Routledge
Taylor & Francis Group

Journal of Research on Educational Effectiveness

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/uree20

Reading Skills Transfer Best from Home Language to a Second Language: Policy Lessons from Two Field Experiments in South Africa

Nompumelelo Mohohlwane, Stephen Taylor, Jacobus Cilliers & Brahm Fleisch

To cite this article: Nompumelelo Mohohlwane, Stephen Taylor, Jacobus Cilliers & Brahm Fleisch (2024) Reading Skills Transfer Best from Home Language to a Second Language: Policy Lessons from Two Field Experiments in South Africa, Journal of Research on Educational Effectiveness, 17:4, 687-710, DOI: 10.1080/19345747.2023.2279123

To link to this article: https://doi.org/10.1080/19345747.2023.2279123

JOURNAL OF RESEARCH ON EDUCATIONAL EFFECTIVENESS
2024, VOL. 17, NO. 4, 687–710
https://doi.org/10.1080/19345747.2023.2279123

Routledge
Taylor & Francis Group

INTERVENTION, EVALUATION, AND POLICY STUDIES

Reading Skills Transfer Best from Home Language to a Second Language: Policy Lessons from Two Field Experiments in South Africa

Nompumelelo Mohohlwane[a], Stephen Taylor[a], Jacobus Cilliers[b] and Brahm Fleisch[c]

[a]Research, Monitoring and Evaluation Directorate, South Africa Department of Basic Education, Pretoria, South Africa; [b]McCourt School of Public Policy, Georgetown University, Washington, DC, USA; [c]Department of Education, University of Witwatersrand, Johannesburg, South Africa

ABSTRACT

In many countries, children need to become proficient in both their home language (L1) and an international language, such as English (L2). Governments face tradeoffs in how to prioritize these two objectives. We provide empirical evidence on cross-linguistic transfer between L1 and L2, using the results of two randomized evaluations of Structured Pedagogy Programs implemented in South Africa. The programs had the same design, implementing organization, and duration. The key difference is that one program targeted the teaching of reading in L1, while the other targeted L2. We find that both interventions had positive effects on the languages they targeted. The L1 intervention also had a positive effect on L2 reading proficiency. In contrast, the L2 intervention had a negative effect on L1 outcomes, for the lower-performing students. These results are consistent with the Simple View of Reading and suggest that decoding skills are best learned in L1. It is thus cost-effective to prioritize learning to read in L1, as well as supporting teachers in this subject, even if becoming proficient in L2 is also regarded as an important policy objective.

ARTICLE HISTORY
Received 27 January 2023
Revised 18 October 2023
Accepted 21 October 2023

KEYWORDS
Language transfer; mother-tongue; language policy; teacher professional development; structured pedagogy

Introduction

The choice of language of instruction (LOI) is one of the most important policy questions in education. This is especially true in developing countries that are typically multilingual with an "international language" such as English or French chosen as the lingua franca. Most countries opt for a bilingual language policy, teaching first in the mother-tongue (L1) and then introducing the foreign/second language (L2), often as the destination language (Eberhard et al., 2020; Kosonen, 2017; Walter & Benson, 2012). This is based on education theory recommending that schooling should begin in the language a child knows best (Ouane & Glanz, 2010; UNESCO, 2003; Walter, 2010;

CONTACT Jacobus Cilliers ejc93@georgetow.educ McCourt School of Public Policy, Georgetown University, Washington, DC, USA.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/19345747.2023.2279123

Walter & Benson, 2012). But policymakers wrestle with reconciling this with the strong demand from parents for their children to be taught in the international language,[1] given the economic returns and associated job opportunities vested in these languages (Casale & Posel, 2011; Giliomee, 2004; Gordon & Harvey, 2019; Wright, 2002). For example, some countries, such as Botswana and Tanzania, prioritize one LOI nationally. Other countries, such as Uganda and Kenya, prioritize one LOI in the urban areas such as English and Kiswahili, respectively.[2] Rwanda *reversed* its language policy in 2021, requiring all primary schools to instruct their students in English after at least three language changes between Kinyarwanda, and French.[3] In South Africa, the mother-tongue is mostly used as the LOI for the first three grades, with English taught as a subject, followed by a switch to mostly English as the LOI.

The evidence for theories of language transition has largely been from the global north with an application for minority language speakers acquiring a majority language. Very little evidence exists for contexts where the majority of the population are mother-tongue speakers attempting to acquire a minority second language. For instance, although seen as an "international language," English is often the minority language in the global south (Eberhard et al., 2020; Kirkpatrick, 2012; Kosonen, 2017; Walter & Benson, 2012). Moreover, the empirical research has largely focused on linguistically similar language pairs such as Spanish to English, and English to French or German (Baker & Stoolmiller, 2011; Goswami et al., 1998; Melby-Lervåg & Lervåg, 2011). Recent evidence examines more linguistically diverse pairs such as English to Korean and English to Chinese (Wang et al., 2006, 2009). Except for limited studies on English and local languages in Kenya, Uganda, and Namibia (de Galbert, 2023; Kim & Piper, 2019; Piper et al., 2018; Veii & Everatt, 2005), there is very little research on the applicability of these language transfer theories for the African context, where there is typically a larger language distance between L1 and L2 within a multilingualism context (Eberhard et al., 2020; Ouane & Glanz, 2010; Prah & Brock-Utne, 2009; UNESCO, 2003).

Do improvements in L1 literacy skills also cause improvements in L2, when the language pair is linguistically distinct? Is there a reverse causal relationship flowing from L2 to L1 literacy skills? And given limited resources, which language, L1 or L2, should teacher professional development programs target?

Our article examines these questions in the context of South Africa, using longitudinal data from two randomized evaluations of structured pedagogy programs targeting early grade literacy, called the Early Grade Reading Studies (EGRS). The first was a mother tongue L1 intervention implemented over three years from 2015 to 2018 (Cilliers et al., 2020, 2022a; Fleisch, 2018; Fleisch et al., 2017; Kotze et al., 2019) and the second was an English L2 intervention also implemented over three years from 2017 to 2019 (Cilliers et al., 2022b). Although these were two separate randomized evaluations

[1]Low-cost private schools in India, for example, often use English as the language of instruction. https://www.epw.in/journal/2021/13/special-articles/learning-and-language.html. Accessed June 6th 2022.

[2]Early grade language of instruction (LOI) in Botswana is either Setswana or English, even though 26 languages are spoken in the country. Kiswahili is the early-grade LOI in Tanzania, and urban centers in Kenya. English is the LOI in urban Uganda. In rural areas the language of the catchment area is used, which is typically the mother tongue.

[3]https://www.worldpoliticsreview.com/in-rwanda-language-change-in-schools-leaves-students-and-teachers-struggling/. Accessed October 10th 2022. https://www.washingtonpost.com/politics/2020/01/24/third-time-11-years-rwanda-changed-language-used-primary-schools/. Accessed May 9th 2023.

rather than two arms within one randomized experiment, there were many factors held constant. In both studies, learners were assessed in their mother tongue as well as in English; and we tracked the same group of learners over a period of four years, or more. Importantly, in both studies, a very similar intervention model was implemented, namely detailed lesson plans, integrated reading materials, and professional support through on-site coaching delivered over a 3-year period to children in grades 1–3 attending non-fee paying schools in rural South African provinces. The key difference was the language being targeted.

We find that targeting the home language leads to improvements in both home language and English literacy skills. In contrast, targeting English causes an improvement in English literacy, but a deterioration in home language literacy for students in the bottom half of the distribution. These results are consistent with the Simple View of Reading.

Literature and Theories of Reading Acquisition

The Simple View of Reading

Becoming literate relies on children understanding that spoken language can be represented by written language, and then learning what symbols represent these sounds. Thus, literacy development rests on both knowledge of the oral language, and the way it is represented in print. The Simple View of Reading (SVR) defines comprehension as the product of two skills: decoding and oral language (Gough & Tunmer, 1986). The continued relevance of SVR has been demonstrated by a meta-analysis of 155 studies (Quinn & Wagner, 2018), and empirical validation (Rodriguez-Segura, 2022). Additional contributors to SVR identified as supplementary predictors to the simple view of reading are background knowledge (Quinn & Wagner, 2018; Spires & Donley, 1998) working memory (Quinn & Wagner, 2018; Schaefer & Kotzé, 2019) and reasoning and inference skills (Ardington et al., 2021; Quinn & Wagner, 2018). Several studies have applied the SVR to L2 acquisition and found a similar relationship for Dutch, Spanish, and English bilingual learners. Both oral language comprehension and decoding were strong predictors of comprehension (Lee et al., 2022).

