Clear evidence, better decisions, more learning.

PREPRINT

Cost-effectiveness and Digital Personalised Learning

Three interventions from Kenya

Date March 2025

Authors Anonymised for Peer Review

DOI 10.53832/edtechhub.1078

Partnership | Progress | Prosperity

@GlobalEdTechHub edtechhub.org

About this document

This is a preprint of an article submitted for journal publication and is the authors' working paper. Any use you make of this preprint is under the conditions of the Creative Commons licence given below, which includes a clause for non-endorsement.

All EdTech Hub research reports are internally and externally peer-reviewed before journal submission. Reviewers of this output are listed below.

Recommended citation	Anonymised for Peer Review. (2025). Cost-effectiveness of digital personalised learning: Three interventions from Kenya [Preprint]. EdTech Hub. https://doi.org/10.53832/edtechhub.1078. Available at https://docs.edtechhub.org/lib/HJ77W2M5.
Preprint DOI	https://doi.org/10.53832/edtechhub.1078
Original preprint publication date	March 2025
Supplemental materials	
Final publication DOI	
Licence	CC-BY Attribution 4.0 International
Corresponding author	Anonymised for the purposes of blind peer review
Reviewers	Daniel Rodriguez-Segura

About EdTech Hub

EdTech Hub is a global research partnership. Our goal is to empower people by giving them the evidence they need to make decisions about technology in education. Our evidence library is a repository of our latest research, findings and wider literature on EdTech. As a global partnership, we seek to make our evidence available and accessible to those who are looking for EdTech solutions worldwide.

EdTech Hub is supported by UKAid, Bill & Melinda Gates Foundation, World Bank, and UNICEF. The views in this document do not necessarily reflect the views of these organisations.

To find out more about us, go to edtechhub.org/. Our evidence library can be found at docs.edtechhub.org/lib/.

Abstract

Recently there has been a significant shift towards better understanding the relative cost-effectiveness of education interventions, particularly those being implemented in LMICs. Work to group similar interventions has added significant value in understanding the types of programmes and forms of technology that are likely to deliver cost-effective impact on learning outcomes. However, this framing also relies on assumptions that programmes grouped within the same category will have similar cost-effectiveness and be broadly similar in the way that they are being implemented to deliver impact on learning. Furthermore, this view misses the key issue that individual features of programmes and how they are contextually implemented are consequential to their cost-effectiveness.

One such category which has gained significant attention in recent years is that of Digital Personalised Learning (DPL). While there are many interventions categorised as offering DPL in some form, it is a difficult category to define owing to the variety of mechanisms through which learning can be personalised. As a result, it can be problematic to assume that the features and cost-effectiveness of DPL programmes can always be compared and grouped, making it difficult to make conclusions around the cost-effectiveness of DPL as a category of interventions.

This lack of consistency is emphasised by the fact there is not a universally agreed definition for what constitutes DPL. Therefore, for the purposes of understanding the cost-effectiveness of DPL programmes, it may be less useful to categorise and assume all interventions with a personalisation element as being comparable. In this paper, we use three examples of DPL interventions being implemented in Kenya to undertake an in-depth analysis of their features and their implications for cost-effectiveness.

Through analysing the key differences between the implementation models, costs, and impact on learning outcomes delivered by these programmes, this paper emphasises significant variance between how DPL programmes are achieving impact. This paper argues that superficial comparisons of cost-effectiveness might miss important information for decision-makers and educators aiming to compare DPL options. As a result of these disparities, DPL interventions should not be conceptualised as its own discrete category when illustrating the cost-effectiveness of groups of interventions.

Introduction

Grouping interventions in cost-effectiveness analysis

In recent years there have been significant developments within the education sector to better understand the relative cost-effectiveness of different groups of interventions, particularly those being implemented in LMICs (e.g. GEEAP, 2023; Angrist et al., 2024). This work has added significant value in helping policymakers identify which interventions may help improve learning outcomes in the most efficient way, and what forms of technology that are likely to deliver cost-effective learning outcomes, particularly in low- and-middle income countries (LMICs).

As a result, many key publications frame discussions around the cost-effectiveness of interventions in relation to their common grouping or categorisation (e.g. Rodriguez-Segura, 2021; GEEAP, 2023; Angrist et al., 2024), with interventions grouped into categories of similar programmes. While this framing is helpful in understanding descriptive patterns in programming or pedagogical approach, by adopting such a wide scope they can sometimes miss some of the nuances within each category that can impact cost-effectiveness on a contextual basis.

Breadth vs. depth of comparison

Taking this approach means that these papers go 'wide' and not 'deep' into what it is that makes each intervention work. This breadth of framing can be useful for multiple reasons, including identifying wider-scale patterns in cost-effectiveness and delivery useful to programme decision-makers. But shortcomings of this approach include missing the contextual nuance of individual interventions, that drive the nature of both cost and effectiveness in each instance.

Furthermore, these generalisations can also lead to false assumptions that there is a similar level of cost-effectiveness in similar types of interventions, or that their unique features are insignificant as long as they belong to that particular category. This is a misrepresentation, as even in the presence of similar cost-effectiveness ratios, interventions may contain specific features impacting cost-effectiveness that are better tailored to one context compared to another.

