The impact of different personalisation algorithms on literacy and numeracy in Kenyan pre-primary education: A comparative study of summative and formative assessments results

The Impact of Different Personalisation Algorithms on Literacy and Numeracy in Kenyan Pre-primary Education: A Comparative Study of Summative and Formative Assessments Results

Chen Sun[1], Louis Major[1], Nariman Moustafa[2], Rebecca Daltry[3], Lazar Obradovic,[4] Aidan Friedberg[4]
[1]University of Manchester, chen.sun@manchester.ac.uk; louis.major@manchester.ac.uk
[2]Open Development & Education, nariman@edtechhub.org
[3]Jigsaw, rebecca@edtechhub.org
[4]EIDU, lazar.obradovic@eidu.com; aidan.friedberg@eidu.com

ABSTRACT: Digital personalisation has demonstrated potential to enhance learning. However, there is limited evidence on the comparative impact of different content personalisation algorithms on early-year numeracy and literacy outcomes, especially in low- and middle-income countries (LMIC). This paper reports an A/B/C test conducted over three weeks via a digital personalised learning tool used by 6479 Kenyan pre-primary learners. Two personalisation algorithms were implemented (maximising learner engagement or score), while expert-curated sequence was used as a control. Learners from 1509 classes were randomly divided across the three partitions. Learning in numeracy and literacy was compared across three metrics: summative assessment, curriculum progress, and formative assessment. Results showed no difference between partitions in the summative assessment. Different partitions tended to progress through the digital curriculum at a different pace. Significant differences between partitions were found in formative assessment scores, with the impact of each algorithmic approach varying according to different learning strands. Findings contribute to a deeper understanding of how different algorithms impact pre-primary education in LMIC contexts, with implications for designing personalised learning approaches tailored to specific learning content and learner profiles.

Keywords: Digital personalised learning, pre-primary, low- and middle-income country, literacy, numeracy

1 INTRODUCTION

Evidence indicates that digital personalised learning (DPL) can have a positive impact on learning outcomes. An important component of personalisation is to sequence learning content to actively engage learners (Diwan et al., 2023) and / or increase knowledge acquisition (Major et al., 2021). Research suggests that content sequencing powered by personalisation algorithms can outperform expert-suggested sequencing (Chau et al., 2018). Among various neural network-based algorithms to sequence content, Long Short Term Memory (LSTM) is commonly used (Huo et al., 2020; Ren & Wu, 2023). Research on LSTM-based algorithms reported two specific purposes: maximising learning outcomes and maximising engagement. This paper contributes to research on digital personalisation in that (1) we implemented and compared two personalisation algorithms (optimising engagement vs. score), evaluated against a default content sequence to assess learning effectiveness, and (2) the study was conducted in an under-researched context, i.e., a low- and middle-income country (LMIC; Kenya). The impact of three different content sequencing methods are investigated, by comparing effects on three learning metrics: summative assessment, curriculum progress, and formative assessment. The main research question is: What is the impact of personalisation (Engagement vs Score vs Expert-curated sequence) on learning for Kenyan pre-primary learners?

2 METHOD

The EIDU DPL platform runs on low-cost Android devices. Learning units align with the Kenyan curriculum in domains (Numeracy and Literacy) and strands (e.g., Classification). An A/B/C test was conducted over three weeks in July 2023 in 1619 low-cost private pre-primary schools in Nairobi, Kenya. Ethical consent for the research was obtained from both the Kenyan government and relevant organisational bodies involved in the study. Parental consent was substituted by institutional and teacher consent. Pre-primary learners (3-6 years old) were randomised and split equally between three experiment partitions (Engagement vs. Score vs. Expert-curated sequence) with anonymised account IDs serving as a hash function seed. The LSTM-based algorithms predicted engagement and scores for each learning unit based on vectors of student performance history. The pre-test data were selected from learners who used EIDU in June 2023, by matching anonymised IDs in the post-test data. The pre-test sample consisted of 5884 learners from 1177 classes. The post-test sample was 6479 learners from 1509 classes. The increase in post-test sample is attributed to inclusion of new schools and classes joining the experiment in July 2023. Learners were assessed using established summative assessments (EGMA, EGRA or MELQO; Friedberg, 2023), curriculum progress (total usage and number of unique units completed), and formative assessment (scores from learning units).

