., Haddad, Y., & De Hoop, T. (2023). *Language of instruction in schools in LMICs: A systematic review, CEDIL Evidence Brief 11*. London and Oxford: Centre of Excellence for Development Impact and Learning. https://doi.org/10.51744/CEB11
Nakamura, P. R. (2023). When transfer transfers: Applying cross-linguistic reading transfer theory to language of instruction policies in low-and middle-income countries. In D. Zhang & R. T. Miller (Eds.), *Crossing boundaries in researching, understanding, and improving language education: Essays in honor of G. Richard Tucker* (pp. 341–359). Cham: Springer International Publishing.
Nakamura, P. R., De Hoop, T., & Holla, C. U. (2018). Language and the learning crisis: Evidence of transfer threshold mechanisms in multilingual reading in South India. *The Journal of Development Studies*, *55*(11), 2287–2305. https://doi.org/10.1080/00220388.2018.1493192
Nakamura, P. R., Koda, K., & Joshi, R. M. (2014). Biliteracy acquisition in Kannada and English: A developmental study. *Writing Systems Research*, *6*(1), 132–147. https://doi.org/10.1080/17586801.2013.855620
Natta, S., Desai, S., & Vanneman, R. (2017). Gender and educational attainment in India: Girls outperforming boys. In *Population Association of America 2017 Annual Meeting*. Population Association of America.
Nishanimut, S. P., Johnston, R. S., Joshi, R. M., Thomas, P. J., & Padakannaya, P. (2013). Effect of synthetic phonics instruction on literacy skills in an ESL setting. *Learning and Individual Differences*, *27*, 47–53.

Oh, M. H., Mancilla-Martinez, J., & Hwang, J. K. (2023). Revisiting the traditional conceptualizations of vocabulary knowledge as predictors of dual language learners' English reading achievement in a new destination state. *Applied Psycho-Linguistics*, *44*(1), 51–75. https://doi.org/10.1017/S0142716422000479

Ouane, A., & Glanz, C. (2010). *Why and how Africa should invest in African languages and multilingual education. An evidence- and practice-based policy advocacy brief*. Hamburg: UNESCO Institute for Lifelong Learning.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Panda, A. K. (2020). Impact of mother tongue based (MTB) education in tribal dominated schools of Odisha. *IOSR Journal of Humanities and Social Science*, *25* (2), 13–18.

Panda, M. (2019). *A sociolinguistic survey of Barwani district, Madhya Pradesh*. New Delhi: National Multilingual Education Resource Consortium (NMRC), JNU & Room to read.

Patel, P., Chatterjee Singh, N., & Torppa, M. (2022). Understanding the role of cross-language transfer of phonological awareness in emergent Hindi–English biliteracy acquisition. *Reading and Writing*, *37*, 887–920. https://doi.org/10.1007/s11145-022-10253-x

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2022). Assessing the effectiveness of a game-based phonics intervention for first and second grade English language learners in India: A randomized controlled trial. *Journal of Computer Assisted Learning*, *38*(1), 76–89.

Pluck, G. (2019). Lexical reading ability predicts academic achievement at university level. *Cognition Brain Behavior An Interdisciplinary Journal*, *22*(3), 175–196. https://doi.org/10.24193/cbb.2018.22.12

Prakash, P., & Joshi, R. M. (1995). Orthography and reading in Kannada: A Dravidian language. In I. Taylor & D. R. Olson (Eds.), *Scripts and Literacy. Neuropsychology and Cognition* (Vol. *7*). Dordrecht: Springer. https://doi.org/10.1007/978-94-011-1162-1_7

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

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. *3*, pp. 545–560). Hillsdale, NJ: Erlbaum.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from: https://www.R-project.org/

Rabiner, D. L., Godwin, J., & Dodge, K. A. (2016). Predicting academic achievement and attainment: The contribution of early academic skills, attention difficulties, and social competence. *School Psychology Review*, *45*(2), 250–267. https://doi.org/10.17105/SPR45-2.250-267

Rao, C., Vaid, J., Srinivasan, N., & Chen, H. C. (2011). Orthographic characteristics speed Hindi word naming but slow Urdu naming: Evidence from Hindi/Urdu biliterates. *Reading and Writing*, *24*, 679–695.

Reddy, P. P., & Koda, K. (2013). Orthographic constraints on phonological awareness in biliteracy development. *Writing Systems Research*, *5*(1), 110–130.