Comparison of DPL cost-effectiveness

The category of Digital Personalised Learning (DPL) is one of the most commonly discussed categories of EdTech interventions, included in the Smart Buys report as 'Using software that allows personalised learning and adapts to the learning level of the child (where hardware is already in schools)', which is identified as having promising but limited evidence for its cost-effectiveness (GEEAP, 2023).

But DPL has significantly varied conceptualisation and implementation modalities that drive its cost-effectiveness, and there are a lack of examples of work dissecting this category of interventions, and truly understanding their cost-effectiveness from their features. In the context of DPL, it can be problematic to assume that their features and cost-effectiveness can be reliably compared and grouped as a result of these differences. Indeed, any categorisation of interventions that don't follow a prescriptive methodological approach must be explored in much greater depth. Therefore, comparisons of the cost-effectiveness of TaRL (Teaching at the Right Level) interventions which follow the Pratham/JPAL approach to TaRL are very different from comparisons of a diverse category like DPL which requires reflection on the commonalities and differences between a wide range of interventions.

Three case studies in DPL analysis

In light of this gap in comparability of DPL models, this paper uses three examples of DPL interventions in Kenya to provide a detailed examination of their features and costs, and the implications of this for their effectiveness and cost-effectiveness. Through analysing the key differences between the implementation models, their costs, and the nature of learning outcomes they improve, this paper finds that superficial comparisons of cost-effectiveness might miss important information for decision-makers and educators aiming to compare DPL options.

This paper contributes to a more robust and comparable process for understanding cost-effectiveness analysis of EdTech interventions which may be categorised together, but have different features and characteristics. Specifically with regards to DPL, this paper presents comparisons of different interventions and platforms which have significantly different outcomes. This comparison allows for discussion to understand some of the drivers of effectiveness in DPL, and how they relate to costs, both in their initial implementation and potentially at scale.

Understanding digital personalised learning

In recent years, using technology to personalise and adapt learning to the level of learners has emerged as a promising approach for improving learning outcomes (GEEAP, 2023). Personalised or adaptive learning programmes, categorised here as DPL, can differ with respect to the extent of their user-responsiveness, adaptability to teach at the level of the child, the extent of their integration in classroom instruction and the extent to which they facilitate collaborative learning (Major and Francis, 2020; GEEAP, 2023).

Definition of DPL

In their systematic review of DPL, Van Schoors et al. (2021) note the wide variety of conceptualisations and operationalisations of the term used in research. Drawing together descriptions from 53 studies, they offer the following general definition:

Unlike conventional learning, digital personalised learning (DPL) takes place in a digital learning environment that adapts to the individual learner in function of optimising individual and/or collaborative learning processes focussing on cognitive, affective, motivational,metacognitive and/or efficiency outcomes. This adaptation/personalisation: (1) can take into account cognitive, affective, motivational and metacognitive characteristics of the learner; (2) can relate to all aspects of the learning environment, more specifically the (nature, number, and sequence of) learning tasks, the content as well as the instruction and support provided by the learning environment; (3) can be the result of information provided by the teacher or the learner himself/herself, but also information collected by the digital environment; and (4) can be enhanced by the teacher through the effective use of data derived from DPL tools.

Current evidence of DPL effectiveness

Meta-analyses have typically found that DPL can be an effective tool for achieving positive outcomes with respect to both learning achievement and perceptions (Alrawashdeh et al., 2024; Major et al., 2021; Zheng et al., 2022). One of its key features, often cited as crucial to this impact, is adapting teaching so that students can learn at their own pace and proficiency level (Major and Francis, 2020; Plaut, 2024; UNICEF, 2022). For example, Major et al. (2021) detailed that DPL in LMICs had a moderate effect size of 0.18 SD when approaches provided personalised support, feedback, and/or assessment, or allowed learners to tailor the interface or content sequence (Bulger, 2016). However, approaches using adaptive systems that adjust to the level of the learner by actively altering the delivery of learning depending on user behaviour and performance (Bulger, 2016), had an effect size of 0.35 SD, representing a significantly greater impact on learning (Major et al., 2021). While it is recognised that any degree of personalisation can be helpful, evidence suggests that the greater the level of personalisation and adaptability afforded by DPL, the greater the likelihood of significant positive effects on learning (GEEAP, 2023; Plaut, 2024).

While most evidence on the impacts of DPL has emerged from high-income countries (HICs), recent studies have showcased promise that DPL can advance learning outcomes in LMICs (Major and Francis, 2020; Major et al., 2021; Pitchford et al., 2019; Van Schoors et al., 2021). In particular, DPL has been cited as offering significant promise for closing educational gaps for lower attaining students, including those who have experienced a period of absence from schooling (UNICEF, 2022). However, the variability in definitions of DPL complicates the extent to which cross-comparisons of the effectiveness of DPL programmes and tools are reliable and can inform the design of products and programmes (Bernacki et al., 2021; UNICEF, 2022; Van Schoors et al., 2021).

