Can Multilingual Environments Promote Scalable EdTech? Evidence from a Randomized Controlled Trial

Phenyo Phemelo Moletsane
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
pmoletsa@andrew.cmu.edu

Christine Kwon
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
ckwon2@andrew.cmu.edu

John Stamper
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
jstamper@cs.cmu.edu

Amy Ogan
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
aeo@andrew.cmu.edu

Paulo F. Carvalho
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
pcarvalh@andrew.cmu.edu

Abstract

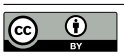
At least half of the world's population uses more than one language daily, and many learners engage with educational content in a language that differs from their primary language. Educational technologies are uniquely positioned to address this linguistic diversity by delivering equivalent instruction in multiple languages, thereby reaching large multilingual populations. However, for this potential to translate into meaningful learning at scale, EdTech systems must support multilingual instruction effectively. To date, there is limited empirical causal evidence on how the language of instruction influences learning outcomes and processes in digital learning environments. In this study, we conducted a small-scale randomized controlled trial to examine the causal impact of instructional language on learning in a radio- and mobile-based engineering course, providing insights relevant to scaling learning in multilingual contexts. We compared learning with instruction in Leb Lango–a predominant local language used by the participants in this study–and in English, one of the country's official languages. Our findings show that learners who received instruction in Leb Lango achieved significantly higher aggregated performance than those who received instruction in English. Instruction in Leb Lango enabled learners to engage more efficiently with the course material and progress through a greater amount of practice content. These findings underscore the role of language in shaping learning and engagement in digital learning systems and suggest that, to maximize learning outcomes at scale, educational technologies should provide instruction and practice opportunities not only in English but also in learners' home languages.

CCS Concepts

• **Applied computing** → **Distance learning**; • **Human-centered computing** → **Empirical studies in HCI**.

Keywords

Multilingual EdTech; Language of Instruction; Mobile Learning; Inclusive EdTech



L@S '26, Seoul, Republic of Korea

ACM ISBN 979-8-4007-2293-6/2026/06
https://doi.org/10.1145/3774398.3811600

ACM Reference Format:
Phenyo Phemelo Moletsane, Christine Kwon, John Stamper, Amy Ogan, and Paulo F. Carvalho. 2026. Can Multilingual Environments Promote Scalable EdTech? Evidence from a Randomized Controlled Trial. In *Proceedings of the Thirteenth ACM Conference on Learning @ Scale (L@S '26), June 29-July 03, 2026, Seoul, Republic of Korea.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3774398.3811600

1 Introduction

The world is linguistically diverse, with more than 7,000 languages spoken worldwide [24]. At least half of the world's population uses more than one language daily [6, 13], and approximately 40% of learners worldwide lack access to education in a language they speak or understand [11, 40]. The difficulties of not learning in one's primary language are well documented in the literature and are often associated with reduced comprehension [2, 5, 36, 42]. Learners in such contexts may face a double burden: developing academic language proficiency, while simultaneously mastering subject-matter content [12]. In-person instruction cannot feasibly accommodate all learners' home languages, as teachers cannot realistically teach in multiple languages simultaneously, and curricula typically lack the flexibility needed to provide differentiated multilingual support. Thus, if instruction is to reach the millions of multilingual learners, it cannot depend solely on traditional classroom-based delivery. Instead, scalable mechanisms are required. Educational technologies (EdTech) are uniquely positioned to address learners' linguistic diversity by delivering equivalent instruction in multiple languages, including local languages. Prior work in classroom contexts has shown that instruction in local languages fosters more relevant, accessible, and meaningful learning experiences, as well as improved learning outcomes [30, 33, 35]. While the benefits of mother tongue instruction are well established in classroom contexts, there is limited evidence on whether these benefits extend to digital learning environments. Moreover, realizing these benefits in digital environments requires effective multilingual design that incorporates local languages and is informed by how learners interact with multilingual platforms and how instructional language shapes their learning.

Prior research examining the integration of local languages into multilingual EdTech at scale suggests positive correlations with participation among learners from disadvantaged backgrounds, faster learning rates, and greater engagement [29]. However, there remains limited empirical causal evidence on how instructional

language influences learning processes and outcomes in digital learning environments.

In this study, we conducted a small-scale randomized controlled trial to examine the causal impact of the language of instruction on learning in a mobile learning environment in Uganda. We compared learning with instruction in Leb Lango, a predominant local language used by the participants in this study, and in English, one of the country's official languages. We conducted the study in an out-of-school low-tech learning environment that delivers a Science, Technology, Engineering, and Mathematics (STEM) course. We employed a quantitative approach to examine the following research questions:

> **RQ1:** What is the effect of Leb Lango instruction (vs. English instruction) on learners' learning outcomes?
> **RQ2:** What is the effect of Leb Lango instruction (vs. English instruction) on learners' efficiency and persistence in the course?

Our findings indicate that learners who received instruction in Leb Lango outperformed those learning in English in overall performance across course assessments. Beyond learning outcomes, our study reveals how language shapes learning processes: Leb Lango instruction enhanced faster comprehension and more efficient learning (as evidenced by fewer practice attempts), as well as greater persistence (as shown by higher course completion). To the best of our knowledge, this is one of the first studies to establish a causal relation between language of instruction and learning in multilingual EdTech environments. Moreover, the study provides evidence that EdTech can—and should—serve as the ideal platform for realizing the benefits of multilingual instruction tailored to learners' home languages.

2 Background and Related Work

In this section, we review prior work and position our study within the broader literature. We begin by examining current practices regarding instructional language in EdTech, emphasizing learners' linguistic needs. Next, we explore research on linguistically and culturally localized EdTech solutions, highlighting learners' experiences and the benefits of incorporating local languages. Finally, we draw from research on scalable EdTech, such as Massive Open Online Courses (MOOCs), which demonstrates that active, practice-based learning enhances learning outcomes. Together, this review highlights key insights and identifies gaps in the literature, particularly the need for research on learning in multilingual, practice-based EdTech environments and for establishing causal evidence on the impact of language of instruction on learning outcomes within EdTech systems.