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Romaine, S. (2013). Keeping the promise of the millennium development goals: Why language matters. *Applied Linguistics Review*, *4*(1), 1–21.
Samuel, S., Roehr-Brackin, K., Pak, H., & Kim, H. (2018). Cultural effects rather than a bilingual advantage in cognition: A review and an empirical study. *Cognitive Science*, *42*(7), 2313–2341. https://doi.org/10.1111/cogs.12672
Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy* (pp. 97–110). NY: Guilford Press.
Share, D. L., & Daniels, P. T. (2016). Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz and Fowler (2014). *Writing Systems Research*, *8*(1), 17–31. https://doi.org/10.1080/17586801.2015.1016395
Sharma, U., Moore, D., & Sonawane, S. (2009). Attitudes and concerns of preservice teachers regarding inclusion of students with disabilities into regular schools in Pune, India. *Asia-Pacific Journal of Teacher Education*, *37*(3), 319–331. https://doi.org/10.1080/13598660903050328
Shenoy, S., Wagner, R. K., Overton, K., & Rao, N. M. (2023). Impact of predictor variables on L2 English reading acquisition for grades 3 and 5. *Reading Psychology*, *44*(3), 242–269.
Shenoy, S., Wagner, R. K., & Rao, N. M. (2020). Factors that influence reading acquisition in L2 English for students in Bangalore, India. *Reading and Writing*, *33*, 1809–1838.
Sierens, S., Slembrouck, S., Van Gorp, K., Agirdag, O., & Van Avermaet, P. (2019). Linguistic interdependence of receptive vocabulary skills in emergent bilingual preschool children: Exploring a factor-dependent approach. *Applied PsychoLinguistics*, *40*(5), 1269–1297.
Simons, G. F., & Fennig, C. D. (2018). *Ethnologue: Languages of the world, twenty*. Dallas: SIL International. Retrieved from: www.ethnologue.com
Singal, N. (2009). *Education of children with disabilities in India*. Paper commissioned for the EFA Global Monitoring Report 2010, Reaching the marginalized.
Sirsa, H. (2014). *First language and sociolinguistic differences in the sound patterns of Indian English*. Doctoral dissertation, University of Oregon.
Skutnabb-Kangas, T. (2000). *Linguistic genocide in education–or worldwide diversity and human rights?* New York: Routledge.
Smith, F., Hardman, F., & Tooley, J. (2005). Classroom interaction in private schools serving low-income families in Hyderabad, India. *International Education Journal*, *6*(5), 607–618.
Souza, L. B., & Leite, A. G. (2014). Profile of phonological awareness in bilingual and monolingual children. *CoDAS*, *26*(1), 61–67.
Tiwari, S., Nair, R., & Krishnan, G. (2011). A preliminary investigation of akshara knowledge in the Malayalam alphasyllabary: Extension of Nag's (2007) study. *Writing Systems Research*, *3*, 145–151. https://doi.org/10.1093/wsr/wsr013
Treffers-Daller, J., Mukhopadhyay, L., Balasubramanian, A., Tamboli, V., & Tsimpli, I. (2022). How ready are Indian primary school children for English medium instruction? *Applied Linguistics*, *43*(4), 746–775. https://doi.org/10.1093/applin/amac003
Tsimpli, I., Mukhopadhyay, L., Treffers-Daller, J., Alladi, S., Marinis, T., Panda, M., … Sinha, P. (2019). Multilingualism and multiliteracy in primary education in India. *Research in Comparative and International Education*, *14*, 54–76.
UNESCO. (2020/2021). Global education monitoring report, 2021/2: Non-state actors in education: Who chooses? Who loses?

UNICEF. (2012). *Why are boys under-performing in education? Gender analysis of Four Asia-Pacific Countries*. UNGEI, United Nations Girls' Education Initiative.
Vaid, J., & Gupta, A. (2002). Exploring word recognition in a semi-alphabetic script: The case of Devanagari. *Brain and Language*, *81*(1–3), 679–690. https://doi.org/10.1006/brln.2001.2556
van Gelderen, A., Schoonen, R., Stoel, R. D., de Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, *99*(3), 477–491. https://doi.org/10.1037/0022-0663.99.3.477
Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29.

APPENDIX A

TABLE A1

Correlation Coefficients for the Relation Between Performance on the ASER Subtasks in Hindi (Data Collapsed Over Delhi and Patna, *n* = 811 Children)

Factor	Letters	Words	Paragraph	Story
Letters				
Words	0.63			
Paragraph	0.54	0.89		
Story	0.50	0.84	0.89	

Note. All correlations are significant at the $p < .001$ level. Correlations with comprehension are reported in Table 2.