However, there are some key evidence gaps relating to the effectiveness of DPL interventions in LMICs, often relating to factors dependent on the contexts of implementation. While the role of teachers in facilitating DPL is crucial, their specific role in each setting varies depending on how DPL is both designed and integrated into classroom instruction (Major et al., 2021; Plaut, 2024). Teachers often have an inactive role in personalisation, with most DPL tools limiting teachers to monitoring progress (UNICEF, 2022), meaning evidence focusing on the classroom integration of DPL that corresponds to teachers' practice is limited. Furthermore, evaluations on the impact of DPL tools are often not publicly available (UNICEF, 2022) and despite the promise of DPL in improving learning outcomes in LMICs, key questions remain around how they can be implemented at scale and in a way that reaches all learners equally (Plaut, 2024). Hence, this paper focuses on three interventions in the same national context, to allow for greater focus and comparisons between different models of implementation.

Current evidence of DPL cost-effectiveness

Using software that adapts to the learning level of the child in contexts where hardware is already available in schools is identified as showing 'promising but limited evidence' in the recently updated Smart Buys report that categorises the cost-effectiveness of groups of educational interventions implemented in LMICs (GEEAP, 2023). Describing these interventions as cost-effective requires more robust evidence of implementation at scale, and implementation by governments. GEEAP (2023) notes that these approaches can be cost-effective if the hardware to support DPL software is already in place and can be reasonably maintained. It is also important to note that most of the evidence behind this position comes from out-of-school programmes that do not crowd out other learning, and the evidence of DPL during schooling is less robust (GEEAP, 2023). Given the financial challenges in many LMICs, DPL with moderate personalisation affordances can still yield significant learning rewards and may represent a cost-effective entry point in marginalised contexts where higher-tech alternatives are unaffordable (Major et al., 2021).

Other assessments of the cost-effectiveness of DPL have also pointed to the need for greater evidence, but have noted that consistently accounting for both fixed (initial and ongoing software development) and variable (hardware) costs is integral to understanding the costs of any DPL programme (Major et al., 2021). In general, access to sufficient infrastructure and hardware is likely to reduce the volume of these variable costs, as well as improve access to DPL programmes (Major and Francis, 2020; Major et al., 2021). A review of DPL products found that most are reliant on a level of enabling ICT infrastructure (UNICEF, 2022), meaning that it is unclear whether costs may be prohibitively expensive in contexts without this level of baseline infrastructure. Furthermore, the uncertainty around implementing DPL at scale has particular implications for the cost-effectiveness and overall affordability of interventions in LMICs, where effectiveness at scale is contingent on having adequate supporting infrastructure and hardware availability, in addition to teacher preparedness to support DPL (Plaut, 2024).

Therefore, there is still a significant degree of uncertainty around the cost-effectiveness of DPL programmes in LMICs owing to a lack of conclusive evidence around both the cost and impact of DPL, and as such the issue merits further interrogation. This paper provides a deeper dive into a regional portfolio of DPL interventions in order to examine the relative cost-effectiveness of different interventions and implementation models. This will provide a thorough examination of whether there are particular aspects of DPL that drive cost-effectiveness, the affordances of calculating LAYS of DPL interventions, as well as caution against drawing cross-context conclusions around the cost-effectiveness of a broad range of loosely linked interventions.

Methodology

Studied interventions

This article draws primarily on three DPL programmes commissioned by EdTech Hub. While there is not a consensus definition of DPL across the sector, we follow Van Schoors et al. (2021) in defining DPL as the use of a digital learning environment that adapts to the individual learner, with the goal of optimising individual and/or collaborative learning processes to enhance cognitive, affective, motivational, metacognitive or efficiency outcomes. The studies that are introduced in this section therefore fall within this conceptualisation of DPL.

Studies in this portfolio also serve as useful case studies to illustrate the significant differences between DPL programmes. The following section uses three interventions from the portfolio of studies in Kenya to emphasise the key differences between DPL interventions in terms of implementation, costs and integration. This section does not attempt to define the detail of how each intervention defines and achieves personalisation, but provides an overview of a range of studies that could be grouped in the same category, and highlights their differences.

Listed below is a brief summary of the three portfolio studies that will form the basis of the discussion around the comparability of cost-effectiveness data for DPL programmes being implemented in similar contexts. For each study the following is provided:

- An overview of the study and its core objectives
- The methodology of the study
- How each study uses personalisation
- How each study's platform is integrated into instruction
- A key publication for further reference

Oppia

Study overview: This study titled "The centrality of peer interaction in technology-supported personalised learning" examined the *Oppia* application, a DPL tool aiming to affect early-grade numeracy outcomes in

Kenya: the teacher-student ratio, the number of devices per child, promotion of peer learning.

Methodology: Multi-factorial RCT design testing 364 students in grade 5 on numeracy with pre/post maths assessments.

Personalisation: The open-source *Oppia* platform supported the creation of personalised numeracy lessons, which offer automatic tutoring with personalised feedback based on responses, with the application being available offline and hosting curriculum-aligned lessons.

Integration: Content delivered through the *Oppia* platform is standalone to classroom teaching.

Key publication: Forthcoming.

EIDU

Study overview: This study titled "Digital personalised learning to improve numeracy outcomes in Kenyan primary school classrooms" in partnership with EIDU and Women Educational Researchers of Kenya (WERK) investigates how a classroom-integrated, DPL tool can most effectively support early-grade numeracy and literacy outcomes in Kenya.

Methodology: Randomised controlled trial, with design-based implementation research. 1995 pre-primary learners sampled from 291 schools, tested with Save the Children's International Development and Early Learning Assessment (IDELA) tool at baseline, midline and endline.