2.1 Current Practices of Language of Instruction and Linguistic Inclusion in EdTech

EdTech is well positioned to broaden access to education and expand participation in multilingual education at scale. However, similar to classrooms, most educational technologies currently remain predominantly monolingual, with instructional content largely delivered in global languages such as English, failing to reflect the multilingual realities of learners. Prior research suggests that language of instruction plays a critical role in shaping who can meaningfully engage with these technologies and who is excluded [26, 34]. This dynamic is particularly pronounced in rural and low-resourced contexts, where learners may have fewer opportunities to formal education or practice English than their urban peers, contributing to persistent literacy and comprehension gaps [14]. A growing body of research emphasizes the importance of aligning EdTech with learners' local linguistic and cultural contexts and highlights the challenges learners face when instructional language does not match their primary language [10, 22, 25, 32, 43]. For example, Kukulska-Hulme et al. (2023) [22], drawing on findings from the ReMaLIC project across Bangladesh, Nepal, Senegal, and Sudan, show that learners routinely navigate multilingual environments in their daily lives, yet encounter significant barriers when interacting with technologies that assume English proficiency. Similarly, Liddicoat and Heugh (2014) [25] suggests a close link between English and technology, showing that students with limited English proficiency struggle to interact effectively with digital learning platforms, which assume competence in English.

Several studies have examined EdTech solutions that consider linguistic inclusion and their relationship to learning outcomes and experiences. For instance, Mohammed and Mohan (2013, 2015) [27, 28] explored culturally and linguistically localized EdTech designs by developing intelligent tutoring systems for teaching programming that incorporated English-based Creoles. Their findings suggest that learners preferred localized instructional content and valued the ability to adjust the density of cultural and linguistic context. Similarly, Jantjies and Joy (2012, 2013) [15–17] deployed bilingual mobile learning technologies in South African high schools, allowing learners to access science and mathematics content-including simplified notes, textbook materials, interactive drill questions, and tests - in both English and local languages. Their results showed that most learners tended to use both languages to engage with content and reported that bilingual access supported conceptual understanding. These studies offer important insights into learner experiences; however, they use self-reported data and qualitative analyses, providing limited evidence on actual learning outcomes.

Another line of research has examined learners' motivation and learning outcomes in multilingual contexts. For example, Barak et al. (2016) [4] studied a MOOC offered in both English and Arabic and found that learners across both language versions exhibited similar levels of intrinsic motivation and self-determination. More recently, Moletsane et al. (2026) [29] conducted a quasi-experimental study in Uganda in which learners selected English, a local language (Leb Lango), or a combination of English and local language. Their results indicate that including local languages in multilingual EdTech was associated with broader access for disadvantaged learners. Furthermore, learners who chose the local language option demonstrated higher participation and faster learning rates, even though they initially performed worse and ultimately reached similar performance levels by the end. While informative, these studies remain correlational and do not isolate the causal impact of language of instruction. Taken together, these studies provide valuable insights into how learners navigate multilingual and culturally inclusive EdTech environments. However, to the best of our knowledge, none employ randomized controlled trials (RCTs) to examine the causal effects of language of instruction in EdTech systems.

2.2 EdTech for Learning at Scale

Research has yet to explore the potential for scaling EdTech to multilingual learners. In recent years, Massive Open Online Courses (MOOCs) such as Coursera, Udacity, and EdX have been widely recognized for their ability to broaden access to educational resources at scale [18]. At the same time, a substantial body of learning sciences research further underscores the value of practice-based learning in scalable MOOCs environments [20, 21, 41]. Learning by doing has long been studied in contexts such as MOOCs to understand how different forms of engagement relate to retention and understanding, and it has been associated with improved learning outcomes. For example, in a large-scale implementation integrating OLI modules into a Coursera MOOC, interactive activities—such as quizzes, problem-solving tasks, and feedback-rich exercises—were more strongly associated with learning gains than passive engagement with videos or text [20, 21]. Similarly, evidence from the testing effect literature supports these findings: practice testing improves learning and retention relative to restudying or passive review [1], and engagement with short, feedback-rich exercises enhances exam performance [8, 9]. Together, this evidence suggests that scalable EdTech systems are most effective when they incorporate structured, interactive practice rather than relying solely on lecture-based content. However, most large-scale platforms remain predominantly monolingual and do not reflect the multilingual realities of learners. We draw from these large-scale, practice-based studies and identify an opportunity: integrating multilingual instruction into practice-based EdTech has the potential to further enhance scalability by enabling a broader range of learners to engage effectively with instructional content. Building on insights from large-scale, practice-based platforms, we investigate how these principles translate to a low-tech, multilingual, practice-based learning environment, providing causal evidence on the effects of language of instruction in EdTech.

3 Method

3.1 Study Recruitment and Context

We adopted a randomized controlled trial (RCT) for this study in collaboration with an educational technology company in Northern Uganda, which develops and delivers practical, remote engineering education courses to rural learners in Northern Uganda. The authors of the paper were not involved in the recruitment, consent, or data collection processes. These activities were carried out by the study partners in Uganda. Ethical approval for the study was obtained from a relevant Ugandan ethics review committee.

All participants were recruited from Lira District, a predominantly rural district in Northern Uganda. Lira District is largely populated by members of the Lango ethnic group, whose primary language is Leb Lango, a minority language spoken by approximately 5% of Uganda's population [39]. The national language-in-education policy (LiEP) formally endorses bilingual education, particularly by promoting the use of local languages as instructional languages in schools [3, 31, 38]. In line with this policy, early primary education is delivered in local languages, after which English becomes the dominant medium of instruction from Primary Four through secondary and higher education [3, 31, 38].

Our study compares learners under two language-of-instruction conditions: Leb Lango (local language) and English. A total of 39 learners fully participated in the study, with 20 assigned to the Leb Lango condition and 19 to the English condition. We obtained written consent from youth participants and their parents for course participation. Participants were between the ages of 9 and 24 years old, not currently enrolled in school, and reported comfort in both English and Leb Lango. Of the participants, 69% identified as women and 31% identified as men. In terms of educational attainment, the majority of learners (84.62%) had completed upper primary schooling (P5-P7), 12.82% lower secondary schooling (S1-S4), and 2.56% of learners had completed lower primary schooling (P1-P4). The RCT was conducted in a classroom in which each language group was split into either a morning or afternoon session and assigned to one of two facilitators who facilitated the experiment. We conducted the experiment by bringing participants into a classroom so we could maintain tight controls but students engaged only with the ed tech. Facilitators did not provide instruction and were present only to register attendance and provide technical support. Attendance among initially registered students in the Leb Lango course was lower than anticipated during the first week. To maintain the integrity of the randomized controlled trial, we conducted a second round of recruitment and randomly assigned the newly recruited participants to the Leb Lango condition.