Although the SVR argues for an underlying similarity in language acquisition and each language requires specific instruction, learning an L2 is dependent on learning an L1. According to the developmental interdependence theory (Cummins, 1979), there are minimum language thresholds that have to be reached in L1 first before successfully learning an L2; without reaching these thresholds, learners may fail to become literate in either language rather than becoming biliterate. In other words, successful transfer only happens under certain conditions. Besides the specific reading skills that transfer from L1 to L2, learners would not need to learn each language entirely, there are some literacy skills that transfer from L1 to L2 including linguistic knowledge about how language works; process knowledge of using and handling of books; story grammar of narrative, and styles of social interaction; decontextualizing language as well as abstract thinking (Cummins, 1979; Kim & Piper, 2019; Macdonald, 1990).

The largest body of evidence on skill transfer from L1 to L2 is in phonemic and phonological awareness—focusing on letters and word reading with transfer established

through correlation studies (Branum-Martin et al., 2006; Koda & Reddy, 2008; Wang et al., 2006). Both correlation and causal evidence from the global south have been limited, an early exception of causal evidence was contributed by Wawire and Kim (2018), based on a randomized control trial in Kenya. They found a positive impact of an 8-week Kiswahili intervention on phonological awareness and letter-sound knowledge in both Kiswahili and English. This result leaves open the possibility that a similar English intervention might also have had a positive spillover to Kiswahili. Our study builds on this by examining both L1 to L2 and L2 to L1 skill transfers, and by doing so in the context of more substantial curriculum-wide programs running over three years (grades 1–3). More recent work measured language transfer in Uganda for two local languages, Luganda and Runyankole-Rukiga, and English (de Galbert, 202) and in Namibia measuring language transfer predictors in Herero and English (Veii & Everatt, 2005).

Deciding on the Language of Instruction (LOI)

There is an ongoing debate on the choice of the LOI–L1 or L2–particularly in the global south. The conversation is not only about the educational benefits but alsoabout heritage and preservation of indigenous languages while reconciling the pressure to transition to a "foreign/international language." The colonial history of these languages contributes to the tension (Alexander, 2005; Giliomee, 2009; Ngcobo, 2009; Wright, 2002). Furthermore, the LOI question is shaped by an understanding of reading acquisition for both L1 and L2.

Several traditional theories inform bilingual education including the immersion and submersion theories that are now widely critiqued as subtractive–replacing L1 with L2 (Cohen & Swain, 1976; Padilla et al., 1990; Parkin et al., 1987). More recent approaches (Ball, 2011; Collier & Thomas, n.d.; Feltes, 2022; Ginkel, 2014) include two-way bilingualism or dual immersion, using two languages as the LOI; multilingualism, using more than two LOIs in the curriculum; and mother-tongue-based education that culminates in a transition to an L2. In the latter, the L1 is used as a bridging language to L2, marked by an early exit (e.g. grade 4) or a late transition to L2 (e.g. grade 7).

Proponents for an early introduction of L2 paired with an early exit from L1 argue that young children are more efficient at acquiring L2, and there is a critical period for L2 acquisition (Colombo, 1982; Lenneberg, 1967, 1969). This has now been disproven (Ball, 2011; Ginkel, 2014; Snow & Hoefnagel-Höhle, 1977). More recent research confirms that young learners are not more efficient than older children or adults at acquiring L2. Adults and children differ in how they learn pronunciation, accent, and syntax. When to introduce L2 and how efficient this will be is more substantially affected by contextual factors such as whether learners are exposed to the L2 outside the classroom, the number of hours of L2 teaching, and teacher competency in teaching L2 (Cummins, 1980; Dutcher & Tucker, 1995; Ginkel, 2014; Walter, 2010).

Establishing when learners have enough mastery for a language to be the LOI is a critical policy question. How long should L1 be maintained before introducing and then transitioning to the L2 as the LOI? There is often an understated distinction between learning a language sufficiently and learning a language proficiently enough to learn

content in it. Cummins (1979, 1980) refers to this as the linguistic threshold hypothesis; it attempts to express the stage at which learners have sufficient skills to successfully learn content in either L1 or L2 as a LOI. He argues that there is a minimum language threshold required to successfully become literate and then biliterate.

Empirical research is ongoing to establish which skills and thresholds fit the linguistic threshold hypothesis. The skills to be measured have moved from grammar to vocabulary, although the exact number of high-frequency words required is still under discussion. This has contributed to the development and use of high frequency words in the English language as well as other benchmarks such as Oral Reading Fluency (Ginkel, 2014; Hasbrouck & Tindal, 2017). Such thresholds have however been underdeveloped in multilingual contexts and for agglutinating languages. Nakamura et al. (2019) contribute to research in multilingual contexts by examining two local South Indian languages, Kannada or Telugu as L1 and English as L2. They propose nascent thresholds for decoding skills including syllable level-phonological awareness and phoneme level phonological awareness in L1 for successful transfer to L2. While tentative, their work contributes to emerging empirical data on skill-based thresholds by context and language. A broader critique of existing work on thresholds (Nakamura et al., 2019; Takakuwa, 2005) is firstly, whether they are absolute or relative (e.g. a specific number of words reached by a specific age or grade or if it's a range in the number of words reached at any age). Recognizing, however, that most skills like vocabulary are continuous or unconstrained and thus difficult to measure. Secondly, how choices on which skills should be considered are made.

While the literature reviewed has been broad, important gaps have been identified. Firstly, there are limited studies validating the range of theories in the global south. Secondly, most of the existing literature lacks causal evidence for language transfer and thirdly, the possibility of language transfer through structured learning programs is limited. Our article contributes to these areas, respectively.

South African Context

South Africa is a middle-income country with nine provinces and corresponding education departments. While the performance of South Africa has improved significantly in the past decade, overall levels of learning remain low. For example, South Africa scored on average 320 in the 2016 Progress in International Reading Literacy Study (PIRLS), whereas the midpoint is 500 points (Gustafsson, 2020a; Howie et al., 2017).

South Africa is similar to many developing countries in its linguistic diversity, with twelve official languages. English is the dominant language used in post-school education and spoken in commerce (Le Cordeur, 2013; Madadzhe, 2019; Nudelman, 2015) although it is a minority L1 language, spoken by <10% of the population as an L1 (Statistics South Africa, 2012). Low English contextual exposure is found in broader social communication as seen through radio and television programs. Out of the top 10 largest radio stations, only two of these use English predominantly. In a typical week with a listenership of 37 million adults, only 30% are listening to English radio stations. Access to international television platforms and streaming services is increasing, however, the public television broadcaster still has a large n average audience of 26 million adults monthly with a legal

mandate to spend 50% of their independently produced program budget on local African language or programs (South African Broadcasting Services, 2020).

Language policy

Notwithstanding the highly constrained English societal context, the language policy balances the need for children to learn to read and write in a language they understand, with a decision to develop proficiency in English. Schools are encouraged to use mother-tongue as the LOI while simultaneously teaching a first additional language, usually English, from the first grade. Although the language policy allows schools to implement this model up to the sixth grade, the language of teaching changes for the majority of learners from the start of the fourth grade. Ninety percent of learners transition to English as the LOI while 10% transition or continue in Afrikaans from the fourth grade with the exception of a handful of schools that retain an African Language L1 as LOI. Approximately 70% of learners learn in their L1 as the LOI for the first three grades, and there is a high match between the LOI and their actual Home Language according to population data (Gustafsson, 2020b). Significantly, however, ~23% of learners start learning in English as their LOI, which is almost four times more than the English home language population figure which stands at 6% (Gustafsson, 2020b). A detailed discussion on the rationale and challenges of maintaining L1 for longer or selecting English as L2 is complex, with a detailed discussion provided by Mohohlwane (2019).

Of note to this study are differences in the South African national curriculum framework between teaching L1 *vs.* L2 early grade literacy (Curriculum & Assessment Policy Statement, 2011). According to the curriculum, teaching in L1 assumes that children have a basic command of the oral vocabulary as they enter grade one, but have not yet developed decoding skills. The primary object then of the early grade curriculum is to build and become fluent in decoding skills. In L1, learners are taught to master phonological awareness, letter naming, letter sounds, phoneme/grapheme relationships, and sounding out syllables and letter blends. Combined, these skills are used to decode words. These skills and strategies are taught from Grade 1 and continue into Grade 3 in the L1 with a maximum time allocation of 8 h per week.

