Economics of Education Review 91 (2022) 102312

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Mother tongue reading materials as a bridge to literacy☆

Margaret Leighton

University of St Andrews, United Kingdom

ARTICLE INFO

Keywords:
Primary education
Language of instruction
Literacy
Intervention

ABSTRACT

Children whose mother tongue is different from the language of instruction at school face additional challenges developing literacy skills. One approach to this favours immersion in the language of instruction, while another favours a transitional period of bilingual education. This paper evaluates the impact of a primary school literacy intervention. In addition to multi-faceted literacy support, the programme included a component of transitional bilingual education for students whose mother tongue is not the usual language of instruction. Over two years, the intervention raised literacy scores in the language of instruction by +0.41 sd, and literacy in mother tongue for minority language speakers by +0.75 sd. Our findings suggests that a light-touch transitional bilingual component can support minority language literacy without crowding out acquisition of the language of instruction.

1. Introduction

Nearly 40% of the world's population cannot access education in their mother tongue (Walter & Benson, 2012). Primary school children whose mother tongue is not the language of instruction face a double hurdle at school: learning the curriculum, and learning the language in which the curriculum is taught. There are two main approaches to integrating language learners into unilingual education systems. The simplest is immersion, in which students are taught immediately, and only, in the new language. Transitional bilingual education, on the other hand, teaches students in both their mother tongue and the new language for some period of time. Proponents of immersion claim that mother tongue and dominant language instruction are substitutes, while those who favour bilingual education argue they are complements.

This paper evaluates the impact of a literacy intervention trialled by Save the Children in a multi-lingual region with a unilingual school system. While the intervention was delivered to all students in the dominant language of instruction, it included an additional element of bilingual education for students from minority language backgrounds. The Bangladesh Children's Book Initiative (CBI) sought to promote literacy through storybooks, and worked throughout the book chain to achieve this: developing and publishing high quality, age-appropriate storybooks; training teachers and school leaders in storybook use; engaging with communities; and seeking to influence policy at a national scale. The programme, which ran for two years, was randomly assigned to 51 schools in the Khagrachari District of Bangladesh.

In addition to the other programme elements described above, the intervention produced books in bilingual editions, combining the language of instruction with local languages. For minority language speakers this was a departure from business-as-usual: speakers of these

☆ This paper would not have been possible without the excellent team from Save the Children who carried out the intervention and data collection: Deba Chakma, Akter Hossain, Holly-Jane Howell, Akidul Islam, Patrick Mroz-Dawes, Kirsten Mucyo, Dalim Tripura, as well as Aliza Nahar, Azmal Hossain, Zulfiqar Bushra, Islam Shahin, Mo Mo Shey and Shishir Mittro. The project on which the paper is based was made possible through the participation of the Government of Bangladesh, participating schools and staff, parents and children. I am indebted to Jakob Henning, Angus Holford, Himangshu Kumar, Lorenzo Neri and Jamin Speer, as well as seminar participants at the University College London Institute of Education, the Royal Economic Society, the Scottish Economic Society and the European Society for Population Economics for very helpful comments and discussions. Comments from the editor and two anonymous referees also helped substantially improved the paper. Research assistance from Rufaida Al Hashmi, Peter Lowley, Saloni Mishra and Duong Tran is gratefully acknowledged. The data used in this paper are drawn from a pilot study of the Children's Book Initiative in Bangladesh. The intervention and the data collection were carried out by Save the Children UK and Save the Children in Bangladesh. The collaboration between the author and Save the Children was facilitated through funding from the Save the Children Research and Technical Fund and the University of St Andrews' Impact and Knowledge Exchange fund. This study has been approved by St Andrews' University Teaching and Research Ethics Committee, approval code: EC14847. Conflict of interest statement: the author worked as a paid consultant to Save the Children during the baseline study for the intervention, and has also worked as a paid or unpaid consultant on other Save the Children monitoring and evaluation reports.

E-mail address: mal22@st-andrews.ac.uk.

https://doi.org/10.1016/j.econedurev.2022.102312
Received 29 June 2021; Received in revised form 31 August 2022; Accepted 5 September 2022
Available online 26 September 2022

languages do not normally encounter their mother tongue in the classroom. Because of this, speakers of minority languages effectively received an enhanced intervention, with respect to their dominant language speaking peers: along with the core components of the intervention described above, these children were also exposed to transitional bilingual education for the first time.

Did this mother tongue component compete with, or complement, literacy development in the language of instruction? The multilingual setting (which includes both native speakers of the language of instruction as well as minority language speakers), combined with the randomised design of the trial, allow us to investigate this. First, we estimate the average treatment effect of the programme on literacy development, both in the language of instruction (for all students) and in mother tongue (for minority language speakers). Second, we test whether minority language students in treatment schools, who received transitional bilingual education in addition to other intervention components, had smaller gains in dominant language literacy than their native-speaker classmates.

Measured at the end of two years of implementation, the intervention led to large and significant increases in literacy in the language of instruction (+0.41 sd). Minority language speakers also saw substantial gains in literacy in their mother tongue (+0.75 sd). We find no statistically significant difference in gains in the language of instruction by native speakers compared to language learners: this suggests that the acquisition of mother tongue literacy did not slow down language learner's acquisition of literacy in the language of instruction. These results, however, are imprecisely estimated: although the difference between groups is not statistically significant, the difference in point estimates is substantially in favour of native speakers. It remains possible that the mother tongue component of the intervention delayed literacy in the language of instruction for at least some minority language speakers.

This paper makes two contributions. First, we report the effectiveness of a multi-faceted literacy intervention that underwent rigorous experimental evaluation. In a systematic review of randomised experiments in primary school in developing countries, McEwan (2015) calculates a mean effect size of 0.12 for teacher training programmes, and 0.08 for provision of instructional materials; compared to these, our estimated treatment effects are large. The costs of the intervention, averaging 50.41 GBP per child, are not trivial; particularly when compared to an annual per-student primary school expenditure of around 62.50 GBP (Bhatta et al., 2019). Back-of-the-envelope calculations suggest the marginal cost is closer to 17.21–34.50 GBP per child, but this is still prohibitively costly for widespread roll-out.

Second, our results show that a modest element of transitional bilingual education can be introduced without crowding out minority language students' acquisition of literacy skills in the dominant language. This is an important finding in light of the debate on mother tongue education. While our results do not suggest that bilingual education *accelerated* literacy in the dominant language (as would be the case if mother tongue education were a strong complement to literacy in another language), neither do we find strong evidence that bilingual education delayed such literacy either (as it would have if the two were substitutes). It should be noted that our comparison across language groups in this domain is noisy, with some evidence suggesting that natives did have higher average literacy gains than language learners. We nevertheless find that the programme led to large gains in literacy for both groups. This suggests that concerns that mother tongue education necessarily competes with the standard curriculum are, in our context, unfounded.

Coming from an RCT, the degree to which these findings would apply in other contexts must be kept in question. The intervention was intense, and closely monitored. Efforts were made to engage a wide-range of stakeholders, including parents and communities: these elements, which could prove difficult to scale, cannot be separated from the overall findings. As others have rigorously demonstrated (Kerwin & Thornton, 2021), a lower-cost version of the same intervention could fail to demonstrate even a fraction of the positive changes found here. It is worth noting, however, that the bilingual component of this intervention did not rely on teachers sharing students' mother tongues. Instead, teachers were trained to facilitate the use of bilingual books regardless their own language skills. The success of this in fostering literacy in minority languages is an important finding for scaling transitional bilingual education elsewhere.