3.2 Course Curriculum and Structure

The curriculum was delivered through a low-tech learning platform designed for resource-constrained contexts, relying on widely available technologies in rural Africa: namely, basic keypad mobile phones and radio broadcasts. All curriculum content - including lessons, assessments, and radio instructions - was identical across the two language conditions and delivered in the respective language. Multiple translators independently produced Leb Lango translations, which were subsequently validated through comparison and consensus-building to ensure consistency across the curriculum. Learners followed radio lessons and interacted with the course using their mobile phones. The program guided learners through the construction of a solar-powered food dehydrator following the steps of the Engineering Design Process (EDP). The EDP was structured into nine ordered steps (or units): Introduction Step: Basics of STEM education; Step 1: Identify; Step 2: Investigate; Step 3: Brainstorm; Step 4: Plan; Step 5: Create; Step 6: Test; Step 7: Improve; Step 8: Launch. Each step consisted of several lessons, with approximately one day allocated to each lesson. These EDP steps were integrated into the weekly lessons, enabling learners to engage in hands-on, problem-solving tasks using locally available materials in their homes and communities. The solar-powered food dehydrator designed to preserve agricultural produce using renewable energy served as the final project. Given the region's reliance on subsistence farming, this project provided a locally relevant application of STEM concepts, and the completed device functioned as a demonstration of both conceptual understanding and practical skills.

3.3 Course Delivery and Activities

The course consisted of radio broadcast lessons, practice questions, weekly tests, a final examination, and a hands-on project. All questions in the course were primarily in multiple-choice format, each with four answer options. During lessons, instructors engaged learners by asking practice questions via the radio, which learners answered using Unstructured Supplementary Service Data (USSD) menus on their phones. USSD is a communication protocol mobile network operators use to send text messages between a mobile phone and an application server for free. It allows learners to submit assessment responses without typing, making it ideal for multiple-choice questions and constrained settings. Learners also completed weekly multiple-choice tests via USSD, aligned with the content covered in the preceding week's lessons. At the end of the program, participants completed a 16-question multiple-choice final assessment delivered through radio instruction and answered via USSD. In addition to assessments, learners design and build an innovative technological product related to the course topic (solar-powered food dehydrator).

4 Analyses and Results

All analyses were conducted in R (Version 4.5.1, R Core Team, 2025).

4.1 What is the effect of Leb Lango instruction (vs. English instruction) on learners' learning outcomes?

We examined differences in course performance by language of instruction. First, we considered overall performance by aggregating learners' responses across all assessed components of the course, including in-lesson practice questions, weekly tests, and the final examination. Performance was measured as the proportion of correct responses on multiple-choice questions. Since learners could make multiple attempts on the same question, we considered only each learner's first attempt. For each learner, we calculated the average first-attempt score across all questions. To estimate the effect of language of instruction, we fitted a linear regression model predicting standardized overall scores as a function of course language (Leb Lango vs. English), controlling for highest completed education level (P1–P4, P5–P7, S1–S4). We then examined performance by assessment type—practice questions, weekly tests, and the final examination. For each assessment type, we calculated the average first-attempt score and fitted a similar linear regression model predicting scores as a function of course language, again controlling for highest completed education level.

Figure 1 summarizes students' average performance by language of instruction. Learners who received instruction in English condition achieved lower overall performance than those who received instruction in Leb Lango (51.3% vs. 56.4%, respectively). This difference was statistically significant in a regression model controlling for prior educational attainment ($b = -0.69$, $p < 0.05$), indicating that learners who received instruction in English scored, on average, 0.69 standard deviations lower than learners instructed in Leb Lango.

When examining performance by assessment type (Figure 2), learners in the English condition scored 0.87 standard deviations lower than learners instructed in Leb Lango on practice questions,

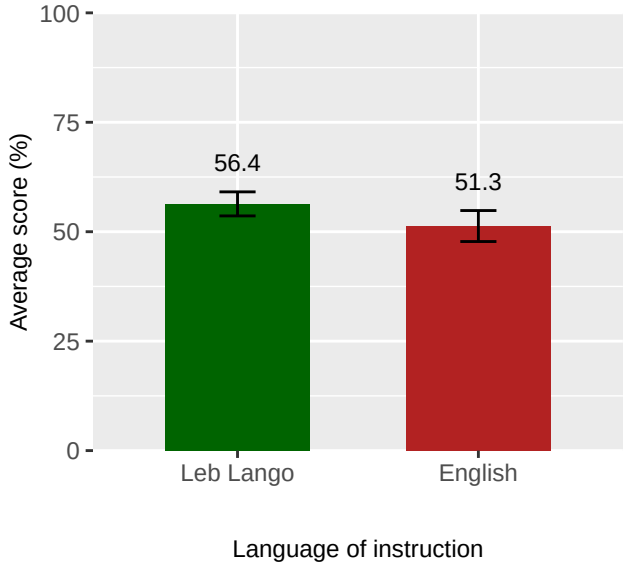


Figure 1: Learners' average overall scores by language of instruction (Leb Lango vs. English) across all course assessments.

and this difference was statistically significant (53.7% vs. 63.6%, $b = -0.87$, $p < 0.05$). In contrast, differences on weekly assessments and the final examination were not statistically significant, with learners in English scoring slightly lower on average (weekly: $b = -0.376$, $p = 0.270$; final examination: $b = -0.448$, $p = 0.222$). In this and all subsequent models, the learners' highest level of education was not significantly associated with performance.

4.2 What is the effect of Leb-Lango instruction (vs. English instruction) on learners' efficiency and persistence in the course?

Given that performance on practice questions differed between the two language groups, we explored our second research question by examining learners' interactions with the course content in greater detail. Specifically, we investigated how the language of instruction was related to learners' efficiency and persistence in the course, focusing on practice questions. For each learner, we first aggregated response data to the question level. We then summarized these data at the learner level by computing two outcome measures: (1) efficiency; and (2) persistence.

We define efficiency as the number of attempts per question, that is, the extent of trial and error during interaction with assessment questions. Fewer attempts indicate higher efficiency, reflecting that learners reached the correct answer with less repeated effort.

Persistence is defined as the number of distinct questions that the learner attempted, capturing how far learners progressed through the course.