In contrast, not only is the teaching of English as a First Additional Language (EFAL) allocated less time in the week, a maximum of 3 h, but the focus of teaching is different. In Grade 1, the focus is not on reading, but on oral vocabulary development. In subsequent years, the curriculum largely assumes that the decoding skills have been mastered in the L1 and would be transferred to L2. Even though there are clearly distinct demands for learners to master different consonant blends in English, the assumption is that decoding knowledge transfer would take place.

Program Description and Evaluation Design

Program Description

We use data from the first and second Early Grade Reading Studies (EGRS I and EGRS II) conducted between 2015 and 2020 in South Africa. The two studies were similar interms of the research team, program design, duration, sample, and service provider.

First, both studies evaluated a structured learning program, which provided daily lesson plans to teachers with integrated materials combined with an on-site coach providing monthly in-classroom support to teachers as well as needs-based workshops. In each study, the lesson plans were fully integrated with the official government curriculum. Second, the studies also had the same duration and target grades: teachers were supported over 1 year with the intervention following the learners for three years: i.e., grade 1 teachers supported in year 1, then grade 2 teachers supported in year 2, and grade 3 teachers supported in year 3. Third, in both studies, the program targeted poor, non-fee-paying public schools, referred to as quintile 1 to 3 schools.[4] Fourth, the studies were initiated and supervised by the Department of Basic Education, but in each case, implementation was outsourced to the same service provider.[5] Fifth, both studies were designed as randomized control trials, the experimental design details are provided in the next section.

The key difference between the two studies is that the first Early Grade Reading Study (EGRS I) targeted teaching of Setswana Home Language (L1) and the second study (EGRS II) targeted teaching English as a First Additional Language (EFAL) (L2).[6] The programs were also implemented in two different provinces in South Africa and were two years apart. EGRS I started in 2015 and was implemented in North-West province, whereas EGRS II started in 2017 and was implemented in Mpumalanga, where the home language is either Siswati or isiZulu. In both studies, schooling was continuing as usual in line with the curriculum, and there were no dedicated interventions targeting other languages or subjects.[7]

The provinces are relatively similar in terms of poverty levels and education performance. In each province, 68.7% of schools were classified as quintile 1 to 3 (South African School Act & 1996 (Act No 84 of 1996), 1996), 2022). The North West province was ranked fourth (out of nine provinces) and Mpumalanga ranked sixth based on their performance in the 2016 PIRLS assessment (van Staden & Bosker, 2014). The provinces also retained a similar ranking in the national school-leaving examination of 2020, at fifth and sixth position, respectively (Department of Basic Education, 2021).

The dominant home language is different in these two provinces. In North West province Setswana is dominant, spoken by 63% of the province. In Mpumalanga isiZulu, spoken by 24%, and Siswati, spoken by 27% of the population, are the two dominant languages (Statistics South Africa, 2012). A comparison of the actual home language of learners and the LOI by province and language in the Foundation Phase shows a 61% Siswati and 64% isiZulu match for learners in Mpumalanga. While in North West 100% of Setswana home language speakers are learning in Setswana as the LOI (Van der Berg et al., 2020).

[4]The majority of schools in South Africa are public (93%) (Department of Basic Education, 2021). Public schools in South Africa are classified by school Socio-Economic Status, referred to as quintiles. These are not perfectly proportional; quintile 1 to 3 schools are the poorest constituting 83% of schools overall. Furthermore, quintile 1 to 3 schools are non-fee paying.

[5]Class Act educational services and Molteno Institute for Language and Literacy.

[6]Neither program had any specific components aimed at transferring literacy skills across languages.

[7]Note that since the curriculum stipulates that teachers should dedicate 7–8 h to teaching home language, compared to 3–4 h teaching EFAL. The different amount of time allocated to the targeted language of the two interventions may have contributed both the main outcomes and transfer.

These languages were also the L1 in this study. All three African languages have transparent orthographies: there is a one-to-one mapping in the grapheme–phoneme relationship, unlike English which has an opaque orthography and a more complex mapping. They are however distinct in their morphology i.e., how words relate to each other. Two of the languages, isiZulu and Siswati, are part of the Nguni language group distinguished by their conjunctive morphology: one word may represent a sentence (Khumalo, 1987). In contrast, Setswana has a disjunctive morphology with short word segments written separately from suffixes and prefixes (Machobane & Mokitimi, 1998). For example, the sentence "there was a stranger who was very hungry" is written as "Kunesihambi esasilambile kakhulu" in isiZulu using only three words and "Go na le moeng o a neng a tshwere ke tlala thata" in Setswana using twelve words.

The morphological differences in the L1s required careful test development within each study as well as across the two studies. The most comparable L1 items were the letter-sound naming, Rapid Object Naming, and Oral Reading Fluency. The one-for-one letter-sound mapping allowed for a simple design aspect across EGRS I and II, with the same sequence of letters followed for 70% of the assessments. The differences were in ensuring that high-frequency letters—less than five letters per language—appeared early in the assessment. The same items were also used for Rapid Object Naming across the languages and the object names were not particularly different in length or complexity. The Oral Reading Fluency passages however were distinct. The passages differed by length, with shorter isiZulu passages, averaging 59 words for isiZulu and 60 words for Siswati. The average length for Setswana was 159 words. This was informed by the specific language structure; we expected learners to read faster in Setswana. The passage length aligns with the newly developed reading benchmarks for African languages, which specify 35 words correct a minute for Nguni languages by the end of Grade 3 and 60 words correct a minute for Sesotho-Setswana languages (Mohohlwane et al., 2022). The passages were piloted and revised to ensure appropriateness as part of the development.

Experimental Design

The studies were evaluated using a clustered randomized control trial. In each study, we created 10 strata of 13 schools that were similar in terms of socio-economic status and exam performance. We then performed stratified random assignment, allocating 50 schools (5 per strata) to the intervention and 80 schools to the control.[8] Previous articles reporting on the results found that both programs were successful at improving literacy in the targeted language (Cilliers et al., 2020, 2022b).

Data Collection

A data collection service provider visited the sample sites multiple times over the period of the evaluation, each time collecting detailed data on student learning and teaching

[8]The studies also included additional treatment arms, to compare the cost-effectiveness of different modalities of implementation. We restrict this paper to the treatment arms that are comparable across studies, and were also the most cost-effective.

practices (see Table 1). In each study, a random sample of 20 grade 1 students was sampled and assessed at baseline (before the start of the invention) and tracked over a period of four years or more. The learner assessments consisted of one-on-one adapted Early Grade Reading Assessments (EGRA) complemented by an exam-type setting written assessment. These were administered by trained fieldworkers using electronic tablets for the one-on-one assessments on the Tangerine or SurveyCTO app. The assessments consisted of tasks assessing letter recognition, listening comprehension, word recognition, oral reading fluency, and comprehension as well as written comprehension.

Table 2 below provides a summary of the learner assessments across the data collections. While EGRS I and II focused on L1 and L2, respectively, learners were assessed in both L1 and L2 across the studies. The assessment instruments are most similar in the round of data collection one year after the completion of the intervention (i.e., when our sample of non-repeating students was in grade 4). The English assessments are exactly the same.[9] Although the words included in the home language assessment are different across studies, because the languages assessed were different, they were each designed to be similar in difficulty. Section A.1 in the Supplementary Appendix provides more detail on the assessment instruments.

Data Analysis

Empirical Strategy

We estimate the impacts on each sample of pupils (EGRS I or II) and each grade year separately, using the following equation:

$$y_{icsb} = \beta_0 + \beta_1(\text{Treatment})_s + X'_{icsb}\Gamma + \rho_b + \epsilon_{icsb}, \quad (1)$$

where y_{icsb} is the outcome of interest for student i who was taught by a teacher in class c, school s, and stratum b; (Treatment)s is a dummy variable equal to one if the student was assigned to the treatment group; ρ_b refers to strata fixed effects; X_{icsb} is a vector of controls;[10] and ϵ_{icsb} is the error term clustered at the school level.