Personalisation: The EIDU tool is aligned with curriculum and teaching practice is central to its effectiveness as a classroom-integrated DPL tool. The sequence of digital learning units is personalised for each learner based on their learning history within the app.

Integration: The tool is integrated into classroom instruction at the directive of the teacher.

Key publication: Major, L., Daltry, R., Otieno, M., Otieno, K., Zhao, A., Sun, C., Hinks, J., & Friedberg, A. (forthcoming). Digital Personalised Learning to improve literacy and numeracy outcomes: A randomised controlled trial in Kenyan pre-primary classrooms.

M-Shule

Study overview: The study titled "Low-tech personalised learning to improve girls' education in Kenya" investigates the elements of personalised learning delivered through low-cost modalities, namely the M-Shule SMS platform.

Methodology: The project used a mixed methods research design, drawing upon primary and secondary data to understand the research context in different ways.

1. Secondary analysis of existing data (effect sizes) about learning outcomes for girls - focusing on different types of EdTech interventions. This showed that personalised learning is effective relative to other types, and a good candidate for research.
2. Analysis of M-Shule data collected during recent initiatives: the Keep Kenya Learning programme and adapting and delivering Tusome content via SMS. We looked at data about how learners interacted with the content, and learning outcomes (ASER scale).
3. A telephone survey about learners' and caregivers' perspectives on using M-Shule throughout. We also worked to design and implement new content with M-Shule.
4. Endline ASER tests were then compared to baseline results.

Personalisation: The M-Shule platform combines SMS with personalised learning, in order to deliver educational content that is adapted to the students' level, without the need for online connectivity.

Integration: Content delivered through the M-Shule platform is standalone to classroom teaching.

Key publication:

Jordan, K., Myers, C., Damani, K., Khagame, P., Mumbi, A., & Njuguna, L. (2024). Supporting equitable access to learning via SMS in Kenya: Impact on engagement and learning outcomes. British Journal of Educational Technology, 00, 1–23. https://doi.org/10.1111/bjet.13533

Areas of study

Outlined below are key areas related to both the effectiveness and cost of these studied interventions, which will form the basis of the structure for the discussion. In this section, the terms are presented and defined to provide clarity. In doing so, this section serves to outline some of the key differences between the included DPL interventions. It does not attempt to detail or define all aspects of personalisation, but rather illustrate key differences between how DPL programmes are implemented to emphasise their variance.

Describing the effectiveness of DPL interventions

'Effectiveness' refers to the level to which studies are able to impact learning outcomes. In all three studies within this paper, this refers to their impact on foundational numeracy and literacy scores. Listed below are the definitions of key terms related to the effectiveness of these interventions, which will form the thematic basis of the discussion in the following section.

Curriculum alignment - This category outlines how well the platform aligns with existing pedagogical approaches and national curricula. Understanding this is crucial to establish the extent to which learning outcomes delivered by each platform represent meaningful progression through national education systems.

Comparability and external validity of effectiveness - The results of any cost-effectiveness analysis are necessarily relative: "cost-effective" compared to what, and for whom? This category for analysis relates to the nature of the data collected, and whether it is comparable to other data, and indeed whether it has some external validity, allowing it to be compared to other data.

Cumulative Impact - The impacts of an intervention are understood through research which may test different hypotheses about which elements or variations of implementation are most effective, and taken together these can be described as the cumulative impact. While longitudinal research may be necessary to understand the impacts of an intervention more broadly, the impacts on learning outcomes are measured in each of the studies.

External drivers - Another crucial factor with DPL is that the impact of each study will inevitably include other factors (outside of the personalisation aspect) that impact upon its effectiveness, the range and significance of which often varies which makes comparisons less accurate.

Describing the costs of DPL interventions

Costs in this paper refer to the total expenditure required to deliver and implement the programme being researched. This does not include research costs, but encompasses the 'total cost of implementation' which are the costs needed to fully implement all programme activities from scratch. Listed below are several key issues related to categorising the total cost of implementation, which are introduced here but form the basis of the discussion of cost results in the following section.

Development costs - One way in which the studies are not directly comparable is through how they incorporate the development costs for the DPL platforms that are central to their interventions.

Challenges with scaling - A key consideration with DPL platforms is how their cost is expected to change with scale beyond the intervention period being considered by any cost-effectiveness analysis.

Long-term maintenance and support costs - Another challenge with assessing the costs of DPL interventions is that the level of maintenance and support of hardware and software are often poorly understood, and broadly underestimated. Few DPL interventions have been implemented for long enough to accurately calculate amortisation of capital expenditure, degradation of devices and ongoing support needs beyond initial implementation cycles.

Cost categorisations - This refers to the way in which studies group and report their costs. Due to significant differences in what studies reported as relevant categories and how they defined key cost components of each intervention, they are incorporated differently in final cost-effectiveness analyses resulting in an inconsistent categorisation of the total 'cost of implementation.'

Comparing findings on cost-effectiveness data

The variation between DPL programmes is further highlighted when looking at the cost and impact data of the three portfolio studies. Emphasising key aspects of their cost and impact demonstrates why trying to conceptualise the cost-effectiveness of DPL as a single category is flawed, as the data demonstrates differentiation between different DPL interventions.