For each outcome (efficiency and persistence), we fit a linear regression model predicting student-level outcome as a function of course language (Leb Lango vs. English), while controlling for

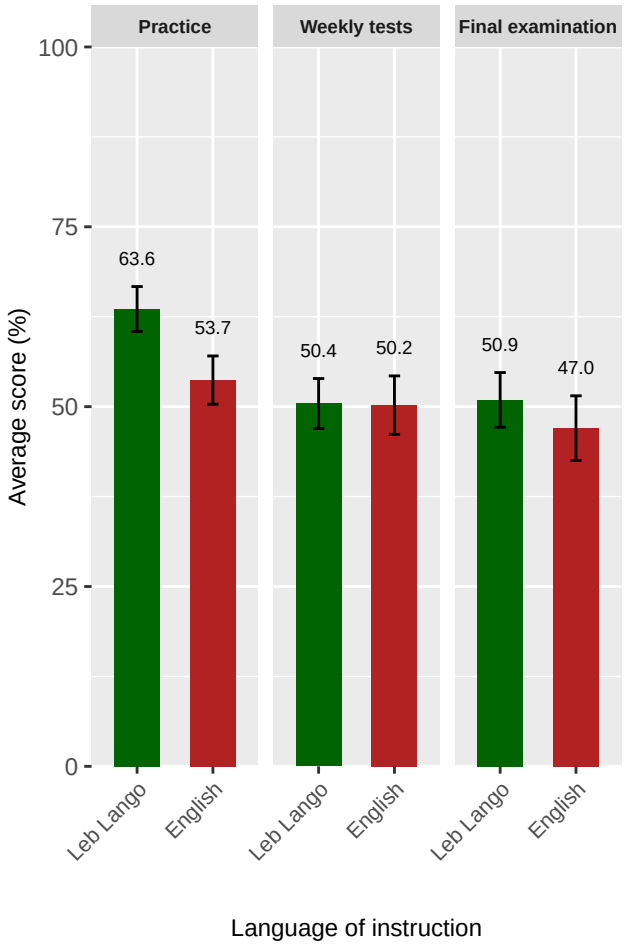


Figure 2: Learners' average scores on practice assessments, weekly tests, and the final examination by language of instruction (Leb Lango vs. English).

learners' highest schooling level (P1-P4, P5-P7, S1-S4). All outcome measures were standardized (z-scored) prior to modeling.

4.2.1 Efficiency. We found clear differences by language of instruction on learner's efficiency. Learners who received instruction in English (mean = 6.47, median = 5.75) were less efficient on questions than those instructed in Leb Lango (mean = 4.07, median = 3.76), and this difference was statistically significant ($b = 1.17$, $SE = 0.31$, $p < 0.001$). The range of attempts was wider for English learners, spanning from about 2.67 to 12.45, while the Leb Lango group ranged from about 2.15 to 7.12 attempts. Several outliers are visible in both groups, with the English group having more extreme outliers exceeding 12 attempts per question (Figure 3). We also controlled for first-attempt score in the efficiency model and found that learners in the English condition were still less efficient than those in the Leb Lango condition ($b = 0.67$, $SE = 0.27$, $p = 0.018$).

We noted that some students in the English group made far more than four attempts on a single question, exceeding the total number of available answer options. This raises questions about

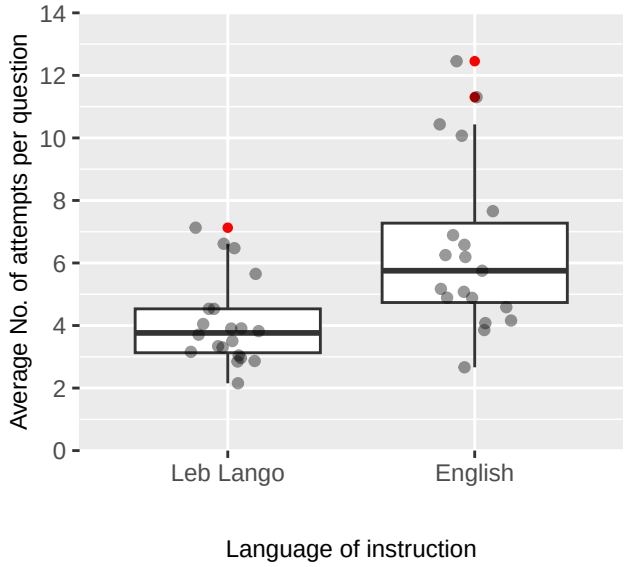


Figure 3: Distribution of learners' average number of attempts per practice question by language of instruction (Leb Lango vs. English).

the factors underlying repeated attempts. One possibility is that learners in the English condition make more attempts to reach the correct answer due to language support needs. Alternatively, repeated attempts could reflect self-regulated learning decisions, such as revisiting questions later for spaced practice or review in preparation for the final assessment. To distinguish between these possibilities, we conducted an exploratory analysis of how attempts were distributed over time. For each learner and question, we classified each attempt based on its timing relative to the first attempt: as a first attempt, a quick retry within the same lesson (less than 12 hours), a repeated lesson attempt after completing the lesson (12-48 hours), or a longer-term revision (more than 48 hours). We then determined the proportion of attempts falling into each category. Figure 4 presents the distribution of attempt types for practice questions by language of instruction.

Among English learners, quick retries accounted for the largest share of attempts (43%), compared to 32% among Leb Lango learners. Repeated lesson attempts represented the largest proportion among Leb Lango learners (35%), and were similar in proportion to English learners (36.6%). First attempts constituted 27% of all attempts in the Leb Lango group and 18% in the English group. Revision attempts were relatively rare in both conditions, though more common among Leb Lango learners (7.6%) than English learners (2.8%). Furthermore, the average number of attempts declined with time for English learners: attempts were highest during quick retries (mean = 2.82), slightly lower during repeated lesson attempts (mean = 2.44) and dropped dramatically for revisions (mean = 0.24). Leb Lango learners showed a flatter early trajectory, with similar numbers of attempts during quick retries (mean = 1.37) and repeated lesson attempts (mean = 1.43), followed by a decline during revision (mean = 0.35). These results are consistent with the possibility that