Indeed, our evaluation of the bilingual component which, although embedded in an intensive overall intervention, is itself relatively light-touch, is the particular contribution of this paper. In theory, such an element could be added or removed from similar interventions at little cost. Much of the previous experimental work on mother tongue or bilingual education in developing country contexts has studied programmes fully or substantially delivered in the mother tongue (see, for example, Kerwin and Thornton (2021), Laitin et al. (2019) and Piper et al. (2018)). Our study is therefore a useful benchmark for policy-makers considering light-touch mother tongue or local language curriculum supplements.

The remainder of the paper proceeds as follows. Section 2 introduces multilingual education, with a focus on mother tongue or transitional bilingual education, and reviews recent work in this area. Section 3 describes the intervention and data collection, and gives an overview of the data. Section 4 sets out the empirical approach. Section 5 presents the main results, while Section 6 presents extensions and robustness checks. Section 7 discusses the results and presents cost effectiveness estimates. Section 8 concludes.

2. Multilingual education

Historically, the relative merits of immersion versus bilingual education have been widely debated. Transitional bilingual education[1] offers children the opportunity to first learn to read in their native language, a skill which, theoretically, can then be more easily transferred to the new language. Immersion, on the other hand, requires students to learn the language of instruction quickly in order to access the curriculum. Proponents argue bilingual education is a stepping stone to literacy, while its detractors fear it sends students on an unnecessary academic detour, delaying progress in other areas.

Much of the literature on bilingual versus immersion education comes from the US, with a focus on integrating Spanish-speaking children into mainstream English education. After decades of debate (see Slavin and Cheung (2005) for an overview), recent evidence from methodologically rigorous studies finds little enduring difference between the two (Chin et al., 2013; Slavin et al., 2011). Based on this evidence, policy advice from the US emphasises the importance of education quality, rather than the immersion/bilingual dichotomy, as a primary driver of English language learner success (Chin, 2015).

In many developing countries, the goal of providing quality education is hampered by scarce resources, from classroom time and teacher skills, to the availability of quality materials. Furthermore, the linguistic landscape in often complex. In areas with multiple local languages, mother-tongue bilingual education is considerably more challenging, as it requires separate arrangements for each language. Many countries also juggle multiple official languages, with students moving through two or more languages of instruction during their education. In contexts such as this, a better understanding of the tradeoffs between immersion and transitional bilingual education could have significant returns.

The evidence on mother tongue education in developing countries is mixed. Some studies suggest that mother tongue education is in fact

[1] Bilingual education used to help language learners adapt to the dominant language of instruction is often referred to as transitional, to distinguish it from other forms of bilingual education, e.g. second or third language acquisition as enrichment.

a complement to skills acquisition in other areas. Using observational data from the Young Lives Survey, Hynsjö and Damon (2016) find that Indigenous children in Peru who attend Quechua-medium primary schools have substantially higher mathematics scores, but find no effect on Spanish language scores. This finding refutes local perceptions that bilingual education slows down the acquisition of Spanish.

Taylor and von Fintel (2016) use within-school changes in language policy to estimate the returns to mother tongue education on English language skills in South Africa. They find that students whose first three years of primary school were taught in an African language, as opposed to English, performed better at English language tests taken after the students transitioned to English medium instruction. Eriksson (2014) reaches a similar conclusion, but this time in the longer term, by considering the effect of South Africa's 1953 Bantu Education Act, which mandated 8 years of mother tongue education (from previously 4 or 6, depending on the state). She finds this longer period of mother tongue education is associated with higher literacy levels, more years of education, and higher earnings at age 28–48. She also finds greater English proficiency for individuals affected by the reform in English-dominated parts of the country (but not everywhere).

In 1994 Ethiopia passed a law which increased the offering of mother tongue primary education. Ramachandran (2017) finds the change substantially increased literacy in one of the five major languages, defined as being able to read a simple sentence. He also shows that this increase in literacy is accompanied by a greater use of written materials and better health knowledge. Seid (2018) investigates the effect of this same reform on test scores in grade 5, after students had transitioned to English-language education. He finds improved grade 5 mathematics scores for students who were first taught in their mother tongue, as opposed to a non-English second language. In contrast, Chicoine (2019) finds that the shift towards mother tongue instruction led to a decrease in years of schooling, particularly in areas which simultaneously adopted a new script. Chicoine's more recent study invokes data indicating the policy was adopted at an earlier date in many communities, and separately estimates the effect of the removal of school fees that occurred around the same time: two factors which can explain the contrasting conclusions from earlier work.

Evidence from an experimental programme in Cameroon, which provided local language instruction in the first three years of primary school as opposed to English, finds evidence slightly in favour of transitional bilingual education. Laitin et al. (2019) estimate that the programme increased test scores in mathematics and English substantially in 1st and 3rd grade, with smaller differences persisting once treated students had reverted to English-medium education in grade 5. Learning outcomes in the sample remained extremely low, however, and the authors conclude that the intervention is unlikely to lead to measurable differences in human capital formation.

In contrast, evidence from a medium-scale randomised control trial in Kenya suggests that mother tongue education may delay skills development in other academic subjects. Piper et al. (2018) carry out a large-scale literacy intervention in Kenya with two treatment arms: one taught in the usual language of instruction, and the other in mother tongue. Compared to the immersion arm, the mother tongue arm did show gains in mother tongue language skills; however, this group performed no better in English or Kiswahili, and worse in mathematics.

A recent study by Kerwin and Thornton (2021) helps make sense of contrasting findings from this literature. Their paper reports on a randomised control trial of a literacy intervention in Northern Uganda. Among other inputs, the intervention delivered the first three years of primary school in the local language. To understand how the details of programme delivery can influence treatment effects, their study had two treatment arms: a full-cost intervention designed to address a wide range of challenges in the local education setting, and a reduced-cost version which followed similar principles but cut back on the most costly elements, reducing the per-student costs by about 60%. The full-cost intervention resulted in substantial gains in reading and writing (0.64 sd and 0.45 sd, respectively); however, the reduced-cost version showed insignificant reading gains, with evidence of negative effects on writing. While the local language component does not appear to be responsible for this difference (both treatment arms increased the use of local language to a similar extent, compared with control), their results demonstrate how modest adjustments to an intervention can have dramatic implications for effectiveness.

Finally, Knauer et al. (2020) evaluate an intervention which trained parents in a low-literacy environment to share books with their children, and also provided books to facilitate this. This study shares two elements with our own: first, a focus on culturally-appropriate, quality storybooks; and second, the provision of these in bilingual editions covering the family's mother tongue as well as other locally-relevant languages. In contrast to our school-based intervention, Knauer et al. (2020)'s study engaged parents specifically. After a short follow-up, the intervention measured substantial success in promoting frequent and quality book sharing in the home, even among illiterate caregivers.

Of the papers reviewed here, Piper et al. (2018)'s study comes closest to our own. Both studies are based on the random assignment of a scalable intervention implemented in government-run schools (Piper et al., 2018's study, however, benefited from two experimental arms as compared to only one here). In addition to being carried out in a different context, the intervention studied here is lighter-touch than (Piper et al., 2018)'s Primary Math and Reading (PRIMR) Initiative, both in scope (PRIMR covered English, Kiswahili and mathematics, while the CBI's focus was language only) and in intensity (the CBI teacher training was 2 vs. 10 days, and the only materials provided were storybooks as compared with the more comprehensive PRIMR materials). Despite these differences, our results on language skills beyond the mother tongue align closely with theirs.