The three portfolio studies included in this analysis are programmes that are categorised by one organisation (EdTech Hub) as being 'Digital Personalised Learning', and are being implemented in the same country context (Kenya). Even with these similarities, the cost data demonstrates clear differences. The remainder of this section highlights several factors that emphasise the difference between how cost and effectiveness data is calculated, and what it represents, for different DPL programmes.

Summary feature chart

The below feature chart (Table 1) categorises the core aspects of the three DPL platforms (Oppia, EIDU and M-Shule) belonging to EdTech Hub's portfolio of DPL programmes in Kenya.

This section does not discuss the differences between the implementation models of the DPL interventions, or define the mechanisms through which they achieve personalisation. Instead, the feature chart and narrative summary provide a descriptive overview to clearly illustrate some of the significant differences between many core aspects of DPL programmes that means they are not reliably comparable as similar interventions. The chart uses the following broad categories to describe the way each platform works:

- **Tool alignment -** this category refers to the alignment of the tool with national curricula, teacher lesson plans and/or other pedagogical frameworks which justify the sequencing and selection of learning activities.
- **User experience** - this category outlines how each user interacts with the DPL platform, the nature of the interaction afforded by the hardware and software of the platform, and the extent of interactivity within the platform. This helps to determine the parameters within which the platform provides functionality and interactivity to the user.
- **Data** - this category outlines the information that is collected and stored by each DPL platform. This is important to help understand how platforms use data to inform their personalisation and experiences for users.
- **Personalisation** - this category emphasises the detail of what aspect within each platform is personalised, and the way in which the platform provides this personalisation. This is critical to understand where within the broad spectrum of personalised learning each platform fits, in order to provide a basis for more reliable separation of different DPL platforms.

Table 1. Feature chart detailing the key differences between how three digital personalised learning programmes are being implemented in Kenya.

	DPL interventions		
	Oppia	**EIDU**	**M-Shule**
Tool alignment			
National curriculum	yes	Yes (KICD approved)	no
Lesson plans	no	yes	yes
Teacher-led	yes	yes	yes
User Experience			
Hardware modality	Smartphone	Android smartphone	SMS
Software	Simple application	Rich media application	text
Interactivity	medium	high	low
Data			
Where is it stored?	cloud	Local & cloud	NA
What is collected?	User learning data	User learning and usage data	none
Personalisation			
What is personalised?	Sequence of learning activities	Sequence of learning activities	feedback
How is it personalised?	Algorithm	Algorithm + LSTM + teacher choice	Teacher choice

Table 1 demonstrates that the nature of programmes categorised as DPL are often significantly varied and represent a range of divergent designs. While there are often commonalities between studies, such as most products being largely consistent with respect to personalising learning activities and pathways, there remain significant differences in relation to integration with curriculum, teachers and pedagogical approaches.

Table 1 provides an overview that the nature of DPL platforms are often significantly varied and represent a range of divergent designs. While most products are largely consistent with respect to personalising learning activities and pathways, and storing user data to inform personalisation, there are particular differences with respect to how the user engages with the product, the hardware and software that is used, as well as the extent to which tools are aligned to existing curriculum and pedagogical practice. Broadly speaking this has significant implications for cost-effectiveness in terms of comparability between these different platforms, and it is important to reflect that any difference between platforms will undermine this. The relationship between upfront costs in developing additional DPL features and functionality remains poorly understood, however, and improvements in learning outcomes are not necessarily correlated.

Discussion of effectiveness

The following section discusses key findings relating to the effectiveness of the three studies in Kenya. The discussion is structured around the same key effectiveness concepts introduced in the previous section.

Curriculum alignment

A key aspect of DPL that is important to understand is what their quantified impact on learning represents, in terms of progression through curriculum. In the context of EIDU, the classroom-integrated DPL tool has had a significant impact on pre-primary learning outcomes (Major et al., forthcoming). A micro-LAYS approach was used, comparing SD gains by treatment groups (over the control) multiplied by the duration and national learning coefficient. The total additional LAYS gained from the intervention is equal to .422 more LAYS per student than the control group over the duration of intervention. This result, derived from the findings of Major et al. (forthcoming), is published in terms of LAYS for the first time here, in collaboration with the authors.

These results indicate the effectiveness of classroom-integrated DPL. This is important because scaling the learning outcomes of DPL requires that they are integrated into national curriculum-focused interventions. While many other DPL applications are supplementary to national curricula and formal education (Major et al., 2021; Van Schoors et al., 2025), these findings show that classroom-integrated DPL can have a greater impact on learning outcomes, while also demonstrating greater potential for scalability. In addition to this teachers were highly positive about the tool, especially its alignment with classroom activities, indicating the effectiveness of classroom-integrated and curriculum-aligned DPL (Daltry et al., forthcoming).

This is important for understanding DPL because even largely similar interventions can have a significant difference in impact depending on whether they are curriculum-aligned. Whether the platform and its content is curriculum aligned is key to understanding what any impact on learning represents. Even if an intervention is highly impactful on learning and cost-effective, if it is not curriculum aligned then this learning gain may be supplementary and not representative of meaningful progression through an educational system. As a result, effect sizes (and the associated cost-effectiveness) may not encapsulate the true contextual value of learning that each programme delivers.