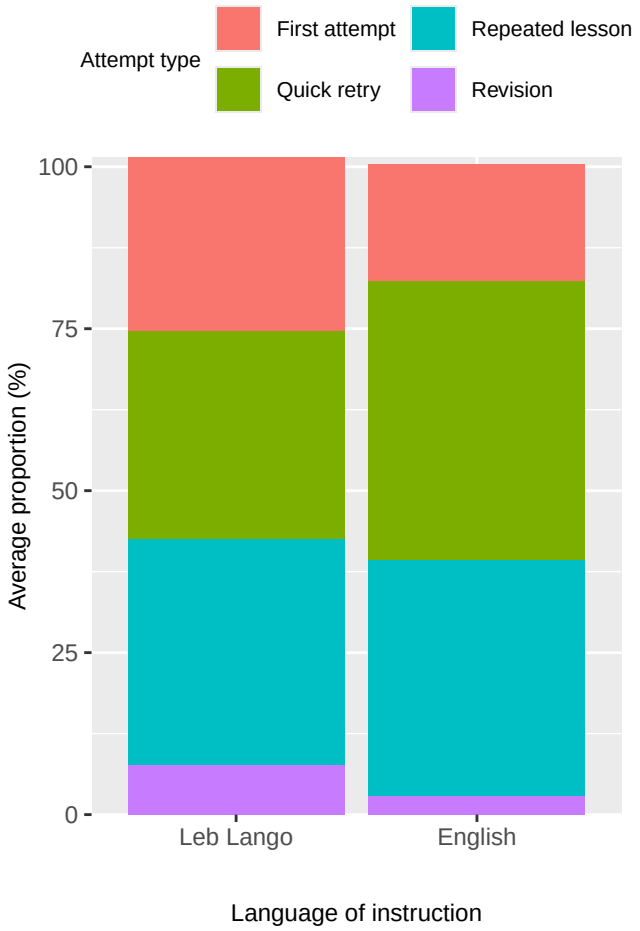


Figure 4: Proportion of attempt types for practice questions by language of instruction.

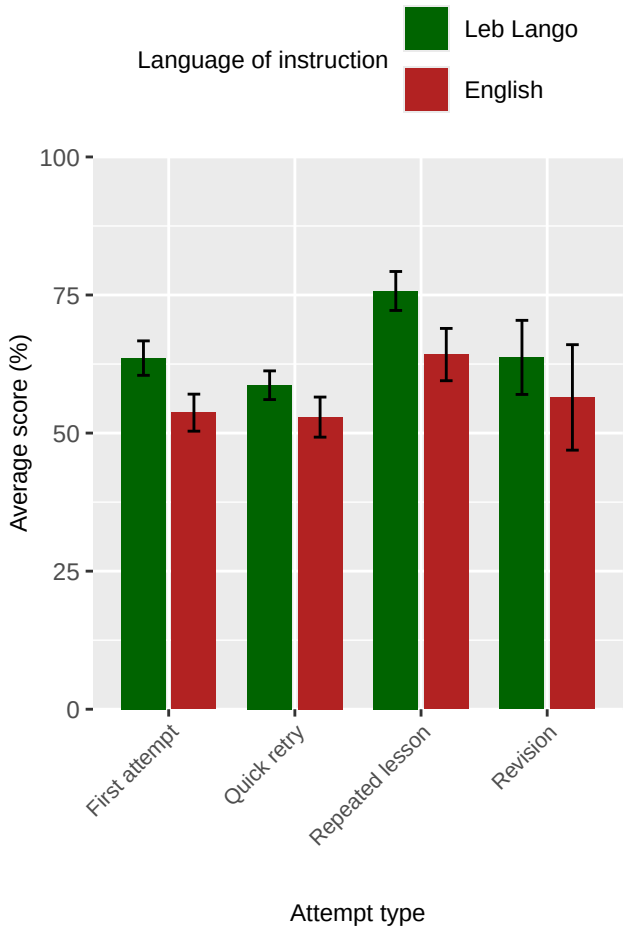


Figure 5: Average score by attempt type and language of instruction on practice questions.

higher attempts among English learners reflect greater reliance on repeated attempts during initial engagement, potentially due to language-related challenges.

Building on the observed differences in attempt trajectories, we next examined whether learners' scores varied by attempt type, considering only the first attempt within each category. This analysis allows us to explore how the learners' scores change with time as they make these attempts. To evaluate this, we calculated the mean proportion of correct responses on the first attempt for each type of attempt, separately by language of instruction. Figure 5 shows the average score by attempt type and language of instruction. Both language groups exhibited a similar temporal pattern in scores across attempt types. Among English learners, first attempt scores was 53.7%, declined to 52.9% on quick retries, increased to 64.2% on repeated lesson attempts, and slightly decreased to 56.5% on revisions. Among Leb Lango learners, correctness was 63.6% on first attempts, dropped to 58.7% on quick retries, increased to 75.7% on repeated lesson attempts, and declined slightly to 63.7% on revisions. In both groups, scores were highest for repeated lesson attempts and lowest for quick retries. However, Leb Lango learners maintained consistently higher scores across all attempt types.

4.2.2 Persistence. Regarding the persistence of the course, we observed differences in progression throughout the course (Figure 6). Learners in the English condition (91 questions) completed less of the course on average than those in the Leb Lango condition (101 questions) ($b = -0.92$, $SE = 0.33$, $p < 0.01$).

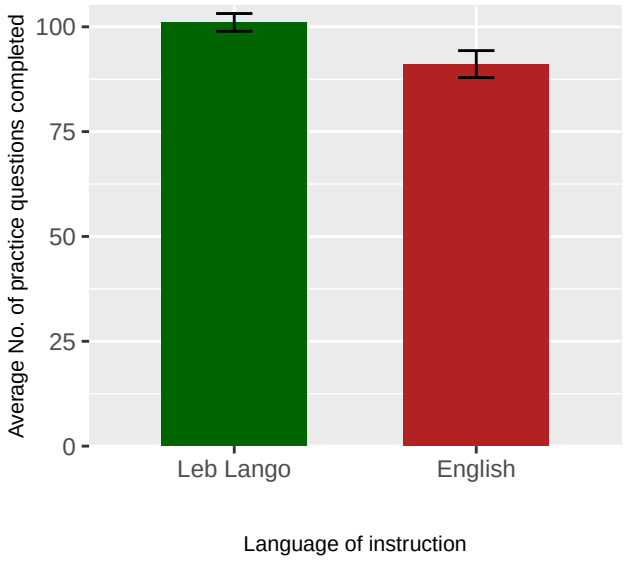


Figure 6: Course persistence by language of instruction.

5 Discussion

We established a causal relationship between the language of instruction and learner performance. Our study also provides insights into how learners engage with course content across the two languages. Below, we discuss the key findings.