There are many reasons for which multilingual education might be desirable that are unrelated to educational efficiency (for an excellent overview, see Ginsburgh and Weber (2020)). While necessarily important in policy making, these considerations are beyond the scope of this paper, which focusses on a technical question: does a period of bilingual education help non-dominant language speakers transition to learning in the language of instruction?

3. Context and data

3.1. The intervention

The data used in this paper come from a trial of the Children's Book Initiative (CBI) literacy intervention carried out by Save the Children in Bangladesh. The intervention aimed to improve poor literacy outcomes in early primary school through the development, provision, and pedagogical use of high quality storybooks. The intervention was delivered to all students in the dominant language of instruction; however, it also included an element of bilingual education targeted at students from minority language backgrounds. This bilingual element was the creation and provision of storybooks in bilingual editions, pairing the language of instruction with local languages. This particular aspect of the CBI was designed with multi-lingual environments in mind; specifically those with little or no children's literature available in local languages.

In addition to this unique bilingual feature, the CBI intervention included a number of elements common to many literacy programmes. The programme ran training workshops for teachers and school leaders (a feature shared with Piper et al. (2018)'s PRIMR intervention), with a focus on the use of storybooks as a tool to teach literacy. At the workshops, teachers were taught how to engage students with books both in group reading and independent reading settings. Parents were engaged through community sensitisation sessions to promote book use at home (Knauer et al. (2020)'s Kenyan study features a similar component, although in their case it was a much more substantial part of the intervention); in some villages, community-based organisations

were supported to hold classes on ethnic scripts. Child-friendly book shelves were installed in participating school classrooms, and populated with books (the provision of books is a feature of many literacy interventions, including those studied by Kerwin and Thornton (2021), Knauer et al. (2020), and Piper et al. (2018)). Classroom teachers were expected to send books home with children, and were taught to keep a book lending register. The intervention team also worked with government stakeholders to influence policy towards bilingual and gender sensitive educational materials.

The programme's particular focus on storybooks makes it distinctive. The CBI espoused a 'whole book chain' approach to the provision of storybooks, working with local writers, illustrators and publishers to develop and produce high quality, gender sensitive storybooks with interesting and accessible text. Storybooks generated by the intervention were an important output in and of themselves, and ownership of these books remained with the authors and publishing houses. The CBI emphasised book quality over quantity, and invested considerable resources in ensuring the storybooks were carefully edited, age-appropriate and attractive. Finally, these books were produced in bilingual editions, incorporating local languages that were rarely, if ever, used in the classroom prior to the intervention.

Our study was carried out in the Chittagong Hill Tracts of Bangladesh, a mostly rural, remote hill country in the southeastern part of the country. The area is ethnically diverse: the schools in which the intervention was trialled enrolled students from four ethnic groups: Bangla (48%), Chakma (15%), Marma (28%) and Tripura (9%). Each of these groups speak their own language and have a distinct identity.[2] School is taught in Bangla, and many minority language speakers arrive at school with little or no knowledge of that language. These children effectively experience Bangla immersion when they attend school.

While teachers recruited from the local area would speak the language of their ethnic group, there is no guarantee a minority language speaking student would ever be taught by a teacher of his or her ethnicity. In our data 32% of minority language speakers attend a school where no teacher speaks their language. On average, 43% of teachers at a minority student's school share their language; this would correspond to 1 or 2 teachers in the median primary school. Minority languages are not taught in schools, and there is no formal opportunity for students to learn to read in languages other than Bangla. Situational analysis highlighted this as a source of alienation for non-Bangla ethnic groups (Save the Children, 2015).

The provision of bilingual reading materials through the CBI introduces an element of transitional bilingual education into what is otherwise a pure immersion setting. Native Bangla speakers experienced the educational enrichment provided by the storybooks, teacher training, and school and community engagement. Non-native Bangla speakers were also privy to these features of the programme, but they received an additional novel educational intervention on top of this: educational materials in their mother-tongue.

3.2. Study design and data collection

The CBI was implemented for two years, between 2016 and 2018, in randomly selected village schools in the Chittagong Hill Tracts. Based on a mapping exercise carried out prior to baseline data collection, 103 schools in 99 villages were identified as meeting criteria for inclusion in the pilot.[3] Prior to randomisation, the schools were stratified according to two criteria: first, the languages spoken by children at the school; and second, the language(s) spoken by teachers. This ensured that the treatment and control groups included a balance of children speaking different languages, and furthermore that minority language-speaking teachers were also balanced across groups. 51 schools were assigned to treatment, and 52 to control.

Data collection was carried out in two waves, approximately 23 months apart, at the start and end of the intervention. The data collection plan called for sampling of up to 14 first grade students in each school, with oversampling of minority-language speakers. In practice, almost half of schools had fewer than 14 first grade students present on the survey day: in these cases, the full class was sampled. Larger samples were drawn at bigger schools to maintain the target sample size. The baseline sample included 1313 children, with per-school sample sizes ranging from 5 to 29. At follow-up, 918 of the original cohort were re-surveyed (70%), and an additional 116 new children were recruited to meet sample targets. When possible this replacement sample was matched on sex, expected grade (in this case, grade 3), and language. One control school was not reachable at follow-up. Both surveys were organised and managed by Save the Children Bangladesh staff familiar with the Chittagong Hill Tracts.

Data used in this paper was collected using the CLASS child literacy assessment. This detailed survey tool covers a range of pre-literacy to literacy skills, starting with letter recognition and introducing more complex literacy skills up to, and including, story reading and comprehension. Bilingual enumerators administered the CLASS assessment to children whose mother tongue was not Bangla. These children completed two versions of the assessment: the Bangla version, and an equivalent translation in their mother tongue.

3.3. Data

3.3.1. Outcome variables

There are two outcomes of interest: literacy in Bangla (for all students), and literacy in a minority language (for minority ethnicity students only). A summary statistic for each of these is generated from the rich information collected by the CLASS tool as follows. First, the score on each section of the class tool is converted to a percentage grade. These grades are then standardised with respect to the mean and standard deviation of the control group for that wave. Finally, an overall score is created by averaging the percentage scores of each section. These in turn are standardised with respect to the control group. All six subsections of the tool are used to generate the Bangla literacy score; however only the first four are retained for the minority language literacy score. Further details on the component sections can be found in Appendix A.

3.3.2. Summary statistics and balance

Table 1 presents basic descriptive statistics from the first wave of data collection, split by treatment and control. The sample includes almost equal numbers of girls and boys; the average age is 6.9 years. The sample is composed of almost one third each Bangla and Marma students, one quarter Chakma students, and 13% Tripura students. Purposeful over-sampling of minority-language students reduced the Bangla share of the sample from the population-average of nearly 50%. At baseline, average Bangla literacy scores were a bit under 40%, while

[2] The languages are quite different from each other. Each of the four languages is written in a different script (Tripura using the latin alphabet). While Chakma and Bangla (also known as Bengali) are both Indo-European languages, Tripura (also known as Kokborok) is in the Sino-Tibetan family, while Marma is a Tibeto-Burman language.