Comparability and external validity of effectiveness

On the research side, the way in which studies are making claims about their cost-effectiveness also varies and needs to be considered. Oppia delivered a cost-effectiveness of one implementation model from the multi-tailed experiment (chosen because of comparability) was 1.4 LAYS per US$100, at a cost per child of ~US$35. While the study used a randomised control trial (RCT), the number of students (n=364) may be too low for the results to be generalised to the broader population.

Conversely, EIDU involved an RCT conducted from October 2022 to October 2023 in Murang'a county, Kenya. This involved 291 schools and a final sample of 1995 pre-primary learners, assessing numeracy and literacy outcomes in the control and treatment groups across three timepoints. The assessment battery used for the RCT was Save the Children's International Development and Early Learning Assessment (IDELA) tool. The data to emerge from this study which has a much larger sample size is therefore more robust, and arguably should be given more weighting when making comparisons or assessments of the cost-effectiveness of DPL.

Cumulative impact

Another notable factor to emerge from EIDU is the age group — low digital literacy in early childhood may present challenges for DPL, because children of this age have less prior exposure to digital technology and devices. While this forthcoming study does not make explicit claims about the cumulative effects of early DPL interventions on digital literacy and learning outcomes more broadly, it should be noted that DPL for foundational literacy and numeracy has compound effects in other contexts, and this study is foundational for demonstrating its impact in Kenya (and possibly other LMICs). Ensuring equitable access to digital technologies in early years can mean that more equitable outcomes are attainable throughout the educational system, as digital technologies are integrated throughout the education system.

External drivers

It is also important to be aware of other factors beyond personalisation that can drive impact within DPL programmes. The findings of Oppia for example demonstrated that peer learning and device sharing can be among the most effective components of DPL, and can lead to not only better learning outcomes, but also reduced costs, suggesting it may be a cost-effective and scalable model for DPL. The alignment of the tool with teaching support and practice is central to engagement, while peer learning and device sharing were significant contributors to improved outcomes.

This study demonstrates that peer learning and device sharing can be one of the most effective aspects of DPL, and can lead to not only better learning outcomes, but also reduced costs, suggesting it may be a cost-effective and scalable model for DPL. However, further research is needed to verify the findings, as the authors note the "sample size is small and students are limited to one school… [and] the intervention is short." Therefore, considering the range of factors that impact upon DPL, such as the extent of peer learning, is essential to make reliable comparisons. Existing groupings of DPL interventions where this aspect varies significantly may not capture just the cost-effectiveness of personalisation, but also the cost-effectiveness of altering engagement with peer learning or other impactful variables.

Discussion of costs

The following section discusses key findings relating to the cost of the three studies in Kenya. The discussion is structured around the same key cost concepts introduced earlier.

Development costs

One way in which the studies are not directly comparable is through how they incorporate the development costs for the DPL platforms that are central to their interventions. The Oppia study was only able to provide limited cost data which does not include some of the key development costs of the programme. As the platform has been operating for a number of years, these development costs were not provided as part of this intervention. However, not factoring in some development costs represents an underestimation of the actual costs (and therefore an overestimation of cost-effectiveness). This undermines the extent to which cost-effectiveness outcomes are comparable to the other interventions in the portfolio which are being implemented at an earlier stage of product development, where initial set-up costs will inevitably be higher and increase costs.

Comparatively, cost data for the EIDU RCT period accounts for significant development costs due to it being the first year of product implementation. It is important to highlight at the outset that this means that the costs reported are not directly comparable to Oppia which did not report these development costs, and that in the context of EIDU these costs are expected to reduce significantly in subsequent iterations and scaling of the product. For example, 36.1% of the costs accrued across the RCT period are related to complementary measures, namely the provision of books and structured pedagogy training, which were part of the RCT delivery but are not directly relevant to the DPL aspect of the programme and were implemented as one-time costs.

Similarly to EIDU, some of the costs for M-Shule relate to initial one-time costs associated with the intervention, in this case adapting learning content to SMS format. These examples demonstrate that the variation in accounting for the development costs of DPL platforms can significantly alter the cost-effectiveness data, and so comparing platforms of similar maturity may be most reliable.

Challenges with scaling

Moreover, a key consideration with DPL platforms is how their cost is expected to change with scale beyond the intervention period being considered by any cost-effectiveness analysis. For EIDU, a number of different factors are relevant to understanding the cost of the study, based on up-front costs as well as ongoing implementation costs. The costs related to training, devices, and set-up lead to a cost of $9.04/learner in year one, which reduces to $5.68 and $5.91 in years two and three due to the schedule for maintenance and support, and associated costs. The average cost per year is therefore $6.88 when considering the specific implementation scenario, however it may also be relevant to compare the ongoing cost of maintaining this rollout, i.e. $5.80. Note that this does not account for further cost savings that would be made from expanding the scale. Specifically in this implementation scenario, and its likely path to scale, where engagement at the school level is only in two grades, but would expand within the same locations to more grade levels, the cost savings are very notable.

Basing the current calculations on actual expenditure of the current EIDU implementation model, we should consider the two figures for cost-effectiveness:

- Total cost of implementation, averaged over three years = $6.88 per learner per year.
- Average cost of continued implementation after initial startup costs (but including ongoing costs of software maintenance and support) = $5.80

These figures, as with any implementation, still include some assumptions but they do represent an accurate estimate of costs of the implementation whose learning outcomes have been measured, while also providing some indication of the potential replicability of this implementation scenario for anyone else designing a similar project.