When conducting mediation analysis, we follow the methods proposed by Imai et al. (2010) and estimate the following in conjunction with Equation (1):

$$y_{icsb} = \gamma_0 + \gamma_1(\text{Treatment})_s \ \gamma_1 M_{icsb} + X'_{icsb}\Gamma + \rho_b + \epsilon_{icsb}, \quad (2)$$

The difference, $\hat{\beta}_1 - \hat{\gamma}_1$, can be interpreted as the causal mediation effect: the impact on y_{icsb} that can be attributed to improvements in the mediating variable, M_{icsb}. We treat this analysis as suggestive, given the strong underlying assumptions.[11]

[9]Eight additional words were added in the word recognition test for EGRS II. We drop these items for our analysis. The same story was used for Oral Reading Comprehension although the time to administer the task changed from 1 to 3 min. We adjust for this and use only 1 min reading across both studies. Any other differences in the test administration are also addressed for the analysis.

[10]The control variables include the different baseline measures of student home language literacy (phonemic awareness, letter recognition, etc.), student gender and age, and some measures of school socio-economic status.

[11]The strongest assumption is that M_{icsb} is independent of y_{icsb}, conditional on treatment and baseline covariates, X'_{icsb}. Even though treatment is random, there might be other factors that cause both M and y.

Table 1. Timelines for the intervention and data collection EGRS I and EGRS II.

	2015	2016	2017	2018	2019	2020
EGRS I						
Intervention	Grade 1 teachers	Grade 2 teachers	Grade 3 teachers	No intervention		
Data collection	Start and end of the year	End of the year	No data collection	End of the year		
EGRS II						
Intervention			Grade 1 teachers	Grade 2 teachers	Grade 3 teachers	No intervention
Data collection			Start and end of the year	End of the year	End of the year	End of the year

Table 2. Common items of the learner assessments across rounds of data collection.

	Grade 2				Grade 3		Grade 4			
	EGRS I		EGRS II		EGRS II		EGRS I		EGRS II	
Construct	*HL*	*EFAL*	*HL*	*EFAL*	*HL*	*EFAL*	*HL*	*EFAL*	*HL*	*EFAL*
Letter-sound recognition	X		X		X		X			
Word reading fluency	X	X		X		X	X	X		X
ORF	X		X	X	X	X	X	X	X	X
ORF comprehension	X		X	X	X	X	X	X	X	X
Written comprehension							X	X	X	X

HL: home language; EFAL: English as a first additional language; ORF: oral reading fluency.
Notes. Light-grey shaded refer to EGRS I assessment instruments. No data was collected at the end of grade 3 for EGRS I.

Internal Validity Tests: Balance and Attrition

As a starting point for the analysis, we show in Table A2.1 in the Supplementary Appendix that, for both studies, the samples are balanced across treatment arms, even after accounting for attrition. Next, Table A2.2 shows that the attrition rates by the end of year four—28% in each study—are balanced across treatment and control. Moreover, the coefficient on "Attrition × Treatment" shows that there is no systematic difference between those who attrited in the Treatment arm and those who attrite in the Control, with the exception of age in EGRS I: attriters in the treatment arm are younger relative to attriters in the control. Taken together, we have confidence that our sample is balanced and that attrition does not bias our results.

Next, we also investigate the similarity of the two samples, by comparing the English reading proficiency levels of the control students in grade 4, when the assessment instruments are most comparable across studies. Balance between these two samples is not necessary for the internal validity of the results but does strengthen the argument that the differences we observe in effect sizes are due to differences in the language targeted in the intervention, rather than due to treatment heterogeneity. Table 3 shows that students in the two samples performed similarly in English word recognition and Oral Reading Fluency (ORF), with the EGRS I sample slightly out-performing the EGRS II sample. Note that the EGRS II grade 4 sample was assessed during the Covid period, so there might have been some learning loss.

As a further comparison, Figure 1 shows the proportion of students in the control groups by grade and study who cannot read a single word, either in L1 or L2. Even after four years of school, over a fifth of learners cannot read a single word in English; and a slightly smaller fraction (15 and 20% in EGRS I and II, respectively) could not read a single word in their home language.

Implementation Quality and Earlier Impacts

Next, we show in Table 4 that the quality of implementation was high and remarkably similar in both programs. For example, 94 and 95% of treated teachers in ERGS I and II, respectively, reported receiving training at the beginning of the year; over 90% of treated teachers in each study reported having access to the graded reading booklets distributed by the program; and 90% of treated teachers in EGRS I report having the lesson plans, with 85% of treated teachers in ERGS II reporting to *use* the lesson plans.

Table 3. Comparison between EGRS I and II samples in grade 4 English literacy.

Variable	(1) EGRS 1 Mean/SE	(2) EGRS II Mean/SE	*t*-Test Difference (1)–(2)
Word recognition	29.412	27.535	1.877
	[1.161]	[1.020]	
Oral reading fluency (ORF)	39.131	36.480	2.652
	[1.722]	[1.475]	
ORF comprehension	0.166	0.230	−0.065***
	[0.010]	[0.012]	
Written compr.	0.184	0.184	0.000
	[0.009]	[0.010]	

Notes. Grade 4 data, with samples restricted to control group students in each study. The value displayed for *t*-tests are the differences in the means across the groups. Standard errors are in square brackets and are clustered at the school level. *** indicates significance at the 1, 5, and 10% critical level.

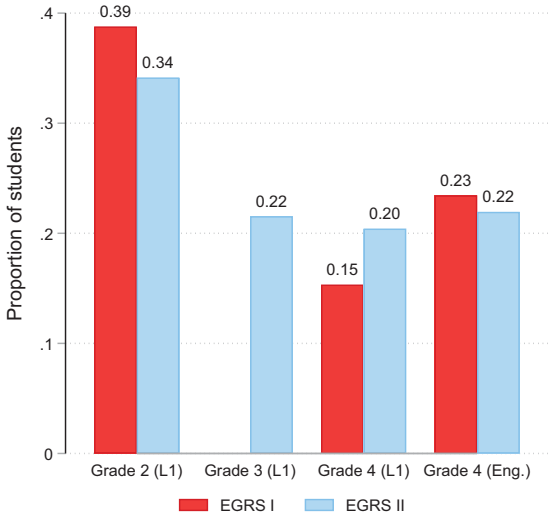


Figure 1. Proportion of students cannot read a single word, by grade, language, and sample.

Moreover, observed teaching practices improved in *both* programs, along similar dimensions, which indicates that the programs succeeded in causing the desired instructional change. For example, data from classroom observations revealed that the proportion of classes where a pupil reads individually to a teacher increased by 40 and 33 percentage points in EGRS I and II, respectively. This suggests that differences in the impacts of the program cannot be attributed to differences in the quality of implementation, or differences in teaching practices emphasized by the respective programs.

As further evidence of the quality of implementation, we show in Figure 2 that each program succeeded in improving the intended literacy outcomes in the targeted language.[12] In EGRS I, students improved their home language literacy by 0.15 and 0.25

[12]The outcomes reported in Figure 2 aggregate scores constructed using principal component analysis, and standardized to have a control mean of one and standard deviation of zero. See Table 2 for the constituent indicators

Table 4. Implementation quality.

(A) Teacher surveys	(1) EGRS I Mean	(2) EGRS II Mean
Received training beginning of the year	0.94	0.95
Access to graded reading booklets	0.9	0.96
Use graded reading booklets		0.93
Access to lesson plans	0.9	
Use lesson plans		0.85
Use lesson plans daily		0.79
(B) Impact on teaching practices	Coef./(SE)	Coef./(SE)
Group-guided reading	0.378***	0.293***
	(0.157)	(0.148)
Pupils read individually to teacher	0.397***	0.333***
	(0.202)	(0.121)

Notes. Data in (A) come from teacher surveys and document inspection conducted in the respective studies, each case restricted to teachers in the treatment arm. "Graded reading booklets" and "lesson plans" are resources provided by the program. See Table 2 in Cilliers et al. (2020) and Figure 1 in Cilliers et al. (2022b) for more details. Data from (B) come from classroom observations conducted in a random sample of 40 teachers in each study: 20 teachers in the treatment and control groups, respectively. Statistics reported are the coefficient of a regression of the dependent variable on treatment, including strata fixed effects. Standard errors are in parentheses. $^{***}p < 0.01$. See Table 6 in Cilliers et al. (2020) and Table 3 in Cilliers et al. (2022b) for more details.

standard deviations, by the end of the first and second years, respectively. EGRS II had the desired positive impact on both English language and literacy skills, by the end of the third year. The positive effect on English language skills was already present at the end of the first year of the program, which is consistent with the fact that English vocabulary is heavily emphasized in the grade 1 L2 curriculum.