[3] Inclusion criteria were determined by the implementation team in Bangladesh, based primarily on location of the villages in one of the participating upazila, and the schools not currently being served by similar literacy interventions. The selected villages are not necessarily representative of the Chittagong Hill Tracts; indeed, it has been suggested that the focus on villages not already enrolled in literacy programmes resulted in a set of villages that are particularly remote and difficult to access.

Table 1
Summary statistics from the first wave.

	Treat			Control			T-C	
	Mean	sd	N	Mean	sd	N	Diff	p-value
Girl	0.51	0.50	705	0.48	0.50	724	0.02	0.36
Age	6.89	1.12	654	6.92	1.27	639	−0.02	0.72
Chakma native	0.25	0.43	705	0.24	0.43	724	0.01	0.68
Marma native	0.29	0.45	705	0.32	0.47	724	−0.03	0.18
Tripura native	0.13	0.34	705	0.13	0.34	724	0.00	0.91
Bangla native	0.32	0.47	705	0.30	0.46	724	0.02	0.43
Bangla score	39.68	20.22	660	38.70	20.87	653	0.97	0.39
Minority language score	0.86	3.68	449	1.24	4.53	460	−0.38	0.17
Bangla score: learners	40.52	20.46	449	37.64	20.95	460	2.88	0.04
Bangla score: natives	37.89	19.64	211	41.24	20.53	193	−3.35	0.09

minority language literacy scores were close to 0. Bangla natives and Bangla learners had similar scores on the Bangla assessment.

The comparison of treatment and control samples in Table 1 shows that the sample is on the whole well-balanced. There are only very small and statistically insignificant differences between treatment and control on demographic characteristics and average test scores. There is, however, a difference in the relative performance of Bangla natives and Bangla learners on the Bangla assessment, with learners out-performing natives slightly in the treatment group, and natives outperforming learners in the control group. To explore whether this could be driving our results comparing these two groups, two alternative specifications are estimated: one which includes individual baseline test score controls, and another which includes student fixed effects. The results are robust to both of these.

3.3.3. Attrition

Of the 1313 children surveyed in the first wave, 918 were followed up in the second wave. This implies an attrition rate of 30%. Our estimation strategy requires that the control group is an appropriate counterfactual for the treatment group: if the two groups experienced different levels or kinds of attrition this could bias the estimated treatment effects. In terms of levels, 30.6% of the sample attrited in treatment schools, and 29.6% in control schools. This shows good balance. It also suggests that the intervention did not substantially reduce drop out in the treatment schools.

To explore which children are attriting, we first regress child characteristics on a dummy for attrition, and then repeat this using child characteristics interacted with treatment status. The results show that attrition itself is not related to treatment status; however, the pattern of attrition is different across treatment and control. Specifically, boys, weaker students and minority language speakers were more likely to attrit in the control group, while in the treatment group attrition is associated with girls and stronger students, but not minority language status (see Appendix Table A.2.)

This suggests that treatment is related to the characteristics of students who attrit from the sample, which could bias estimated treatment effects. Two approaches are taken to address this. First, the primary specification controls for child characteristics. It is reassuring to note that the inclusion of these controls is inconsequential to the main results. Second, in a robustness check, the sample observations are re-weighted according to the inverse of their probability of attrition. The main results are also robust to this.

4. Empirical approach

4.1. Overview

The estimation of the effect of the programme proceeds in two steps. First, the overall impact of the intervention is established by comparing the literacy scores of students in the treatment and the control groups. Second, the additional impact of the mother tongue component is estimated by comparing minority language versus dominant language speakers across treatment and control. The intuition and the identifying assumptions for these two steps are outlined below.

The effect of the intervention on Bangla literacy (for all children) and mother tongue literacy (for speakers of minority languages only) is estimated using difference-in-differences. This estimation relies on the randomisation of treatment. Balance checks at wave 1 suggest that randomisation was largely successful (see Table 1).

The additional impact of the mother tongue component is estimated by comparing the treatment-induced Bangla literacy gains of native Bangla speakers and minority language speakers. This is done using a triple-difference specification. The intuition is as follows. While all students in the treatment schools received the Children's Book Initiative intervention, Bangla speakers have always had learning materials in their mother tongue, while minority language speakers received these through the intervention for the first time. This means that, compared with Bangla speakers, minority language speakers in treated schools received an additional element of the intervention: these students now have access to learning materials in their mother tongue.

The validity of this estimation strategy relies on a parallel trend assumption. This requires that the difference in literacy scores between Bangla natives and minority language speakers in the control schools is a valid counterfactual for the difference in literacy scores between these two groups in the treatment schools. This could be violated if there are substantial non-linearities in learning gains that make gaps between different groups incomparable. While this assumption cannot be tested directly, it is reassuring to see that Bangla-natives and minority language speakers have broadly similar achievement levels at baseline (see Table 1).

4.2. Estimating equations

The estimation proceeds in two steps, as described above. The first estimating equation takes the following form:

$$y_{i,t} = \alpha_0 + \alpha_1 \text{PostTreat}_i + \alpha_2 \text{Post}_i + \alpha_3 \text{Treat}_i + \Gamma X_i + \epsilon_{i,t}, \quad (1)$$

where y_{it} is a student's literacy measure at time $t = 1, 2$, $Post = 1$ in the second wave, $Treat = 1$ for treated schools, $PostTreat$ is the interaction of $Post$ and $Treat$, and X_i is a short list of predetermined individual-level controls (sex and relative age, along with a language dummy). The coefficient of interest when estimating Eq. (1) is $\hat{\alpha_1}$, the estimated effect of treatment on literacy measures.

The effect of the transitional bilingual learning materials on Bangla literacy is estimated using a triple difference specification as follows:

$$y_{i,t} = \alpha_0 + \beta_1 \text{MinorPostTreat}_i + \Lambda_1 \text{Levels}_i + \Lambda_2 \text{Interactions}_i + \Delta X_i + \eta_{i,t}, \quad (2)$$

where $MinorPostTreat$ is the triple interaction between post, treat and the minor language dummy, $Levels$ includes dummy variables for post, treat and minority language, $Interactions$ includes the two-way interactions of the dummy variables in $Levels$, and X_i is as above. This triple difference equation estimates the effect of the intervention on

Table 2
Treatment effects: literacy in Bangla and minority language.

	(1) Bangla	(2) Bangla	(3) Other language	(4) Other language
Treat*Post	0.395***	0.405***	0.730***	0.745***
	(0.109)	(0.108)	(0.165)	(0.168)
Treat	0.0466	0.0285	−0.0831	−0.0863
	(0.121)	(0.117)	(0.0972)	(0.0978)
Post	9.14e−10	−0.0178	1.03e−08	0.0154
	(0.0720)	(0.0722)	(0.111)	(0.110)
Girls		0.0980**		0.0700
		(0.0492)		(0.0581)
Age at wave 2		−0.0266		0.0162
		(0.0271)		(0.0223)
Marma		−0.219*		0.0719
		(0.114)		(0.0777)
Tripura		−0.492***		0.981***
		(0.121)		(0.112)
Bangla		−0.190*		
		(0.105)		
Constant	−9.58e−10	0.391	2.93e−09	−0.404**
	(0.0915)	(0.252)	(0.0855)	(0.194)
Observations	2347	2339	1642	1636
R^2	0.032	0.059	0.072	0.201

Each column reports estimates from a separate OLS regression with the dependent variable a normalised literacy outcome. The dependent variable in Columns 1 and 2 is the overall Bangla literacy score, while the dependent variables in Columns 3 and 4 is the overall minority language score. All control variables are shown. Standard errors (in parentheses) are clustered at the school level.

minority language speakers relative to Bangla natives, given by the estimated coefficient $\hat{\beta}_1$.