TABLE A2

Correlation Coefficients for the Relation Between Performance on the ASER Subtasks in Telugu (in Hyderabad, *n* = 286 Children)

Factor	Letters	Words	Paragraph	Story
Letters				
Words	0.61			
Paragraph	0.48	0.89		
Story	0.44	0.81	0.87	

Note. All correlations are significant at the $p < .001$ level. Comprehension data were not obtained.

TABLE A3

Correlation Coefficients for the Relation Between Performance on the ASER Subtasks in English and in Hindi (Data Collapsed Over Delhi and Patna, *n* = 811 Children)

Factor	Letters English	Words English	Paragraph English	Story English	Comprehension English
Letters Hindi	0.74	0.42	0.25	0.23	0.11
Words Hindi	0.68	0.68	0.45	0.40	0.15
Paragraph Hindi	0.59	0.65	0.46	0.42	0.12
Story Hindi	0.55	0.67	0.53	0.48	0.11
Comprehension Hindi	0.55	0.62	0.42	0.39	0.17

Note. All correlations are significant at the $p < .01$ level.

TABLE A4

Correlation Coefficients for the Relation Between Performance on the ASER Subtasks in English and in Telugu (in Hyderabad, *n* = 286 Children)

Factor	Letters English	Words English	Paragraph English	Story English	Comprehension English
Letters Telugu	0.57	0.34	0.26	0.22	0.12
Words Telugu	0.52	0.46	0.42	0.38	0.16
Paragraph Telugu	0.44	0.46	0.42	0.35	0.14
Story Telugu	0.42	0.48	0.46	0.40	0.18

Note. All correlations are significant at the $p < .05$ level.

TABLE A5

Overview of the Statistical Results from the Linear Model Analyses on Measures of Decoding and Reading Comprehension in the Regional Language

Factor	Letters	Words	Paragraph	Story	Comprehension
Language in the home (majority/ minority)	$B = -30.61$; $t = -5.80$; $p < .001$***	$B = -32.31$; $t = -3.38$; $p < .001$***	$B = -35.60$; $t = -3.18$; $p = .002$**	$B = -34.40$; $t = -3.07$; $p = .002$**	$B = -42.09$; $t = -3.33$; $p < .001$***
Household type (monolingual/ multilingual)	$B = 2.07$; $t = 1.24$; $p = .22$n.s.	$B = 4.68$; $t = 1.54$; $p = .12$n.s.	$B = 5.77$; $t = 1.62$; $p = .10$n.s.	$B = 5.44$; $t = 1.53$; $p = .13$n.s.	$B = 5.95$; $t = 1.54$; $p = .12$n.s.
Interaction Language in the home * Household type	$B = 22.94$; $t = 4.00$; $p < .001$***	$B = 24.12$; $t = 2.32$; $p = .02$*	$B = 28.01$; $t = 2.30$; $p = .02$*	$B = 23.58$; $t = 1.94$; $p = .053$n.s.	$B = 29.91$; $t = 2.17$; $p = .03$*
City (Hyderabad)	$B = -0.76$; $t = -0.47$; $p = .64$n.s.	$B = -8.75$; $t = -2.98$; $p = .003$**	$B = -5.88$; $t = -1.71$; $p = .09$n.s.	$B = -2.46$; $t = -0.72$; $p = .47$n.s.	–
City (Patna)	$B = -1.68$; $t = -1.13$; $p = .26$n.s.	$B = -1.18$; $t = -0.44$; $p = .66$n.s.	$B = 7.79$; $t = 2.46$; $p = .01$*	$B = 14.03$; $t = 4.44$; $p < .001$***	$B = -4.45$; $t = -1.38$; $p = .17$n.s.

TABLE A5 *(Continued)*

Factor	Letters	Words	Paragraph	Story	Comprehension
Age	$B = -0.38$; $t = -0.60$; $p = .55$n.s.	$B = 0.04$; $t = 0.04$; $p = .97$n.s.	$B = 0.70$; $t = 0.53$; $p = .60$n.s.	$B = -1.21$; $t = -0.90$; $p = .37$n.s.	$B = 1.20$; $t = 0.74$; $p = .46$n.s.
Gender	$B = -1.07$; $t = -0.87$; $p = .39$n.s.	$B = -6.28$; $t = -2.81$; $p = .005$**	$B = -6.96$; $t = -2.67$; $p = .008$**	$B = -7.56$; $t = -2.90$; $p = .004$**	$B = -7.60$; $t = -2.49$; $p = .013$*

Note. n.s. = not significant; * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