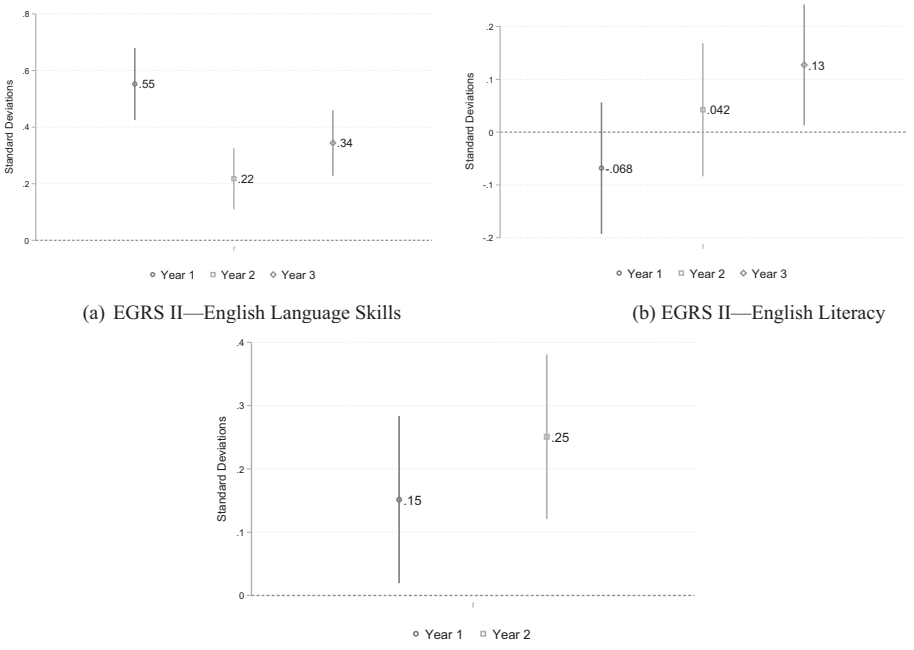
Main Results: Language Transfer

Our main results are reported in Table 5. We focus on grade 4 results when the assessment instruments are most comparable. There is a striking asymmetry in language transfer, depending on which language is targeted. Targeting the home language caused *positive* spillovers onto English literacy (Panel A, columns (3) and (4)), but targeting English caused a *reduction* in home language literacy (Panel B, columns (1) and (2)). For example, EGRS I caused a 9% increase in English ORF, but EGRS II caused an 11% reduction in home language ORF compared to the control group. Moreover, the intervention targeting home language instruction had a *larger* impact on English literacy than the intervention targeting English instruction. Notably, all of these outcomes are observed one year after the respective programs ended, so they suggest persistence in learning gains/loss.

Quantile Regressions

Figure 3 shows quantile regression for EGRS II at the end of grade 3—the final year of program implementation. Figure 3a shows that targeting L2 actually had a

for the literacy scores. There were no floor effects in home language literacy, but there were floor effects in English reading proficiency.



(a) EGRS II—English Language Skills

(b) EGRS II—English Literacy

(c) EGRS I—Home Language Literacy

Figure 2. Treatment effects, by study and year. (a) EGRS II—English Language Skills. (b) EGRS II—English Literacy. (c) EGRS I—Home Language Literacy. *Notes*. Coefficient estimates and 90% confidence intervals for $\hat{\beta}_1$, estimated using Equation (1). Separate regressions are run for each study, outcome, and year. The dependent variables are aggregate scores constructed using principal component analysis, and standardized to have a control mean of one and standard deviation of zero. See Table 2 for the constituent indicators for the literacy scores. The indicators for English Language are (i) productive vocabulary and (ii) English comprehension.

positive significant effect on L1 letter recognition, for students in the top half of the performance distribution. This is evidence of positive spillovers in basic decoding skills even though there were negative spillovers in higher-level L1 reading proficiency. But there are stark inequalities in who benefited or suffered from the EGRS II intervention. Only students in the top half of the distribution improved their English literacy skills (Panel B), and only students in the bottom half of the distribution experienced a reduction in L1 ORF (Panel A).[13] This pattern of result is not present in EGRS I.

Mediation Analysis

As a final step, we perform mediation analysis on the EGRS I sample, to determine how much of the improvements in English in grade 4 can be attributed to earlier gains in home language literacy. We apply the methods proposed by Imai et al. (2010), which allows one to decompose the overall effect size on English literacy between an indirect effect that

[13]Note that there is no effect, positive or negative, for the bottom quintile of students because of floor effects: in both the treatment and control arms, the bottom fifth of students cannot read a single word in their home language.

Table 5. Impacts on home language and English literacy at the end of grade 4.

	(1)	(2)	(3)	(4)
	Home language		English	
	ORF	Reading compr.	ORF	Reading compr.
(A) Improving home language (L1) instruction (EGRS I)				
Treatment	7.159***	0.058***	3.561*	0.024*
	(1.972)	(0.015)	(2.026)	(0.013)
Control mean	47.357	0.298	39.131	0.166
Observations	1846	1846	1846	1846
R-squared	0.176	0.163	0.157	0.128
(B) Improving English second language (L2) instruction (EGRS II)				
Treatment	−2.774***	−0.033*	−1.151	0.030*
	(1.024)	(0.019)	(2.025)	(0.017)
Control mean	25.093	0.448	36.480	0.230
Observations	1729	1729	1729	1729
R-squared	0.284	0.220	0.253	0.237

EGRS: early grade reading study; ORF: oral reading fluency.

Notes. Each column in each panel is a separate regression, estimated using Equation (1). (A,B) Use data from the ERGS I and II, respectively. Data collection took place one year after the program ended, four years after the start of the program, when non-repeating students in our sample were in grade 4. The dependent variables in the first two columns relate to home language literacy; the dependent variables in the remaining columns relate to English literacy. Standard errors are in parentheses and clustered at the school level. All estimations include strata fixed effects and the following control variables: different baseline measures of student home language literacy (phonemic awareness, letter recognition, etc.), student gender and age, and some measures of school socio-economic status, $^{*}p < 0.1$ and $^{***}p < 0.01$.

operates through a mediating variable (i.e., improvements in English that are due to improvements in L1), and a direct effect (i.e., improvements in English that cannot be attributed to improvements in L1). The estimating equations are shown in Section Empirical Strategy, and full analyses can be found in Section A.2 in the Supplementary Appendix.

Results are shown in Table 6. Columns (1) and (2) show the direct and indirect effects, respectively. The mediating variables are different indicators for home language literacy, measured at the end of grade 2. For comparison, Column (3) shows the overall effect of the program on Grade 4 English ORF, restricting the sample to the same learners who were assessed in both grades 2 and 4. These results suggest that *all* of the observed gains in English are explained by the earlier gains in home language literacy. In fact, column (1) suggests that there would have been a *negative* impact on English literacy, were it not for the improvements in word recognition, ORF, or comprehension. We cannot state with certainty which of these different indicators for home language literacy are most important for the development of English literacy, because they are all highly correlated with each other and fundamentally relate to the same underlying construct.

Discussion and Mechanisms

Many studies have shown correlations between literacy measures in L1 and L2, but have lacked a way to identify causal relationships between L1 skills and L2 skills. In this article, we have been able to measure the causal impact of L1 skills on L2 reading outcomes, and the causal impact of L2 skills on L1 reading outcomes. The experimental design means that we can rely on an externally caused improvement in one language to