While the triple difference facilitates testing directly for a negative interaction effect between the intervention and minority language status, a split-sample re-estimation of Eq. (1) provides a more intuitive comparison of the treatment effect for Bangla natives versus minority language speakers. This is presented as a complementary approach.

5. Main results

5.1. Impact of the intervention: treatment vs. control

The impact of the intervention on Bangla and minority language literacy is estimated using Eq. (1), with the summary literacy scores as dependent variable.[4] Table 2 presents the results for Bangla and other languages, with and without controls for student characteristics. From the preferred specification with controls, the intervention increased Bangla literacy by 0.405 standard deviations (Column 2), and literacy in mother tongue for minority language speakers by 0.745 standard deviations (Column 4). Both estimates are strongly significant, and substantial. Inclusion of basic controls has no effect.

5.2. Impact of transitional bilingual materials

Table 2 shows large and significant treatment effects of the intervention on both Bangla and minority language literacy skills. The question remains: did the mother tongue aspect of the intervention help or hinder minority language speakers' acquisition of literacy skills in Bangla? Would they have made more progress on literacy in Bangla if they had not had materials in their mother tongue? To answer this question we estimate first the triple difference Eq. (2), and second re-estimate Eq. (1) on the sample split by language status. The allows us to compare the gains made by Bangla and minority language speakers, with respect to their peers in the control group.

These results are presented in Table 3. Column 1 shows estimates of Eq. (2). The parameter of interest here is the triple interaction of treatment, post period and minority language status. Although the estimated coefficient is negative (and not trivial at 0.175 standard deviations), it has a large standard error (0.192) is not statistically significant. Columns 2 & 3 show results from a re-estimation of the simple treatment effect Eq. (1), split by Bangla natives (Column 2) and minority language speakers (Column 3). The key point of interest here is a comparison of the treatment effect estimates (interaction of treatment and post period) across the two groups. As in Column 1, the difference between these is not inconsequential ($0.524 - 0.347 = 0.177$); however, a test of equality between these coefficients in a seemingly unrelated equations framework fails to find a statistically significant difference between the two (chi squared = 0.88, p-val = 0.349). While these tests fail to demonstrate a statistically significant difference between the gains in Bangla literacy across the two groups, it is important to note that the estimates are noisy. The point estimate on the interaction term in Column (1) is not a precisely estimated zero, but rather a large coefficient with a large standard error. We conclude from this that, while there is no statistically significant difference between the gains in Bangla literacy made by Bangla natives, there is some evidence of large gains for natives, perhaps attributable to sub-groups not identified in this study.

6. Extended results and robustness checks

6.1. Extended results

6.1.1. Heterogeneous effects

Substantial heterogeneity could exist in both the overall treatment effects, and in the relative size of the treatment effect across language groups. To explore this, the fully interacted (triple difference) Eq. (2) is re-estimated on Bangla literacy using subsets of the panel split by: baseline test score, child sex, and child age. This yields two coefficients of interest: the treatment effect on Bangla natives, and the treatment effect on minority language speakers relative to that.

These two parameters are shown in Table 4, for each split of the panel. With one exception (the younger half of the panel), the estimates are consistent with the overall trend: a large and significant treatment effect on Bangla natives, and a statistically insignificant interaction term between treatment and minority language status. The main (Bangla native) treatment effect for younger children remains positive and substantial, but is statistically insignificant: this suggests a smaller treatment effect on younger versus older children in the cohort. While the coefficient on the interaction term is statistically insignificant in all subsets, it is worth nothing that the point estimates are moderately large and negative for boys, younger children, and those who had higher baseline test scores. This suggests that the bilingual component may have slowed the acquisition of Bangla literacy skills for some children in these categories.

6.1.2. Teacher languages

While all teachers would teach in Bangla, teachers recruited from the local area may also speak one of the local languages: this could help minority language speakers acquire literacy skills in their mother tongue, and could also potentially help such students transition to Bangla education. Summary data on teachers at sample schools was collected during a mapping exercise several months prior to the baseline data collection in 2016; this data was not updated, and we do not have a mapping of children to teachers over the three academic years spanned by the study. The mapping data does, however, allow us to determine the composition of teachers at each school at the start of the study. From this data we calculate whether there was *any* teacher at the school at that time who spoke the child's mother tongue. The schools in

[4] Treatment effects on literacy score components are shown for Bangla in Table B.1, and for minority languages in Table B.2.

Table 3
Interacted treatment effects: literacy in Bangla.

	(1) Triple difference	(2) Bangla speakers	(3) Non-Bangla speakers
Treat = 1	−0.173	−0.168	0.117
	(0.156)	(0.156)	(0.139)
Post = 1	−0.120	−0.122	0.0245
	(0.115)	(0.116)	(0.0862)
Treat = 1 × Post = 1	0.524***	0.524***	0.347***
	(0.157)	(0.157)	(0.131)
Minority language = 1	−0.473***		
	(0.151)		
Treat = 1 × Minority language = 1	0.290		
	(0.188)		
Post = 1 × Minority language = 1	0.145		
	(0.137)		
Treat = 1 × Post = 1 × Minority language = 1	−0.175		
	(0.192)		
Girls	0.0973**	0.192*	0.0586
	(0.0489)	(0.109)	(0.0545)
Age at wave 2	−0.0266	0.0307	−0.0529*
	(0.0271)	(0.0456)	(0.0316)
Chakma	0.490***		
	(0.120)		
Marma	0.273**		−0.208*
	(0.124)		(0.115)
Tripura			−0.478***
			(0.120)
Constant	0.322	−0.238	0.580**
	(0.278)	(0.434)	(0.287)
Observations	2339	699	1640
R^2	0.062	0.036	0.076

Each column reports estimates from a separate OLS regression with the overall Bangla literacy score as dependent variable. Column 1 estimates Eq. (2) on the full sample. Columns 2 and 3 estimates the primary treatment effect specification (Eq. (1)), with Column 2 restricted to Bangla natives, and Column 3 to minority language speakers. All control variables are shown. Standard errors (in parentheses) are clustered at the school level.

Table 4
Heterogeneous treatment effects: split sample estimates.

	(1) High score	(2) Low score	(3) Girls	(4) Boys	(5) Old	(6) Young
Treat = 1 × Post = 1	0.514***	0.450**	0.476**	0.537***	0.648***	0.320
	(0.155)	(0.180)	(0.198)	(0.203)	(0.181)	(0.195)
Treat = 1 × Post = 1 × Minority language = 1	−0.266	0.0829	−0.0377	−0.298	−0.00198	−0.235
	(0.190)	(0.230)	(0.236)	(0.238)	(0.228)	(0.239)
Observations	1131	1092	1187	1152	1228	1111
R^2	0.132	0.206	0.099	0.044	0.088	0.060

Each column reports estimates from a separate OLS regression of Eq. (2), with the overall Bangla literacy score as dependent variable. Each pair of columns presents results from an estimation on roughly one half of the sample, split as follows: high (1) vs. low (2) baseline Bangla scores; girls (3) vs. boys (4); older (5) vs. younger (6) children. Only selected estimates are displayed: other coefficients include a full set of interactions between treatment, post period and minority language status, as well as basic controls, as shown in Table 3. Standard errors (in parentheses) are clustered at the school level.

our sample are small, with a mean and median of 4 teachers per school, and 91% of schools having 5 or fewer teachers: if teacher turnover is modest, there is a reasonable chance the sample child would be taught by any given teacher over a 3 year period.