TABLE A6

Average Percentage of Correct Responses (with Standard Deviation) on the ASER Subtasks and Comprehension Questions in the Regional Language (Hindi or Telugu) at Each Site and for Each Gender

Location	Letters	Words	Paragraph	Story	Comprehension
Delhi	94 (20)	77 (36)	68 (42)	58 (40)	69 (43)
Hyderabad	92 (20)	68 (37)	62 (47)	55 (48)	–
Patna	92 (22)	77 (37)	77 (41)	73 (42)	66 (44)
Boys	92 (21)	71 (39)	65 (44)	58 (44)	63 (44)
Girls	93 (20)	78 (35)	73 (42)	67 (43)	71 (42)

TABLE A7

Overview of the Statistical Results from the Linear Model Analyses on Measures of Decoding and Reading Comprehension in English

Factor	Letters	Words	Paragraph	Story	Comprehension
Language in the home (majority/ minority)	$B = -10.85$; $t = -2.47$; $p = 0.014$*	$B = -9.69$; $t = -1.24$; $p = 0.22$n.s.	$B = -6.94$; $t = -0.81$; $p = 0.42$n.s.	$B = -14.59$; $t = -1.86$; $p = 0.063$n.s.	$B = 10.13$; $t = 1.87$; $p = 0.062$n.s.
Household type (monolingual/ multilingual)	$B = 6.19$; $t = 3.85$; $p < 0.001$***	$B = 11.58$; $t = 4.05$; $p < 0.001$***	$B = 13.63$; $t = 4.36$; $p < 0.001$***	$B = 13.70$; $t = 4.78$; $p < 0.001$***	$B = 2.44$; $t = 1.23$; $p = 0.22$n.s.
Interaction Language in the home * Household type	$B = 3.74$; $t = 0.78$; $p = 0.44$n.s.	$B = 1.15$; $t = 0.13$; $p = 0.89$n.s.	$B = -3.55$; $t = -0.38$; $p = 0.71$n.s.	$B = 1.59$; $t = 0.19$; $p = 0.85$n.s.	$B = -9.47$; $t = -1.59$; $p = 0.11$n.s.
City (Hyderabad)	$B = 2.21$; $t = 1.42$; $p = 0.16$n.s.	$B = -1.66$; $t = -0.60$; $p = 0.55$n.s.	$B = 13.68$; $t = 4.51$; $p < 0.001$***	$B = 8.10$; $t = 2.91$; $p = 0.004$**	$B = 4.15$; $t = 2.16$; $p = 0.031$*
City (Patna)	$B = -5.22$; $t = -3.32$; $p < 0.001$***	$B = -4.95$; $t = -1.77$; $p = 0.077$n.s.	$B = 3.83$; $t = 1.25$; $p = 0.21$n.s.	$B = 2.52$; $t = 0.90$; $p = 0.37$n.s.	$B = -11.14$; $t = -5.76$; $p < 0.001$***
Age	$B = 0.27$; $t = 0.46$; $p = 0.65$n.s.	$B = 0.74$; $t = 0.71$; $p = 0.48$n.s.	$B = 1.51$; $t = 1.33$; $p = 0.19$n.s.	$B = 0.34$; $t = 0.32$; $p = 0.75$n.s.	$B = -0.11$; $t = -0.16$; $p = 0.88$n.s.

TABLE A7 *(Continued)*

Factor	Letters	Words	Paragraph	Story	Comprehension
Gender	$B = -0.61$; $t = -0.50$; $p = 0.62$n.s.	$B = -3.57$; $t = -1.64$; $p = 0.10$n.s.	$B = -2.60$; $t = -1.09$; $p = 0.28$n.s.	$B = -1.09$; $t = -0.50$; $p = 0.62$n.s.	$B = -3.73$; $t = -2.48$; $p = 0.013$*

Note. n.s. = not significant; * = $p < .05$; ** = $p < .01$; *** = $p < .001$.

TABLE A8

Average Percentage of Correct Responses (with Standard Deviation) on the ASER Subtasks and Comprehension Questions in English at Each Site and for Each Gender

Location	Letters	Words	Paragraph	Story	Comprehension
Delhi	90 (19)	57 (39)	31 (38)	26 (34)	13 (25)
Hyderabad	92 (18)	56 (36)	46 (42)	34 (39)	17 (36)
Patna	85 (27)	53 (41)	36 (45)	29 (43)	2 (12)
Boys	89 (21)	54 (38)	37 (42)	29 (39)	9 (25)
Girls	89 (22)	57 (39)	40 (43)	31 (39)	12 (29)