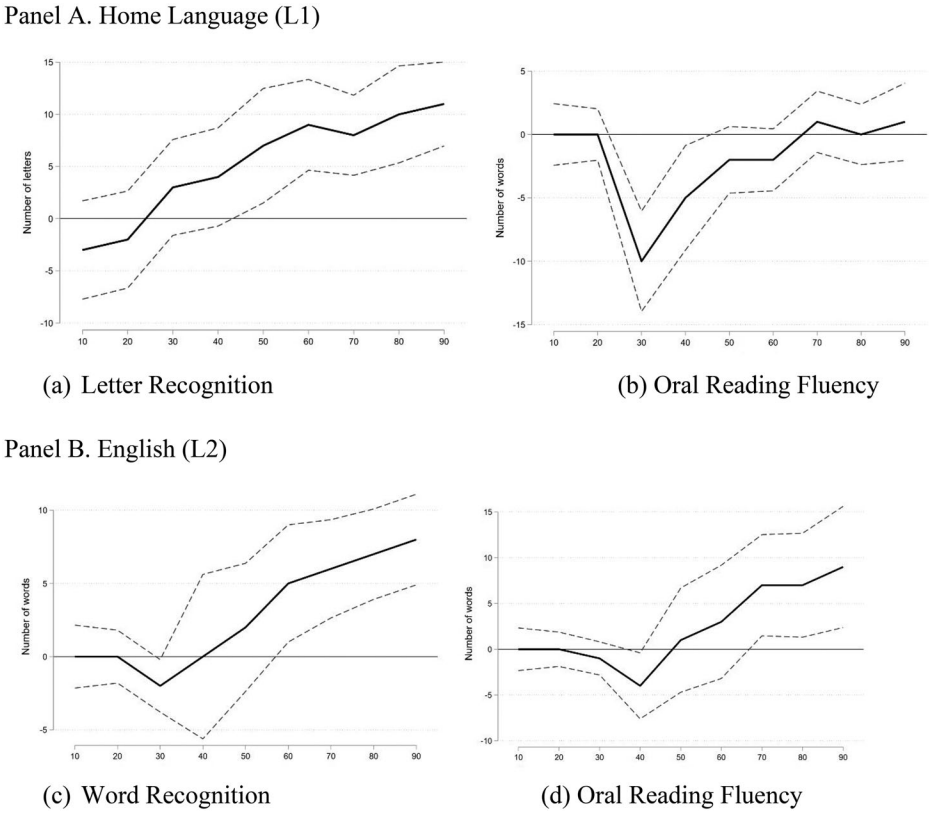


Figure 3. Quantile regressions for ERGS II on grade 3 literacy outcomes. (A) Home Language (L1). (a) Letter Recognition. (b) Oral Reading Fluency. (B) English (L2). (c) Word Recognition. (d) Oral Reading Fluency. *Notes*. Quantile regressions for each decile of student performance. In (A) the dependent variables are home language letter recognition and Oral Reading Fluency (ORF); in (B) the dependent variables are English word recognition and ORF.

Table 6. Mediating effects of grade 2 home language literacy on grade 4 English ORF.

Mediator (M)	(1) Direct effect $\hat{\beta}_1 - \hat{\gamma}_1$	(2) Indirect effect $\hat{\gamma}_1$	(3) Overall effect $\hat{\beta}_1$
Letter recognition	0.028	3.785	3.813
Word recognition	−2.083	5.896	3.813
ORF	−2.203	6.016	3.813
Paragraph compr.	−1.486	5.299	3.813

ORF: oral reading fluency.
Notes. The overall effect is estimated using Equation (1). The indirect is measured using equation (2). The direct effect is the difference between the two.

identify the impact of reading skills in that language on reading skills in another language. Moreover, we have two reciprocal experiments allowing us to observe the causal relationships between L1 and L2 skills in both directions. The close similarity in research and program design between EGRS 1 and EGRS 2 means that we are able to rule out many possible reasons for different results between the two experiments, other than the nature of language transfer between L1 and L2.

Ruling Out Competing Mechanisms

We are able to rule that differences in the quality of program implementation, or spillovers in teaching practices or teaching time, explain the results. First, we show in section Data Analysis above that the quality of implementation was high and remarkably similar in both studies. Second, the negative effect on home language literacy in EGRS II is unlikely due to crowding out of teaching time. We asked teachers at the end of grade 3 how many hours they dedicate to teaching home language literacy. There was a small, statistically insignificant decrease of 13 min in the amount of time that treated teachers reported to spend teaching home language in a week (~156 s a day), relative to a control mean of 7 h. This is unlikely to explain the magnitude of the observed negative effects. We also do not see any reduction in mathematics, providing further evidence that crowding out of teaching time does not explain the results.[14]

Third, although it is theoretically possible that the improvements in English reading skills in EGRS I were due to a transfer of teaching skills from one subject to another, it is unlikely to explain the magnitude of the results, for four reasons. First, if general teaching practices improved one would expect to see improvements in all the subjects taught by the treated teachers, but in both EGRS I and II there were no positive spillovers in mathematics. Second, these are bundled interventions combining coaching with the provision of resources, such as lesson plans and reading aids. So even if teachers in EGRS I improved their general pedagogy, they would have been unable to fully apply them in the English class without the additional learning aids. Third, positive spillover would require teachers to be familiar with the differing expectations of the curriculum for L1 and L2 and be able to navigate through these successfully. Fourth, if generic improvements in teaching practice prompted by an L1 structured learning program caused improvement in both L1 and L2 in EGRS 1, then one would also expect it to apply the other way around in EGRS II, which is not the case.

This combination of evidence leads us to conclude that the mechanism operates through a transfer of skills at a *student* level, not a teacher level.

Theoretical Explanations

The results of this study are consistent with the simple view of reading (SVR), which argues that reading comprehension requires *both* strong decoding skills and oral vocabulary skills. A weakness in either will lead to a weakness in reading for meaning. In the case of EGRS I, the gains in English reading fluency and comprehension are likely to have been influenced by the improved decoding skills in Setswana that were transferred to the L2. Most learners had sufficiently high levels of oral language proficiency in L1 and thus improved their word decoding skills in their L1 when teachers were supported by a structured pedagogical program. The mediation analysis confirms that almost all of the gains seen in English came from learning to decode in L1. The findings suggest that all of the SVR assumptions and conditions were met for L1 instruction, and the curriculum was based on correct assumptions.

[14]In South Africa one teacher teachers all the subjects at early grade. So, the same grade 1 teacher teaches Home Language, English and Mathematics.

In the case of EGRS II, the focus of L2 was mostly on oral vocabulary, and students improved their vocabulary skills.[15] However, because L2 teaching did not stress the teaching of decoding skills, and L1 decoding skills were also weak, there were limited improvements in reading fluency. This is evidenced by the fact that only the high-performing students improved their L2 reading skills. In other words, only those with a sufficient grasp of decoding skills as acquired through L1 seem to have benefited from the L2 intervention, which complemented those decoding skills with improved English vocabulary. The fact that there was no positive spillover from L2 to L1 reading fluency and comprehension is also consistent with the SVR since there were limited improvements in L2 decoding skills. There were positive spillovers in L1 letter sound recognition for the top-performing skills, likely because letter sounds are written in the same way across both L1 and L2.

The negative impact of EGRS II on L2 ORF and reading comprehension for the students in the bottom half of the distribution might be due to the fact that teachers do not have sufficient knowledge of the orthographic rules for both L1 and L2. As discussed above, there is a large linguistic distance between indigenous South African languages and English. EGRS II teachers might have applied the same sequencing of decoding skills—i.e., the same sounding out syllables and letter blends—that they received for L2 to their L1 classes. But this is the wrong starting point, given the different orthographic rules in the different languages. This could have confused students who are already struggling with basic decoding skills. We can only speculate on this interpretation, and more research is required to fully understand the result.

It is possible that the differences in morphology between the various home languages in the two studies account for some of our results. In particular, the home languages spoken in the L2 intervention, Siswati and isiZulu, have a conjunctive morphology, while the home language spoken in the L1 intervention, Setswana, possesses a disjunctive morphology that is more similar to English. Plausibly a positive transfer of ORF is easier when the orthographies of the two languages are similar. Consequently, we might have observed positive (or lack of negative transfer) transfer in ORF from L2 to L1 in EGRS II if the home language were Setswana.

Nevertheless, three pieces of empirical evidence and theoretical considerations suggest that the differences in morphology between the home languages are not sufficient to fully explain the results. Firstly, there is evidence of negative transfer from L2 to L1 ORF. Secondly, the English literacy skills of the control groups in both samples are remarkably similar. Table 3 demonstrates that there is no statistically significant distinction between the two samples regarding English ORF at grade 4, and, in fact, the EGRS II sample performed better in the comprehension test. Thus, the evidence does not support the notion that the transfer of ORF skills between L1 and L2 is easier when the home language is Setswana. Thirdly, students tend to develop decoding skills at a slower pace in English due to the cognitive load of learning both a new language and decoding skills. This is substantiated by the notable discrepancy in the impact on L1 literacy in EGRS I compared to the impact on English literacy in EGRS II (refer to Figure 2).

[15]As discussed above, the L1 curriculum places a large emphasis on mastering decoding in grade 1, the L2 curriculum emphasises oral vocabulary. The assumption is that students already have a basic command of the oral vocabulary in their L1, and that the decoding skills they learn in L1 will transfer to L2.

Considering the lesser impact on L2 decoding skills, one would anticipate a smaller transfer of L1 decoding skills as well.