To explore the potential mediating role of teachers who share a minority language student's mother tongue, the primary treatment effects (Eq. (1)) are re-estimated on subsets of the sample defined by whether any teacher at the school speaks the child's mother tongue. By default, all Bangla speakers fall into this category, so the analysis is run with and without the Bangla natives when relevant.[5]

The results are presented in Table 5. Surprisingly, having a teacher at the school who shares their language results in smaller treatment effects for minority language speakers. Minority language speakers at a school with no teachers who speak their language (Column 3) see 0.621 standard deviations improvement in Bangla literacy scores, compared with their peers at schools with teachers who do speak their language (Column 2), whose improvement is statistically insignificant. Even more surprisingly, this pattern holds even for minority language literacy (although in this case the difference is smaller, and the effect is significant for both groups), where presumably having a teacher who can speak the language would be a particular asset.

Unfortunately the data do not allow us to explore this result further. It may be that the languages spoken by teachers are correlated with other characteristics we do not observe: for example, it could be that average education levels differ between those who speak local languages and those who speak only Bangla. It could also be that teacher

[5] Because of this, we cannot estimate the triple difference Eq. (2) on a subset of students at schools where no teachers speak their language, as that would exclude all Bangla students.

Table 5
Sample split: at least one teacher at the school speaks child's mother tongue.

	(1)	(2)	(3)	(4)	(5)
	Yes*	Yes	No	Yes	No
	Bangla literacy			Minority language literacy	
Treat*Post	0.337***	0.209	0.621***	0.684***	0.840**
	(0.113)	(0.145)	(0.224)	(0.176)	(0.355)
Observations	1821	1122	518	1119	517
R^2	0.064	0.090	0.083	0.174	0.265

Each column reports estimates from a separate OLS regression of Eq. (1), with either Bangla literacy score (Columns 1, 2 & 3) or minority language literacy score (Columns 4 & 5) as the dependent variable. Sample is split according to whether at least one teacher at the school speaks the child's native language (Yes) or not (No). Bangla natives are excluded from all but Column 1 (Yes*).

composition varies according to the remoteness of schools, for example if lay teachers are appointed in rural areas when qualified teachers are not available. In any event, these findings are intriguing and invite further study.

6.2. Robustness checks

6.2.1. Inverse probability weighting

Differences were observed in average characteristics of attriters across treatment and control schools. While the primary specification controls for observed child characteristics, as an alternate specification we also estimate the main results using inverse attrition-probability weights. To do this, we first estimate the probability of attrition on the first wave sample using a probit model with the full set of controls. We then predict the probability of attrition, and re-weight the data by $1/(1-p)$. The results are not sensitive to this alternate approach (see Appendix Table B.6).

6.2.2. School and student fixed effects

To check whether unmeasured community characteristics are driving differences in achievement, the main specification is re-estimated with school fixed effects. The estimated treatment effects are very similar (see Appendix Table B.4). To control for unobserved student characteristics, the results are separately re-estimated using a fixed effects model. Although this restricts the sample to students observed at both waves, there is no qualitative change in the results (see Appendix Table B.5).

6.2.3. Lagged dependent variable

As an alternate approach, the main and interaction effects are estimated using a lagged dependent variable specification. As above, we first estimate the overall effect of the intervention; we then estimate the effect of the transitional bilingual element through an interaction between treatment and minority language status. The first estimating equation takes the following form:

$$y_{i,t} = \tilde{\alpha}_0 + \tilde{\alpha}_1 y_{i,t-1} + \tilde{\alpha}_2 \text{Treat}_i + \tilde{\Gamma} X_i + \tilde{\epsilon}_{i,t}, \quad (3)$$

where y_{it} is a student's literacy measure at time $t = 1, 2$, $Treat = 1$ for treated schools, and X_i are the student controls. The coefficient of interest is the treatment effect $\tilde{\alpha}_2$.

We estimate the effect of the transitional bilingual learning materials as follows:

$$y_{i,t} = \tilde{\beta}_0 + \tilde{\beta}_1 y_{i,t-1} + \tilde{\beta}_2 \text{Treat}_i + \tilde{\beta}_3 \text{Minor}_i + \tilde{\beta}_4 \text{Minor}_i * \text{Treat}_i + \tilde{\Delta} X_i + \tilde{\eta}_{i,t}, \quad (4)$$

where $Minor = 1$ for children who are minority language speakers and $Minor = 0$ for Bangla natives. This lagged dependent variable equation estimates the effect of the intervention on minority language speakers relative to Bangla natives, given by the coefficient $\tilde{\beta}_4$.

The results of these regressions align very closely with the primary specification: the point estimates of the treatment effect are very similar, and the interaction between treatment and minority language status remains statistically insignificant, with an even smaller point estimate (see Appendix Table B.3.).

7. Discussion

7.1. The intervention

Our estimates show that, measured at the end of two years of implementation, the intervention resulted in substantial improvements in literacy. The effect sizes we estimate are quite large. In a review of randomised controlled trials to promote learning in developing countries, McEwan (2015) finds a mean effect size for teacher training interventions of 0.12 sd, and for instructional materials of 0.08 sd. The 0.41 sd gain in literacy skills found in our study could be due to a number of features. First, this intervention was quite long, lasting two years (the average duration in McEwan's study was 12.9 months), and combined multiple intervention elements: teacher training, development and provision of instructional materials, school and community involvement. Second, our outcomes were measured immediately after the end of the intervention, before they had time to fade (this is also true for the majority of the studies in McEwan's sample, but not all).

While larger than the average in McEwan's review, our effect sizes are not out of line with recent findings from mother-tongue interventions. Indeed, starting from very low levels of attainment, Laitin et al. (2019) found that mother tongue instruction in Cameron increased math scores by over 1 sd, and English scores by only slightly less; (Kerwin & Thornton, 2021) measure a 0.64 sd improvement in reading scores after one year of their mother-tongue-first literacy intervention in Uganda.

We find even larger standardised gains in mother tongue literacy, but these come from a very low baseline: the average baseline literacy score in minority languages was less than one point on a hundred-point scale, with a standard deviation of 3.68 points; in contrast the Bangla score was 39.7, with a standard deviation of 20.2 (see Table 1). Our 0.75 sd gain in mother tongue literacy remain comparable to the recent literature: in Kenya, Piper et al. (2016) estimate a 0.3–0.6 sd increase in mother tongue literacy from the PRIMR programme; in Uganda, Kerwin and Thornton (2021) measure 1.0 sd improvements in mother tongue letter recognition, with their largest finding being a 1.3 sd increase in the ability to write one's first name.

How might we expect these improvements to effect labour market outcomes, particularly for minority language speakers? (Berman et al., 2003) study Hebrew (dominant language) fluency among immigrants from the Soviet Union to Israel. They find that Hebrew fluency had no effect on wages in low-skilled occupations, but a substantial effect in high-skilled occupations. Using linguistic distance as an instrument for the difficulty of acquiring fluency in the dominant language, Laitin and Ramachandran (2016) estimate the returns to dominant language fluency in India. They find that fluency in the dominant language is associated with greater educational attainment and better occupational outcomes.