3 RESULTS

Summative Assessment. 1661 learners completed the summative assessment in the Engagement partition, 1702 in the Score partition, and 1640 in the expert-curated sequence group. Possible score range was 0 to 1 for each assessment unit. Scores were averaged across all test units and aggregated to overall scores for literacy and numeracy. ANOVA tests did not reveal significant group differences for pre-test on Literacy learning ($F(2,4106) = 2.96$, $p = .052$) and on Numeracy learning ($F(2,1906) = .11$, $p = .89$). Similarly, post-test analysis did not show any group difference in Literacy ($F(2,4361) = .96$, $p = .38$) and in Numeracy ($F(2,2186) = .58$, $p = .56$).

Curriculum Progress. No differences were found in total usage between partitions ($F(2,6488) = 2.00$, $p = .13$), although there were significant differences in the number of unique learning units completed ($F(2,6488) = 1509.58$, $p < .001$). See Table 1. The Engagement partition progressed through the highest number of unique units. The Engagement partition solved significantly more unique units (13.45) than the Score partition and significantly more units (3.11) than the expert-curated group. The Score partition, however, completed significantly fewer units (10.34) than the expert-curated partition.

Table 1: Curriculum Progress: Usage and Number of unique learning units completed.

	Total duration (in minutes)	Average progress (*Mean*)
Engagement	96.73	21.6
Score	99.73	8.2
Expert-curated sequence	103.74	18.5

Formative Assessment. ANOVA comparing the three groups in the pre-test sample showed no significant differences between the partitions across all eight strands, suggesting the post-test sample is comparable. In total, 6371 learners participated across all three partitions: 2089 Engagement, 2117 Score, and 2165 expert-curated. They collectively played 216 common learning units. ANOVA revealed group differences occurred within all nine learning strands. Analysis using Tukey post hoc tests allowed for pairwise group comparisons (Table 2), suggesting different personalisation strategies may benefit learning in different ways, depending on learning strand or measure in question. The results corroborate the positive effects of content sequencing literature.

Table 2: Post hoc Tukey test for group comparisons

Partition	**Eg vs. Expert Mean (SE)**	**Score vs. Expert Mean (SE)**	**Eg vs. Score Mean (SE)**
Classification	.033 (.004) ***	.044 (.004) ***	-.011 (.003) **
Listening	.028 (.007) ***	-.068 (.007) ***	.096 (.006) ***
Measurement	.056 (.008) ***	.042 (.008) ***	.014 (.009)
Numbers	-.044 (.008) ***	-.162 (.008) ***	.119 (.009) ***
Phonological awareness	-.017 (.007)	-.022 (.007) **	.006 (.007)
Reading	-.013 (.008)	-.117 (.008) ***	-.104 (.007) ***
Speaking	.043 (.007) ***	-.034 (.007) ***	.076 (.006) ***
Writing	-.001 (.001)	-.005 (.001) ***	.006 (.001) ***

Note: ** $p < .01$, *** $p < .001$. Eg = Engagement, Score = Score partition, Expert = expert-curated

4 DISCUSSION AND FUTURE WORK

This work contributes to a deeper understanding of how low-cost DPL benefits literacy and numeracy learning for pre-primary learners in LMICs. The findings demonstrate varied effects of different content sequencing algorithms on specific learning content. Personalisation had no impact on the summative assessment, but may affect learning pathways (e.g., Engagement partition went through learning units faster) and improve certain content learning. Future research should focus on investigating and identifying algorithms that are more beneficial for pre-primary learners in LMICs, taking into account the specific subject matter. Further investigation is needed to pinpoint the exact effects of content sequencing algorithms, by comparing different LSTM-based algorithm designs.

REFERENCES

Diwan, C., Srinivasa, S., Suri, G., Agarwal, S., & Ram, P. (2023). AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Computers and Education: Artificial Intelligence*, *4*, 100110.

Friedberg, A. (2023). Can A/B Testing at Scale Accelerate Learning Outcomes in Low- and Middle-Income Environments?. In (Eds.) *Communications in Computer and Information Science*: *Vol. 1831* (pp. 780-787). Springer.

Huo, Y., Wong, D. F., Ni, L. M., Chao, L. S., & Zhang, J. (2020). Knowledge modeling via contextualized representations for LSTM-based personalized exercise recommendation. *Information Sciences*, *523*, 266-278.

Major, L., Francis, G. A., & Tsapali, M. (2021). The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis. *British Journal of Educational Technology*, 52, 1935–1964.

Ren, J., & Wu, S. (2023). Prediction of user temporal interactions with online course platforms using deep learning algorithms. *Computers and Education: Artificial Intelligence*, *4*, 100133.