Conclusion

Over the past decade evidence on the impact of structured learning programs has included contributions from developing countries. However, one critical problem relates to the subject and language sequencing, and the extent to which intervention designed to improve instructional practices in one school subject and/or language would transfer or spill over into other subjects and language. This study examines the question of reciprocal language transference between L1 and L2 in the African context, taking advantage of two large-scale randomized evaluations of structured pedagogy programs in South Africa. We have a unique opportunity to understand the nature of language transfer between L1 and L2, in both directions, since we have two interventions with the same modality, dosage, and service provider, but one targeting L1 and one targeting L2.

We find that in both studies the structured learning program was successful at changing teaching practice and improving literacy in the targeted language. However, it was only in the case of the L1 intervention that a positive spillover to L2 was observed. When the program targeted English as the L2, there were gains in English language and reading skills, but these gains were modest and concentrated amongst the top half of the performance distribution. Furthermore, despite some initial positive effects on letter sound recognition in L1, there were no positive spillovers to L1 reading fluency and comprehension, and ultimately a negative impact on these L1 reading skills for students in the bottom half of the performance distribution. The results can at least be partly explained by, and support, the SVR.

Taken together, these results suggest that decoding skills are best taught in the L1 since children already possess sufficient oral language comprehension. Moreover, decoding skills are more easily transferable across languages with similar orthographies, whereas oral language skills, such as vocabulary, do not transfer in the same way. Furthermore, teaching of decoding skills in L2 may worsen *students'* decoding skills in L1, especially if there is a large orthographic distance between languages and students have not sufficiently mastered L1 decoding skills.

This has important policy implications for multilingual settings where children do not enter school with sufficient prior exposure to the L2. First, students should be taught in their home language, especially in the early grades. Second, programs aimed at improving early grade reading should not prioritize L2 literacy instruction. Our studies did not include a program targeted at both L1 and L2, so we do not know whether such an intervention would be more or less cost-effective. But if resource constraints mean that governments or implementing organizations can only intervene in one language, then L1 should be prioritized.

One caveat is that our results are drawn from two experiments in two different populations, and not one experiment in one population. Even though the interventions themselves were almost exactly equivalent, the different populations might have had different responses to the treatment. We show in the article that the two provinces are very similar in terms of socio-economic status and education outcomes, and that both studies

sampled the same type of schools. But a key difference is the students' home language. Theoretically, the extent of language transferences from L2 to L1 might depend on the degree of similarity between the two languages. These are only conjectures, of course, and future studies comparing L1 and L2 interventions within the same language or the same language group would provide further insights.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Open Research Statements

Study and Analysis Plan Registration

There is no study and analysis plan registration associated with this manuscript.

Data, Code, and Materials Transparency

The data and code underlying the results reported in this manuscript are openly available in Harvard Dataverse: https://doi.org/10.7910/DVN/RSUADW.

Design and Analysis Reporting Guidelines

This manuscript is not accompanied by a completed JREE Randomized Trial Checklist.

Transparency Declaration

The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Replication Statement

This manuscript reports an original study.

Open Scholarship

This article has earned the Center for Open Science badge for Open Data. The data are openly accessible at https://doi.org/10.7910/DVN/RSUADW.

References

Alexander, N. (2005). Language, class and power in post-apartheid South Africa. *Harold Wolpe Memorial Trust Open Dialogue Event*.

Ardington, C., Wills, G., Pretorius, E., Mohohlwane, N., & Menendez, A. (2021). Benchmarking oral reading fluency in the early grades in Nguni languages. *International Journal of Educational Development*, *84*, 102433. https://doi.org/10.1016/j.ijedudev.2021.102433

Baker, D. L., & Stoolmiller, M. (2011). Effect of reading comprehension on passage fluency in Spanish and English for second-grade English learners. *School Psychology Review*, *40*(3), 331–351. https://doi.org/10.1080/02796015.2011.12087702

Ball, J. (2011). Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in the early years. *UNESCO, Education Sector*, 87.

Branum-Martin, L., Mehta, P. D., Fletcher, C. D., Carlson, A., Ortiz, M., Carlo, M., & Francis, D. J. (2006). Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergarteners in transitional bilingual education classrooms. *Journal of Educational Psychology*, *98*(1), 170–181. https://doi.org/10.1037/0022-0663.98.1.170

Casale, D., & Posel, D. (2011). English language proficiency and earnings in a developing country: The case of South Africa. *The Journal of Socio-Economics*, *40*(4), 385–393. https://doi.org/10.1016/j.socec.2011.04.009

Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, M., & Taylor, S. (2022a). The challenge of sustaining effective teaching: Spillovers, fade-out, and the cost-effectiveness of teacher development programs. *Economics of Education Review*, *87*, 102215. https://doi.org/10.1016/j.econedurev.2021.102215

Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, N., Taylor, S., & Thulare, T. (2022b). Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning in South Africa. *Journal of Development Economics*, *155*, 102815. https://doi.org/10.1016/j.jdeveco.2021.102815

Cilliers, J., Fleisch, B., Prinsloo, C., & Taylor, S. (2020). How to improve teaching practice?: An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*, *55*(3), 926–962. https://doi.org/10.3368/jhr.55.3.0618-9538R1

Cohen, A. D., & Swain, M. (1976). Bilingual education: The 'immersion' model in the North American context. *TESOL Quarterly*, *10*(1), 45. https://doi.org/10.2307/3585938

Collier, V. P., & Thomas, W. (n.d.). *How Quickly Can Immigrants Become Proficient in School English?* 11.

Colombo, J. (1982). The critical period concept: Research, methodology, and theoretical issues. *Psychological Bulletin*, *91*(2), 260–275. https://doi.org/10.1037/0033-2909.91.2.260

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 222–251. https://doi.org/10.3102/00346543049002222

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, *14*(2), 175. https://doi.org/10.2307/3586312

Curriculum and Assessment Policy Statement (2011). *Curriculum and assessment policy statement English home language. Intermediate phase Grade 4-6 Grade 4-6*. Department of Basic Education.

de Galbert, P. G. (2023). Language transfer theory and its policy implications: Exploring interdependence between Luganda, Runyankole-Rukiga, and English in Uganda. *Journal of Multilingual and Multicultural Development*, *44*(1), 1–19. https://doi.org/10.1080/01434632.2020.1788038

Department of Basic Education (2021). *National Senior Certificate 2020 examination report*.

Dutcher, N., & Tucker, R. (1995). *The use of first and second languages in education. A review of international experience* [Pacific Island Discussion Paper Series]. World Bank. Retrieved from http://documents1.worldbank.org/curated/en/131161468770987263/pdf/multi-page.pdf

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2020). *Ethnologue: Languages of the world* (23rd ed.). SIL International.

Feltes, J. M. (2022). *La doble inmersión en dos lenguas nacionales: Una metodología pedagógica para el desarrollo de habilidades de biliteracidad crítica y el diálogo intercultural para*

estudiantes de escuelas primarias indígenas de México. Universidad Iberoamericana AC. Retrieved from https://www.scielo.org.mx/pdf/sine/n43/n43a5.pdf

Fleisch, B. (2018). *The education triple cocktail: System-wide instructional reform in South Africa* (1st ed.). Juta.

Fleisch, B., Taylor, S., Schöer, V., & Mabogoane, T. (2017). Failing to catch up in reading in the middle years: The findings of the impact evaluation of the Reading Catch-Up Program in South Africa. *International Journal of Educational Development*, *53*, 36–47. https://doi.org/10.1016/j.ijedudev.2016.11.008

Giliomee, H. (2004). The rise and possible demise of Afrikaans as public language. *Nationalism and Ethnic Politics*, *10*(1), 25–58. https://doi.org/10.1080/13537110490450764

Giliomee, H. (2009). A note on Bantu education, 1953 to 1970. *South African Journal of Economics*, *77*(1), 190–198. https://doi.org/10.1111/j.1813-6982.2009.01193.x

Ginkel, A. (2014). *Additive language learning for multilingual settings*. USAID.

Gordon, S. L., & Harvey, J. (2019). Choice of language in education: Do we know what South Africans want? *Language and Education*, *33*(3), 226–243. https://doi.org/10.1080/09500782.2018.1488865

Goswami, U., Gombert, J. E., & Barrera, L. F. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics*, *19*(1), 19–52. https://doi.org/10.1017/S0142716400010560

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

Gustafsson, M. (2020a). *A revised PIRLS 2011 to 2016 trend for South Africa and the importance of analysing the underlying microdata* [Working Paper of the Department of Economics and the Bureau for Economic Research]. Stellenbosch University.