Other studies also demonstrated the importance of considering scale when accounting for cost and impact. The findings from Oppia demonstrated the greater impact and potential to scale for DPL integrated within formal schooling (in this case teacher support). M-Shule found that learners had significantly better learning outcomes as a result of the intervention with a total effect of 0.181 Learning-Adjusted Years of Schooling (LAYS) per learner. This study also shows the potential for scaling DPL through low-cost messaging technologies. This has equity implications for girls and marginalised learners in particular. If the scale of the programme were to be increased, the ongoing implementation costs would be reduced and the cost-effectiveness of the intervention enhanced.

Therefore with respect to DPL, as with many other interventions, the cost-effectiveness of the programmes is often highly dependent on their scale. When making comparisons between different programmes, and assumptions around the cost-effectiveness of their category, it is therefore also important to consider how cost-effectiveness changes with scaling.

Key cost categorisations

There were also significant differences in what studies reported as the 'cost of implementation' and therefore how many key cost components of each intervention are incorporated in the final cost-effectiveness analysis.

Oppia, for example, was unable to provide costed equivalents for volunteer software development time spent on developing the platform, and although these categories were estimated from time allocation and average salaries of such roles, the lack of direct cost reports on these categories is likely to undermine the accuracy of the cost data.

Furthermore, whether DPL interventions are relatively 'high' or 'low' cost can mean that their impact and purpose is significantly different. For example, M-Shule examined elements of personalised learning that can be implemented in a more sustainable and cost-effective manner through lower-cost modalities, in order to understand the scalability of DPL in low-income contexts. Low-cost modalities for delivering DPL (such as mobile phones in this case) are important for expanding the benefits to learning outcomes equitably. This study addresses a gap in the literature around the combination of SMS and personalised learning, specifically with regard to gender equity and reaching marginalised groups. This emphasises that DPL interventions can have a range of costs of implementation for a range of reasons, such as their modality, purpose, or way costs are categories.

Calculating LAYS within DPL

The exploration of cost and impact data from the portfolio of DPL programmes indicates that, more generally, calculating the cost-effectiveness of DPL interventions needs to account for additional variation and nuances.

Learning-Adjusted Years of Schooling are used as a metric in different ways - when specific interventions are analysed with learning outcomes data, "micro-LAYS" are calculated[1]. Within the context of the DPL programmes investigated here, micro-LAYS are based on the change in learning outcomes defined in standard deviations, by the learning coefficient relevant to the country context, which is 0.79 for Kenya. Yet this learning adjustment, which determines an additional year of schooling in Kenya to be worth 0.79 years of quality schooling is based on relatively few data points from standardised tests which are not consistently implemented at a broad enough scale in Kenya to be accurately representative of learning outcomes nationally (Patrinos and Angrist., 2018). When comparisons to other contexts are introduced, the validity of this co-efficient adjustment must be considered with some caution. Additionally, the specific learning trajectories and expected gains at different grade levels are not accounted

[1] micro-LAYS calculations are distinct from, but comparable to macro-economic estimates of the returns on education used in determining LAYS at a national level.

for within this data. Given these uncertainties, the comparison of cost-effectiveness of these three interventions should be considered as illustrative estimates.

Even with these caveats, it is clear that the cost-effectiveness of the EIDU implementation is a positive outlier. Regardless of how costs are accounted for, it is at least three times more cost effective. With the most rigorous total-cost accounting, EIDU's DPL implementation comes to a figure of 4.61 LAYS/$100 for this implementation, or 5.47 LAYS/$100 when accounting for only the ongoing cost after setup. By contrast, M-Shule produces only 1.679 LAYS/$100 and the estimate around Oppia's cost-effectiveness, albeit less precise, comes to only 1.4 LAYS/$100. When compared to other DPL interventions, which average 2.8 LAYS/$100, the EIDU intervention stands out among this class of interventions - however wide variation from 0 to 10 LAYS/$100 has been reported (GEEAP, 2023).

These calculations of LAYS are not yet able to take into account some of the additional nuance around learning progression and learning trajectories which data from DPL interventions could provide. However, they are able to provide a baseline for comparison with other studies, and to demonstrate illustrative comparative values of different implementation models.

Conclusion

These comparative findings demonstrate some of the difficulties of comparing different types of DPL interventions, as well as the need for more data that can support the Harmonised Learning Outcomes database. This suggests a key role for DPL in generating improved data around learning progression, learning trajectories and how these compare in different contexts, so that more granular and accurate comparisons can be made. For example, these interventions have the potential to unlock answers to questions such as "How do Kenyan primary students progress from simple arithmetic to long division? How is this affected by their teacher and curriculum and how does it compare to progression in other countries?" Answers to questions like these will be critical to enhancing the quality of education, and adapting teaching and learning to different contexts.

However, the comparison between the studies also underscores the importance of improved clarity on different platforms and implementation models of DPL. These studies demonstrate the complexities of comparative cost-effectiveness analysis, even within one intervention class, such as DPL, where cost categories and effectiveness measures are theoretically similar. Although general comparison based on such estimates can be made, and may be considered informative for decision-making within the contextual parameters discussed, direct quantitative comparisons are more limited and require further data. Further exploration of these gaps is necessary to facilitate more rigorously costed decision-making of EdTech design and implementation.