5.1 Multilingual EdTech Has Potential to Improve Learning Outcomes and Scale Learning

Our results suggest that, overall, instruction in Leb Lango improved learner performance relative to English instruction. Importantly, these differences are not attributable to variation in learners' formal schooling levels, suggesting that language of instruction plays an independent role in shaping learning outcomes. One plausible explanation for our finding is that instruction in Leb Lango reduced linguistic barriers to learning. For example, by using the local language, learners were able to understand concepts more quickly [23], engage more actively with the content, and connect ideas more effectively [30, 33, 37]. Our findings align with prior qualitative studies comparing Leb Lango and English instruction in EdTech [23], which report that learners and facilitators perceived Leb Lango learners as understanding course content more quickly and accurately than those receiving English instruction, likely due to greater linguistic familiarity. We extend this work by providing causal evidence that these perceived differences translate into measurable differences in learning outcomes. Related work on multilingual delivery in EdTech, based on a study of 2,931 students, shows that instruction in local languages can broaden access and participation among disadvantaged learners [29]. While our study is relatively small in scale, it provides experimental evidence on learning in a multilingual EdTech context and complements prior large-scale correlational findings by showing that integrating local languages into EdTech instruction not only broadens access but can also enhances learning outcomes.

This causal link highlights the potential role of EdTech in supporting learning among multilingual populations at scale. This is particularly important for multilingual learners and learners outside formal schooling systems, who primarily use local languages in their everyday lives. If designed to incorporate learners' home languages, EdTech systems may help expand access to education and promote equity in multilingual contexts where the language of instruction differs from learners' home language.

Our results are also consistent with prior research on local language instruction in classroom settings, which shows that teaching in learners' home languages can support conceptual understanding and foster more meaningful learning experiences, particularly for marginalized students [30, 33, 35]. However, these benefits have not been fully realized in practice, as implementing multilingual education in classrooms is challenging. Even when language-in-education policies formally incorporate local languages in the early years of schooling, their implementation has often been limited and uneven. Integrating local languages into educational technologies offers a practical pathway to advance these policy goals by enabling instruction in learners' home languages at scale.

Whereas prior work using quasi-experimental designs found highest schooling level to be a strong predictor of performance, with larger effects than language of instruction [29], we did not find highest schooling level to be significantly associated with learner performance in our study. While one might expect learners with more formal schooling to outperform those with less schooling, randomization balances educational backgrounds across groups and reduces the selection biases that can affect quasi-experimental estimates. Our findings therefore suggest that instruction in Leb Lango supported learners across all education levels in demonstrating their knowledge. This highlights the potential of multilingual instruction to promote equity in learning. However, additional research is needed to better understand the relative contributions of schooling and language across diverse contexts.

5.2 Practice-Based Learning in Local Languages EdTech Enhances Learning Efficiency and Persistence

Practice activities appear to have played an important role in shaping overall performance: learners in the Leb Lango condition benefited more from practice, which may help explain the cumulative advantage observed in aggregated course outcomes. These patterns suggest differences in engagement and learning processes across the two language conditions.

We found that learners in the two language conditions exhibited different learning behaviors in the digital environment. Firstly, learners in the Leb Lango condition were more likely to answer practice questions correctly on their first attempt and therefore made fewer attempts per practice question, indicating greater efficiency. This suggests that instruction in the local language supported faster comprehension and more accurate initial responses. In addition to the direct benefits of greater comprehension, instruction in the local language also led to indirect benefits. Leb Lango learners progressed through a greater number of practice questions, allowing

them to engage with more varied concepts and persist further in the course. Prior work shows that varied practice is strongly related to the best learning outcomes compared to repeatedly completing the same questions up to correctness [8]. Our findings extend prior evidence to multilingual contexts and suggest that providing practice in learners' local languages can potentially be scaled through EdTech and may enable learners to engage with a wider range of course content and achieve improved learning outcomes.

Secondly, we found that learners' attempts at practice questions reflected a combination of learning strategies. Learners in the English condition engaged more in quick retries, whereas learners in the Leb Lango condition engaged more in repeated lesson attempts. This pattern is consistent with learners receiving instruction in English using repeated practice to get the correct answer, whereas Leb Lango learners repeated practice to review and revise. The distinction between quick retries and repeated lesson attempts suggests that language barriers may lead learners to rely on trial-and-error behaviors rather than review-based engagement. These results highlight important nuances in how language may shape engagement with practice content.

Learners in both the English and Leb Lango groups achieved higher scores when revisiting questions during repeated lesson attempts, and even attempts made after longer intervals during revisions showed strong performance. Notably, learners in the Leb Lango group scored higher than the English group across all attempt types. This suggests another indirect benefit of instruction in learners' home language; by making instruction more accessible and comprehensible, it may provide learners with more time and opportunity to self-regulate their practice. In this case, it allowed students in the Leb Lango group to allocate a greater proportion of their activity to revision attempts and to space their practice more [7, 19], with better learning outcomes.

6 Implications

Overall, integrating instruction and practice opportunities in learners' home languages within EdTech platforms creates strong potential for scaling high-quality learning to broader populations whose linguistic backgrounds align with the language of instruction. Using scalable digital delivery, multilingual EdTech can extend access to STEM content beyond traditional classroom constraints, enabling more learners to explore a wider range of topics and progress further in the curriculum. In turn, this approach may contribute to more equitable participation and learning in STEM at scale.

Our findings provide causal evidence that multilingual environments can promote scalable EdTech even in low-resource settings. Our study provides causal evidence that complements prior research that successfully implemented multilingual EdTech at larger scale [29]. First, language shapes how learners engage with content in digital learning environments, indicating that scalable EdTech should incorporate local languages to promote more efficient comprehension and engagement. Second, it is not sufficient to translate content alone; platforms must support sustained, practice-rich learning experiences in learners' local languages: combining local language instruction with varied practice activities has potential to support better conceptual understanding and enables learners to progress through more topics and learn more. Third, scalable multilingual EdTech must account for differences in how learners interact with content in different languages to ensure effectiveness. For example, the repeated attempts observed among English learners likely reflect a strategy to consolidate understanding in a less familiar language. Therefore, scalable systems should prioritize both the integration of local languages and the provision of adaptive supports–such as contextual translations, scaffolded practice, structured feedback–to accommodate diverse linguistic pathways to learning. Finally, integrating local languages into digital platforms can make language-in-education policies actionable at scale, enabling broader access to quality instruction for multilingual learners.