In our context of Bangladesh, the language of instruction is also the dominant language at a national level. Improvement in Bangla fluency is therefore of particular relevance for future labour market outcomes. Bangla fluency is likely to be especially critical for minority language speakers with the prospect of entering higher skilled occupations, as well as those who migrate away from their home communities.

It should be noted that primary school learning remains modest in the sample, even in the treatment group. Mean endline literacy scores for basic literacy skills such as simple word reading and word recognition are only 78%–86%, despite the children having been in school for at least 3 years. This is consistent with UNICEF estimates of literacy in Bangladesh suggesting that, by the end of 2nd and 3rd grade, only 25% of children could read 90% of words in a story and answer inferential questions. For comparison, 57% of children at a similar grade level in Thailand could perform the same reading comprehension

task.[6] While interventions such as this one can help children make significant progress on literacy, there remains much to be done.

7.2. Transitional bilingual materials

Are transitional bilingual materials a substitute or a complement to the production of literacy skills in the dominant language? Estimates from Table 3 demonstrates that the additional mother tongue component of the intervention did not demonstrably hinder, and nor did it accelerate, non-native speakers' development of Bangla literacy. These results are similar to those found by Piper et al. (2018) in Kenya.

One of the primary arguments in favour of immersion over bilingual education is that it will bring students up to speed in the language of instruction more quickly, without taking a costly detour through bilingual literacy. Our results suggest that, at least for a light touch bilingual programme such as this, such concerns are unfounded. Despite substantial gains in literacy in their mother tongue, non-Bangla children in the treated schools saw equivalent growth in Bangla literacy skills as their Bangla-native peers.

Does this mean that mother tongue literacy and literacy in the dominant language are neither substitutes nor complements? Our analysis does not answer this definitively. Specifically, we do not know how teachers used these bilingual materials, and to what extent they did or did not divert class time from instruction in Bangla towards instruction in minority languages. It could be that Bangla instruction time was decreased for minority language children, but that complementarities between the two meant that these children easily kept up. It could also be that teachers did not adjust teaching to make use of these bilingual materials, with no resulting loss in Bangla-focussed instruction. The bilingual learning materials could instead have brought new learning resources for minority language children, either through their parents or through an increased motivation to engage with their more-inclusive school.

Ultimately, our results are (mildly) good news for bilingual education: minority language learners benefited by acquiring literacy in their mother tongue, and did not suffer very large negative consequences with respect to their literacy skills in the language of instruction. It should be noted, however, that while we do not find statistically significant differences in the literacy improvements between Bangla natives and language learners, these results are imprecisely estimated. While our results suggest that Bangla natives may have had larger gains than language learners, both groups show large and positive improvements in Bangla literacy due to the intervention. Given the potential benefits of mother tongue education beyond its role as a bridge to mastery of the dominant language, this is a success in itself. We do not find evidence, however, that a light touch bilingual programme provides language learners a fast track to literacy.

7.3. Cost effectiveness

The Bangladesh Children's Book Initiative was carried out as a pilot study over two years. The project is estimated to have directly reached 5893 children, 5513 parents, and 1031 teachers and school leaders. As part of this, 20,123 bilingual storybooks were distributed.

Calculating the cost effectiveness of complex interventions such as this is challenging, particularly in a pilot setting where fixed costs tend to be a large share of the total budget. It is nevertheless helpful to put the effects estimated here into context with respect to the cost of the programme. The most straightforward approach is to calculate the total cost of the intervention per child served. Reaching 5893 children with a total budget of 297,047 GBP gives a per-child total cost of 50.41 GBP (as the project ran for two years, 25.21 GBP per child-year). To put this into context, the average annual per-student primary school expenditure in Bangladesh in 2016 was approximately 62.54 GBP (Bhatta et al., 2019).

Estimating the cost per unit of outcome is more challenging, particularly when the programme has multiple outcomes. If the sole purpose of the intervention were to raise Bangla literacy scores, the cost effectiveness would be: 50.41 GBP/0.41 sd = 122.95 GBP/sd improvement in Bangla literacy (0.0081 sd/GBP). Assuming that the gains in mother tongue literacy were equally important, and assuming (as found in the mapping exercise) that 52% of treated children were minority language speakers, would give a different calculation. Suppose 5893 children gained 0.41 sd in Bangla, and $(0.52 * 5893 = 3064)$ children gained 0.75 sd in minority language literacy. Then the cost per sd in literacy improvement is closer to 297,047 GPB/$(0.41 * 5893$ sd $+ 0.75 * 3064$ sd) = 297,047 GPB/(2416 sd + 2298 sd) = 297,047 GPB/(4714 sd) = 63.01 GBP/sd improvement in literacy in any language (0.016 sd/GBP).

Some assumptions about which costs are fixed versus variable allows us to approximate a marginal cost estimate. Using broad cost categories (e.g. output level costs, non-thematic costs), gives an estimate of fixed costs as 32% of the total budget; a line-by-line assessment (e.g. development of books, production of books, teacher training, manager salary) gives a higher estimate of fixed costs, at 66% of total. With these figures as upper and lower bound estimates, reaching an additional child would cost the programme between 17.21–34.50 GBP. If the sole outcome of interest is Bangla literacy gains, the marginal cost per sd is therefore between 41.98–84.15 GBP (0.024–0.012 sd/GBP). Assuming both minority language and Bangla literacy are equally important, the marginal cost per sd in literacy improvement is 21.51–43.13 GBP (0.046–0.023 sd/GBP).[7]

8. Conclusion

This paper presents evidence from a literacy intervention piloted in a multi-lingual area of rural Bangladesh. The complex literacy intervention was implemented in the language of instruction for all students, but also included a light-touch transitional bilingual component for minority language speakers. It achieved substantial gains in literacy: after two years of implementation, treated students had 0.41 sd higher literacy scores than students in control schools. Minority language students, who benefited from the provision of storybooks in both the dominant language and their mother tongue, saw considerable gains in early literacy skill development in their mother tongue, scoring 0.75 sd higher than control group students.

We do not find strong evidence that the bilingual element of the programme either accelerated or decelerated non-native Bangla speakers' acquisition of Bangla literacy skills. There is no statistically significant difference between the effect of the intervention on Bangla literacy gains for native and non-native Bangla speakers; however, these results are imprecisely estimated, with point estimates suggesting a substantially larger improvement in literacy among native speakers. In light of this, we hesitate to conclude that there are no important differences in the treatment effects across the two groups. What we can say with confidence is the intervention led to large improvements in both groups. On the one hand, this is positive news for bilingual education: minority

[6] Data from the Multiple Indicators Cluster 6 Survey (MICS6) 'UNICEF Global database on foundational learning skills' (April 2021 version), accessed at https://mics.unicef.org/surveys.