The current research aims to shed light on the gaps in data, as well as the imperatives for greater consistency in cost-capture, and clearer categories for DPL, and systematic characterisation of core components to what has become an increasingly fragmented category. While important work on this has been done from a theoretical and conceptual level, data from implementation should also be used to support these categories. The need for this expanded evidence base can be addressed in future phases of this same research, including the second phase of EIDU's implementation in Kenya, and comparative examples from other countries which share greater similarities to the DPL platform employed. This research agenda also points to the need for longitudinal data on learning outcomes, and research on the compounding effects of DPL at foundational levels.

Bibliography

Alrawashdeh, G. S., Fyffe, S., Azevedo, R. F. L., & Castillo, N. M. (2024). Exploring the impact of personalized and adaptive learning technologies on reading literacy: A global meta-analysis. *Educational Research Review*, *42*, 100587. https://doi.org/10.1016/j.edurev.2023.100587

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, H., & Sabarwal, S. (2024). *How to Improve Education Outcomes Most Efficiently? A Review of the Evidence Using a Unified Metric.* https://www.cgdev.org/sites/default/files/how-improve-education-outcomes-most-efficiently-evidence-review.pdf

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A Systematic Review of Research on Personalized Learning: Personalized by Whom, to What, How, and for What Purpose(s)? *Educational Psychology Review*, *33*(4), 1675–1715. https://doi.org/10.1007/s10648-021-09615-8

Bulger, M. (2016). *Personalized learning: The conversations we're not having* (p. 29) [Working Paper]. Data & Society Research Institute. https://www.datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf

Daltry, R., Hinks, J., Sun, C., Major, L., Otieno, M., & Otieno, K. (forthcoming). Integrating digital personalised learning into early-grade classroom practice: a teacher-researcher design-based research partnership in Kenya.

Global Education Evidence Advisory Panel. (2023). *2023 Cost-effective approaches to improve global learning: What does recent evidence tell us are 'Smart Buys' for improving learning in low- and middle-income countries?* World Bank. https://thedocs.worldbank.org/en/doc/231d98251cf326922518be0cbe306fdc-0200022023/related/GEEAP-Report-Smart-Buys-2023-final.pdf

Jordan, K., Myers, C., Damani, K., Khagame, P., Mumbi, A., & Njuguna, L. (2024). Supporting equitable access to learning via SMS in Kenya: Impact on engagement and learning outcomes. *British Journal of Educational Technology*, *n/a*(n/a). https://doi.org/10.1111/bjet.13533

Major, L., Daltry, R., Otieno, M., Otieno, K., Zhao, A., Hinks, J., Sun, C., & Friedburg, A. (2023). *Randomised Controlled Trial Protocol: Digital* *personalised learning to improve literacy and numeracy outcomes in Kenyan classrooms*. EdTech Hub. https://doi.org/10.53832/edtechhub.0176

Major, L., Daltry, R., Otieno, M., Otieno, K., Zhao, A., Sun, C., Hinks, J., & Friedberg, A. (forthcoming). Digital Personalised Learning to improve literacy and numeracy outcomes: A randomised controlled trial in Kenyan pre-primary classrooms.

Major, L., & Francis, G. (2020). *Technology-supported personalised learning: Rapid Evidence Review* [Report]. EdTech Hub. https://doi.org/10.17863/CAM.59538

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*(5), 1935–1964. https://doi.org/10.1111/bjet.13116

Patrinos, H. A., & Angrist, N. (2018). *Global Dataset on Education Quality: A Review and Update (2000-2017)* (Policy Research Working Paper No. WPS8592). World Bank Group. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/390321538076747773/Global-Dataset-on-Education-Quality-A-Review-and-Update-2000-2017

Pitchford, N. J., Chigeda, A., & Hubber, P. J. (2019). Interactive apps prevent gender discrepancies in early-grade mathematics in a low-income country in sub-Sahara Africa. *Developmental Science*, *22*(5), e12864. https://doi.org/10.1111/desc.12864

Plaut, D. (2024). *How Can Implementers Apply Digital Personalised Learning in Schools?* EdTech Hub. https://doi.org/10.53832/edtechhub.1008

Rodriguez-Segura, D. (2021). EdTech in developing countries: A review of the evidence. *The World Bank Research Observer*, *lkab011*. https://doi.org/10.1093/wbro/lkab011

UNICEF. (2022, May 26). *Trends in Digital Personalized Learning | UNICEF Office of Global Insight & Policy*. https://www.unicef.org/globalinsight/reports/trends-digital-personalized-learning

Van Schoors, R., Elen, J., Raes, A., & Depaepe, F. (2021). An overview of 25 years of research on digital personalised learning in primary and secondary education: A systematic review of conceptual and methodological trends. *British Journal of Educational Technology*, *52*(5), 1798–1822. https://doi.org/10.1111/bjet.13148

Zheng, L., Long, M., Zhong, L., & Gyasi, J. F. (2022). The effectiveness of technology-facilitated personalized learning on learning achievements and learning perceptions: A meta-analysis. *Education and Information Technologies*, *27*(8), 11807–11830. https://doi.org/10.1007/s10639-022-11092-7