7 Limitations and Future Work

Our study has several limitations. Firstly, the sample size was small, with only 39 participants, which may limit the precision of our estimates. Secondly, due to platform constraints and the low-tech nature of the system, our analysis was constrained to multiple-choice items, which restricted the range of question types and limited our ability to capture nuanced insights into how learners use language and interact with the technology in free-text. Thirdly, we did not directly measure learners' language proficiency, and therefore cannot determine how proficiency in English or Leb Lango may have influenced the results. Lastly, we did not evaluate whether the English and Leb Lango versions were equivalent in difficulty. Translation choices and potential differences in item difficulty may have influenced how learners navigated the course and engaged with the content.

These limitations point to several directions for future work. Future studies could incorporate baseline measures of language proficiency to better understand how fluency interacts with instructional language in shaping learning outcomes. Furthermore, ensuring equivalence in content difficulty across language versions would strengthen comparability between language conditions. Future studies with larger sample sizes could further validate these findings in broader contexts.

Our results should be interpreted with these limitations in mind. Nevertheless, our work suggests important implications for scaling EdTech to linguistically diverse learners.

8 Conclusion

In this study, we provided causal evidence on the effect of language of instruction on the learning outcomes and learning processes in an EdTech environment through a randomized controlled trial. We compared instruction in Leb Lango with instruction in English and found that learners receiving Leb Lango instruction outperformed English learners on aggregated STEM course performance in non-formal settings. Specifically, learners in the Leb Lango condition were more likely to answer questions correctly on practice activities, required fewer attempts per practice question, and progressed through a greater number of practice questions, indicating more efficient engagement and persistence with the learning material. English and Leb Lango learners performed similarly on weekly tests and the final examination. Our findings reveal different ways in which language shapes learning and engagement. This study provides evidence that the potential of educational technologies to

scale instruction can be achieved by including equivalent and high-quality instruction and practice opportunities in different languages in parallel.

Acknowledgments

We would like to thank the Mastercard Foundation and the Jacobs Foundation (Award No. 2023151900) for supporting this research. We thank our collaborator, Yiya Solutions, Inc., for granting access to the data and offering extensive support throughout this research. We also sincerely thank the students who participated in the course.

References

[1] Olusola O. Adesope, Dominic A. Trevisan, and Narayankripa Sundararajan. 2017. Rethinking the Use of Tests: A Meta-Analysis of Practice Testing. *Review of Educational Research* 87, 3 (June 2017), 659–701. doi:10.3102/0034654316689306

[2] Hassana Alidou, Aliou Boly, Birgit Brock-Utne, Yaya Satina, Kathleen Heugh, and H Ekkehard Wolff. 2006. Optimizing Learning and Education in Africa – the Language Factor. (2006).

[3] Hülya Kosar Altinyelken, Sarah Moorcroft, and Hilde van der Draai. 2014. The dilemmas and complexities of implementing language-in-education policies: Perspectives from urban and rural contexts in Uganda. *International Journal of Educational Development* 36 (May 2014), 90–99. doi:10.1016/j.ijedudev.2013.11.001

[4] Miri Barak, Abeer Watted, and Hossam Haick. 2016. Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education* 94 (March 2016), 49–60. doi:10.1016/j.compedu.2015.11.010

[5] Juliana Bernhofer and Mirco Tonin. 2022. The effect of the language of instruction on academic performance. *Labour Economics* 78 (Oct. 2022), 102218. doi:10.1016/j.labeco.2022.102218

[6] Ellen Bialystok, Fergus I.M. Craik, and Gigi Luk. 2012. Bilingualism: Consequences for Mind and Brain. *Trends in Cognitive Sciences* 16, 4 (April 2012), 240–250. doi:10.1016/j.tics.2012.03.001

[7] Shana K Carpenter, Nicholas J Cepeda, Doug Rohrer, Sean HK Kang, and Harold Pashler. 2012. Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review* 24, 3 (2012), 369–378.

[8] Paulo F. Carvalho, Elizabeth A. McLaughlin, and Kenneth R. Koedinger. 2022. Varied practice testing is associated with better learning outcomes in self-regulated online learning. *Journal of Educational Psychology* 114, 8 (Nov. 2022), 1723–1742. doi:10.1037/edu0000754

[9] Stephen H. Edwards, Krishnan P. Murali, and Ayaan M. Kazerouni. 2019. The Relationship Between Voluntary Practice of Short Programming Exercises and Exam Performance. In *Proceedings of the ACM Conference on Global Computing Education (CompEd '19)*. Association for Computing Machinery, New York, NY, USA, 113–119. doi:10.1145/3300115.3309525

[10] Jenny Eppard, Amir Kaviani, Michael Bowles, and Jason Johnson. 2021. EdTech Culturation: Integrating a Culturally Relevant Pedagogy into Educational Technology. *Electronic Journal of e-Learning* 19, 6 (2021), 516–530. https://eric.ed.gov/?id=EJ1330679 ERIC Number: EJ1330679.

[11] GEM Report UNESCO. 2025. *The right multilingual education policies can unlock learning and inclusion.* GEM Report UNESCO. doi:10.54676/KDJZ7041

[12] C. Goldenberg. 2008. Teaching English language learners: What the research does – and does not – say. *American Educator* 2 (2008), 8–44.

[13] François Grosjean. 2010. *Bilingual: Life and Reality.* Harvard University Press.

[14] M Hossain. 2016. English Language Teaching in Rural Areas: A Scenario and Problems and Prospects in Context of Bangladesh. *Advances in Language and Literary Studies* 7, 3 (April 2016). doi:10.7575/aiac.alls.v.7n.3p.1

[15] Mmaki Jantjies and Mike Joy. 2012. Multilingual mobile learning - A case study of four South African high schools. *CEUR Workshop Proceedings* 955 (Jan. 2012), 208–211.

[16] Mmaki Jantjies and Mike Joy. 2013. Mobile learning through indigenous languages: learning through a constructivist approach, Vol. 2013. Hamad bin Khalifa University Press (HBKU Press), 14. doi:10.5339/qproc.2013.mlearn.14

[17] Mmaki Jantjies and Mike Joy. 2015. Mobile enhanced learning in a South African context. *Educational Technology & Society* 18, 1 (2015). http://hdl.handle.net/10394/25709

[18] Kasch Julia, Kalz Marco, et al. 2021. Educational scalability in MOOCs: Analysing instructional designs to find best practices. *Computers & Education* 161 (2021), 104054.

[19] Sean HK Kang. 2016. Spaced repetition promotes efficient and effective learning: Policy implications for instruction. *Policy insights from the behavioral and brain sciences* 3, 1 (2016), 12–19.