[7] For comparison, Kerwin and Thornton (2021)'s pilot study in Northern Uganda achieved an improvement in reading of 0.032 SD per USD (full-cost version of programme) and 0.018 SD per USD (reduced-cost version). The intervention for that study took place in 2013, while the CBI started in 2016: in that year the USD:GPB conversion ranged from approximately 0.7 to 0.8. Assuming a conversion rate of 0.75, the Uganda intervention had a GBP cost effectiveness of 0.043 (full-cost) to 0.024 (reduced-cost). These measures are for reading in mother tongue only, as the programme relied on a mother-tongue-first approach, with the dominant language of instruction introduced in later years.

language speakers in the study gained skills in their native language, with the intervention nevertheless leading to a net improvement to their literacy in the language of instruction. On the other hand, it suggests that hopes that bilingual education could be a highly-efficient bridge to literacy in other languages may be unfounded.

This study has a number of limitations. First, the intervention ran for two years, with data collection taking place immediately afterwards. It is possible that complementarities between mother tongue literacy and the acquisition of other skills would grow with time — or indeed that the costs of investing in bilingualism might emerge at higher grade levels. Due to this timeline, we also have no information on how long the measured literacy gains persist, in either language. Second, learning outcomes for other subjects have not been collected. In Kenya, Piper et al. (2018) found that mother tongue education reduced achievement in mathematics: a channel we cannot test with our data. Finally, the study was designed to evaluate the intervention as a whole, and as such further assumptions are required for us to analysis of the mother tongue component separately.

Furthermore, there are a number of limitations in applying the lessons learned from a pilot study such as this to transitional bilingual education programmes more widely. As Kerwin and Thornton (2021) starkly demonstrate, the context and the design of the intervention matter. The bilingual education introduced through this intervention was very minimal: students were provided with some learning materials, but may or may not have had any literate speakers of their language at home or at school in order to make these materials relevant. While our data show that these materials nevertheless had some effect, it is not straightforward to extrapolate these findings to other contexts — or even to the same context with a modified intervention design.

This study adds to the growing body of work suggesting that bilingual education is not a panacea for the challenges faced by low-attainment, multi-lingual schools. Language of instruction is, however, an issue with implications far beyond the scores on literacy assessments. It would be beyond the scope of this research to make recommendations on whether mother tongue education is in the best interest of students from minority language groups. With respect to literacy outcomes in the dominant language of instruction, we find no benefits from a small bilingual addition to the curriculum: on the other hand, we find no costs either. The benefits in other domains, from mother tongue literacy to community buy-in for primary education, could be large.

The results presented above raise two important questions for further research. First, the gains in mother tongue literacy are large, given the nature of the intervention: how did they come about? Bilingual storybooks are a small input into the production of literacy skills: through what mechanism did they help children gain literacy in their mother tongue? In particular, it would be useful to estimate the relative contributions of school and household factors in this process. A better understanding of this would be useful when applying lessons from this study to other contexts.

Second, the study here was limited to short-run outcomes: literacy in Bangla and in the mother tongue were both measured immediately after the two-year intervention. This is a limitation our study shares with the majority of evaluations of education interventions: 90% of the instruction-focused studies in McEwan (2015) review collected endline data less than one month from the end of the study. Evidence from early child development interventions suggests that fade-out is a serious issue (see, e.g., Andrew et al. (2018)); there is good reason to be concerned about similar issues here. The long-term impacts of transitional bilingual education are critical to language policy decisions. Unfortunately further follow-ups are not possible with this study; future research should prioritise establishing middle to long-term effects on a range of outcomes.

CRediT authorship contribution statement

Margaret Leighton: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Data availability statement

The data used in this paper are property of Save the Children. Researchers interested in using the data are encouraged to contact the author.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econedurev.2022.102312.

References

Andrew, A., Attanasio, O., Fitzsimons, E., Grantham-McGregor, S., Meghir, C., & Rubio-Codina, M. (2018). Impacts 2 years after a scalable early childhood development intervention to increase psychosocial stimulation in the home: A follow-up of a cluster randomised controlled trial in Colombia. *PLOS Medicine*, *15*(4), e1002556 – 19.

Berman, E., Lang, K., & Siniver, E. (2003). Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics*, *10*, 265–290.

Bhatta, S. D., Genoni, M. E., Sharma, U., Khaltarkhuu, B. E., Maratou-Kolias, L., & Asaduzzaman, T. M. (2019). *Bangladesh education sector public expenditure review: Tech. rep.*, World Bank.

Chicoine, L. (2019). Schooling with learning: The effect of free primary education and mother tongue instruction reforms in Ethiopia. *Economics of Education Review*, *69*, 94–107.

Chin, A. (2015). Impact of bilingual education on student achievement. *IZA World of Labor*, 1–10.

Chin, A., Daysal, N. M., & Imberman, S. A. (2013). Impact of bilingual education programs on limited english proficient students and their peers: Regression discontinuity evidence from Texas. *Journal of Public Economics*, *107*(C), 63–78.

Eriksson, K. (2014). Does the language of instruction in primary school affect later labour market outcomes? Evidence from South Africa. *Economic History of Developing Regions*, *29*(2), 311–335.

Ginsburgh, V., & Weber, S. (2020). The economics of language. *Journal of Economic Literature*, *58*(2), 348–404.

Hynsjö, D., & Damon, A. (2016). Bilingual education in Peru: Evidence on how Quechua-medium education affects indigenous children's academic achievement. *Economics of Education Review*, *53*(C), 116–132.

Kerwin, J. T., & Thornton, R. L. (2021). Making the grade: the sensitivity of education program effectiveness to input choices and outcome measures. *The Review of Economics and Statistics*, *103*(2), 251–264.

Knauer, H. A., Jakiela, P., Ozier, O., Aboud, F., & Fernald, L. C. (2020). Enhancing young children's language acquisition through parent-child book-sharing: A randomized trial in rural Kenya. *Early Childhood Research Quarterly*, *50*, 179–190.

Laitin, D. D., & Ramachandran, R. (2016). Language policy and human development. *American Political Science Review*, *110*(3), 457–480.

Laitin, D. D., Ramachandran, R., & Walter, S. L. (2019). The legacy of colonial language policies and their impact on student learning: Evidence from an experimental program in Cameroon. *Economic Development and Cultural Change*, *68*(1), 239–272.

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, *85*(3), 353–394.

Piper, B., Zuilkowski, S. S., Kwayumba, D., & Oyanga, A. (2018). Examining the secondary effects of mother-tongue literacy instruction in Kenya: Impacts on student learning in English, Kiswahili, and mathematics. *International Journal of Educational Development*, *59*, 110–127.

Piper, B., Zuilkowski, S. S., & Ong'ele, S. (2016). Implementing mother tongue instruction in the real world: Results from a medium-scale randomized controlled trial in Kenya. *Comparative Education Review*, *60*(4), 776–807.

Ramachandran, R. (2017). Language use in education and human capital formation: Evidence from the Ethiopian educational reform. *World Development*, *98*, 195–213.

Save the Children (2015). *Language use and literacy skills: READ baseline assessment and situational analysis: Tech. rep.*

Seid, Y. (2018). The impact of learning first in mother tongue: evidence from a natural experiment in Ethiopia. *Applied Economics*, *51*(6), 577–593.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research*, *75*(2), 247–284.

Slavin, R. E., Madden, N., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Educational Evaluation and Policy Analysis*, *33*(1), 47–58.

Taylor, S., & von Fintel, M. (2016). Estimating the impact of language of instruction in South African primary schools: A fixed effects approach. *Economics of Education Review*, *50*, 1–15.

Walter, S. L., & Benson, C. (2012). Language policy and medium of instruction in formal education. In B. Spolsky (Ed.), *The Cambridge handbook of language policy* (pp. 278–300).