Gustafsson, M. (2020b). *Teacher demand in South Africa in 2019 and beyond* (No. 1). Department of Higher Education and Training. Retrieved from https://resep.sun.ac.za/wp-content/uploads/2022/03/DHET-Supply-and-Demand-Report-Phase-1-1.pdf

Hasbrouck, J., & Tindal, G. (2017). *An update to compiled ORF norms* (Technical No. 1701). Behavioral Research and Teaching University of Oregon. Retrieved from https://intensiveintervention.org/sites/default/files/TechRpt_1702ORFNorms%20FINAL.pdf

Howie, S., Combrinck, C., Roux, K., Tshele, M., Mokoena, G., & Palane, N. M. (2017). *Progress in International Reading Literacy Study 2016* (p. 220). Centre for Evaluation and Assessment, Faculty of Education, University of Pretoria.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10.1037/a0020761

Khumalo, J. S. M. (1987). *An autosegmental account of Zulu Phonology*, 82.

Kim, Y.-S G., & Piper, B. (2019). Cross-language transfer of reading skills: An empirical investigation of bidirectionality and the influence of instructional environments. *Reading and Writing*, *32*(4), 839–871. https://doi.org/10.1007/s11145-018-9889-7

Kirkpatrick, A. (2012). English in ASEAN: Implications for regional multilingualism. *Journal of Multilingual and Multicultural Development*, *33*(4), 331–344. https://doi.org/10.1080/01434632.2012.661433

Koda, K., & Reddy, P. (2008). Cross-linguistic transfer in second language reading. *Language Teaching*, *41*(4), 497–508. https://doi.org/10.1017/S0261444808005211

Kosonen, K. (2017). *Language of instruction in Southeast Asia. Background paper prepared for the 2017/8 Global Education Monitoring Report, Accountability in education: Meeting our commitments*.

Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. *International Journal of Educational Development*, *66*, 203–213. https://doi.org/10.1016/j.ijedudev.2018.09.004

Le Cordeur, M. (2013). Afrikaans as medium of instruction within a transformed higher education system in South Africa with special reference to Stellenbosch University. In C. C. Wolhuter (Ed.), *Conference Proceedings: Educational Research in South Africa: Practices and Perspectives*, 25–30 January (pp. 55–76). Oxford University Press.

Lee, H., Jung, G., & Lee, J. H. (2022). Simple view of second language reading: A meta-analytic structural equation modeling approach. *Scientific Studies of Reading*, *26*(6), 585–603. https://doi.org/10.1080/10888438.2022.2087526

Lenneberg, E. (1967). *Biological foundations of language* (Vol. 13). Wiley. https://doi.org/10.1080/21548331.1967.11707799

Lenneberg, E. (1969). On explaining language. *Science*, *164*(3880), 635–643. https://doi.org/10.1126/science.164.3880.635

Macdonald, C. (1990). *Crossing the threshold into standard three.pdf* [Threshold Project].

Machobane, M., & Mokitimi, M. (1998). Problems in the development of Sesotho orthography. In K. K. Prah (Ed.), *Between distinction and extinction: The harmonisation and standardisation of African language* (pp. 203–212). Wits University Press.

Madadzhe, R. N. (2019). Using African languages at universities in South Africa: The struggle continues. *Stellenbosch Papers in Linguistics Plus*, *58*, 205–218. https://doi.org/10.5842/58-0-843

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*(1), 114–135. https://doi.org/10.1111/j.1467-9817.2010.01477.x

Mohohlwane, N. L. (2019). How language policy and practice sustains inequality in education. In *South African Schooling: The enigma of inequality*: *A study of the present situation and future possibilities* (pp. 127–146).

Mohohlwane, N., Wills, G., & Ardington, C. (2022). A review of recent efforts to benchmark early reading skills in south African languages. In *Early grade reading in South Africa* (Vol. 1, pp. 83–108). Oxford University Press.

Nakamura, P. R., De Hoop, T., & Holla, C. U. (2019). 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

Ngcobo, M. (2009). The politics of compromise and language planning: The case of South Africa. *Language Matters*, *40*(2), 205–214. https://doi.org/10.1080/10228190903188575

Nudelman, C. (2015). *[Language in South Africa's higher education transformation: A study of language policies at four universities]* [Masters]. University of Cape Town.

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*. UNESCO Institute for Lifelong Learning.

Padilla, A. M., Fairchild, H. H., & Valadez, C. M. (Eds.). (1990). *Bilingual education: Issues and strategies*. Sage Publications.

Parkin, M., Morrison, F. C., & Watkin, G. (1987). *French immersion research relevant to decisions in Ontario*. Ministry of Education: Publications Sales, Ontario Institute for Studies in Education.

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. https://doi.org/10.1016/j.ijedudev.2017.10.002

Prah, K. K., & Brock-Utne, B. (2009). *Multilingualism: An African advantage: A paradigm shift in African language of instruction policies*. Centre for Advanced Studies of African Society.

Quinn, J. M., & Wagner, R. K. (2018). Using meta-analytic structural equation modeling to study developmental change in relations between language and literacy. *Child Development*, *89*(6), 1956–1969. https://doi.org/10.1111/cdev.13049

Rodriguez-Segura, D. (2022). A closer look at reading comprehension: Experimental evidence from Guatemala. *International Journal of Educational Development*, *93*, 102630. https://doi.org/10.1016/j.ijedudev.2022.102630

Schaefer, M., & Kotzé, J. (2019). Early reading skills related to Grade 1 English Second Language literacy in rural South African schools. *South African Journal of Childhood Education*, *9*(1), 1–13. https://doi.org/10.4102/sajce.v9i1.644

Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and Speech*, *20*(4), 357–365. https://doi.org/10.1177/002383097702000407

South African Broadcasting Services (2020). *South African Broadcasting Services Annual Report 2020* (p. 172).

South African School Act, 1996 (Act No 84 of 1996) (2022). Amended National Norms and Standards for School Funding, Pub. L. No. 1730, 45878.

Spires, H. A., & Donley, J. (1998). Prior knowledge activation: Inducing engagement with informational texts. *Journal of Educational Psychology*, *90*(2), 249–260. https://doi.org/10.1037/0022-0663.90.2.249

Statistics South Africa (2012). *Census 2011: Census in brief*.

Takakuwa, M. (2005). *Lessons from a paradoxical hypothesis: A methodological critique of the threshold hypothesis*, 12.

UNESCO (2003). *Education in a multilingual world: UNESCO education position paper* (p. 35).

Van der Berg, S., Gustafsson, M., & Burger, C. (2020). *School teacher supply and demand in South Africa in 2019 and beyond* (No. 1). Department of Higher Education and Training. https://resep.sun.ac.za/wp-content/uploads/2022/03/DHET-Supply-and-Demand-Report-Phase-1-1.pdf

van Staden, S., & Bosker, R. (2014). Factors that affect South African Reading Literacy Achievement: Evidence from prePIRLS 2011. *South African Journal of Education*, *34*(3), 1–9. https://doi.org/10.15700/201409161059

Veii, K., & Everatt, J. (2005). Predictors of reading among Herero–English bilingual Namibian school children. *Bilingualism: Language and Cognition*, *8*(3), 239–254. https://doi.org/10.1017/S1366728905002282

Walter, S. (2010). *Mother tongue-based education in developing countries: Some emerging insights*, 26.

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* (1st ed., pp. 278–300). Cambridge University Press. https://doi.org/10.1017/CBO9780511979026.017

Wang, M., Park, Y., & Lee, K. R. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Educational Psychology*, *98*(1), 148–158. https://doi.org/10.1037/0022-0663.98.1.148

Wang, M., Yang, C., & Cheng, C. (2009). The contributions of phonology, orthography, and morphology in Chinese–English biliteracy acquisition. *Applied Psycholinguistics*, *30*(2), 291–314. https://doi.org/10.1017/S0142716409090122

Wawire, B. A., & Kim, Y.-S G. (2018). Cross-language transfer of phonological awareness and letter knowledge: Causal evidence and nature of transfer. *Scientific Studies of Reading*, *22*(6), 443–461. https://doi.org/10.1080/10888438.2018.1474882

Wright, L. (2002). Language as a 'resource' in South Africa: The economic life of language in a globalising society. *English Academy Review*, *19*(1), 2–19. https://doi.org/10.1080/10131750285310031