[20] K. Koedinger, Elizabeth Mclaughlin, Julianna Zhuxin Jia, and Norman L. Bier. 2016. Is the doer effect a causal relationship?: how can we tell and why it's important. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (2016). https://api.semanticscholar.org/CorpusID:13031796

[21] Kenneth R. Koedinger, Jihee Kim, Julianna Zhuxin Jia, Elizabeth A. McLaughlin, and Norman L. Bier. 2015. Learning is Not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale* (Vancouver, BC, Canada) *(L@S '15)*. Association for Computing Machinery, New York, NY, USA, 111–120. doi:10.1145/2724660.2724681

[22] Agnes Kukulska-Hulme, Ram Ashish Giri, Saraswati Dawadi, Kamal Raj Devkota, and Mark Gaved. 2023. Languages and technologies in education at school and outside of school: Perspectives from young people in low-resource countries in Africa and Asia. *Frontiers in Communication* 8 (Feb. 2023). doi:10.3389/fcomm.2023.1081155

[23] Christine Kwon, Yuchen Yao, Yuhan Che, John Stamper, and Amy Ogan. 2025. Navigating Local versus Colonial Languages of Instruction in Out-of-School Contexts: Insights from a Randomized Controlled Trial in Uganda. (2025). https://repository.isls.org//handle/1/11333

[24] William R. Leben. 2018. Languages of the World. In *Oxford Research Encyclopedia of Linguistics*. doi:10.1093/acrefore/9780199384655.013.349

[25] Anthony J. Liddicoat and Kathleen Heugh. 2014. Educational Equity for Linguistically Marginalised Students. In *The Routledge Handbook of Educational Linguistics*, Fiona Copland and Angela Creese (Eds.). Routledge, 79–91.

[26] Anthony J. Liddicoat, Kathleen Heugh, Timothy Jowan Curnow, and Angela Scarino. 2014. Educational responses to multilingualism: an introduction. *International Journal of Multilingualism* 11, 3 (July 2014), 269–272. doi:10.1080/14790718.2014.921174 _eprint: https://doi.org/10.1080/14790718.2014.921174.

[27] Phaedra Mohammed and Permanand Mohan. 2013. A Case Study of the Localization of an Intelligent Tutoring System. Memphis, USA.

[28] Phaedra Mohammed and Permanand Mohan. 2015. Dynamic Cultural Contextualisation of Educational Content in Intelligent Learning Environments using ICON. *International Journal of Artificial Intelligence in Education* 25, 2 (June 2015), 249–270. doi:10.1007/s40593-014-0033-9

[29] Phenyo Phemelo Moletsane, Michael W Asher, Christine Kwon, Paulo F. Carvalho, and Amy Ogan. 2026. Inclusive Mobile Learning: How Technology-Enabled Language Choice Supports Multilingual Students. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. Association for Computing Machinery, New York, NY, USA, Article 340, 17 pages. doi:10.1145/3772318.3791139

[30] Audrey Msimanga and Anthony Lelliott. 2014. Talking Science in Multilingual Contexts in South Africa: Possibilities and challenges for engagement in learners home languages in high school classrooms. *International Journal of Science Education* 36, 7 (May 2014), 1159–1183. doi:10.1080/09500693.2013.851427 _eprint: https://doi.org/10.1080/09500693.2013.851427.

[31] Mathias Bwanika Mulumba and Fred Masagazi Masaazi. 2012. Challenges to African development: the medium of instruction in Uganda's education system. *Pedagogy, Culture & Society* 20, 3 (Oct. 2012), 435–450. doi:10.1080/14681366.2012.712056

[32] Sisanda Nkoala. 2024. Educators' Experiences of Using Multilingual Pedagogies During Emergency Remote Teaching: A Case Study of South African Universities. *International Journal of Multilingualism* 21, 1 (2024), 534–547. doi:10.1080/14790718.2022.2063988

[33] Vuyokazi Sylvia Nomlomo. 2007. *Science teaching and learning through the medium of English and isiXhosa: a comparative study in two primary schools in the Western Cape.* Doctor of Philosophy. University of the Western Cape.

[34] Oladokun Omojola. 2011. English-oriented ICTs and ethnic language survival strategies in Africa. *Global Media Journal African Edition* 3, 1 (Aug. 2011). doi:10.5789/3-1-18

[35] Benjamin Piper, Stephanie S. Zuilkowski, and Salome Ong'ele. 2016. Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya. *Comparative Education Review* 60, 4 (Nov. 2016), 776–807. doi:10.1086/688493

[36] Margie Probyn. 2006. Language and Learning Science in South Africa. *Language and Education* 20, 5 (Sept. 2006), 391–414. doi:10.2167/le554.0 _eprint: https://doi.org/10.2167/le554.0.

[37] Marissa Rollnick and Margaret Rutherford. 1996. The use of mother tongue and English in the learning and expression of science concepts: a classroom-based study. *International Journal of Science Education* 18, 1 (1996), 91–103.

[38] Medadi Ssentanda. 2014. The Challenges of Teaching Reading in Uganda: Curriculum guidelines and language policy viewed from the classroom. *Apples - Journal of Applied Language Studies* 8, 2 (May 2014), 1–22. https://apples.journal.fi/article/view/97862

[39] Uganda Bureau of Statistics. 2025. County and Subcounty Profiles – Lango Sub Region – Volume 2. https://www.ubos.org/wp-content/uploads/2025/06/Ankole-Sub-Region-Census-2024-Profile-Report.pdf

[40] UNESCO. 2025. *Languages matter: global guidance on multilingual education.* United Nations Educational, Scientific and Cultural Organization. doi:10.54675/

MLIO7101

[41] Rachel Van Campenhout, Bill Jerome, Jeffrey S. Dittel, and Benny G. Johnson. 2023. The Doer Effect at Scale: Investigating Correlation and Causation Across Seven Courses. In *LAK23: 13th International Learning Analytics and Knowledge Conference* (Arlington, TX, USA) *(LAK2023)*. Association for Computing Machinery, New York, NY, USA, 357–365. doi:10.1145/3576050.3576103

[42] E. Van Laere, K. Aesaert, and J. van Braak. 2014. The Role of Students' Home Language in Science Achievement: A multilevel approach. *International Journal of Science Education* 36, 16 (2014), 2772–2794. doi:10.1080/09500693.2014.936327

[43] Mark Warschauer. 2003. *Technology and Social Inclusion: Rethinking the Digital Divide.* The MIT Press, Cambridge.