EEDI 2024

Impact Report

A study to evaluate the effectiveness of Eedi on raising attainment in mathematics at KS3 (Year 7)

DR WAYNE HARRISON
DR EMMA DOBSON
PROFESSOR STEVE HIGGINS
GERMAINE UWIMPUHWE
RAHIL KHOWAJA

www.interventions.whatworked.education

Table of Contents

Executive Summary

This report evaluates the impact of Eedi, a digital mathematics platform, on raising maths attainment amongst Key Stage 3 students (Year 7).

Eedi is an innovative EdTech platform offering over 60,000 diagnostic maths questions, engaging 15,500 monthly users through its school and tutoring services. Enhanced by AI and video explanations, Eedi helps learners tackle misconceptions and provides additional support. Premium users benefit from live, one-on-one tutoring, allowing students to request and receive assistance through the platform's chat function.

This impact evaluation included 20 schools, with randomisation at school level. The Eedi programme was used in intervention schools from Autumn 2023 to June 2024 to supplement classroom teaching, with the control schools not using the programme. Eedi is set by teachers for students to complete in lesson time or as an additional homework task. The evaluation also included a light touch implementation and process evaluation, using online teacher surveys and Eedi usage data.

Key Findings

Learning Outcomes: Children using Eedi achieved higher average scores on the NWEA assessment than those in control schools using other or no math platforms. The difference was statistically significant, and the consistent positive trend, after accounting for prior attainment, indicates encouraging early evidence of Eedi's potential to support math learning.

Efficacy: Eedi had an effect size of 0.17, meaning students in Eedi schools scored, on average, 0.17 standard deviations higher than those in control schools using other or no maths platforms. This suggests that Eedi was more effective than other approaches.

Pupil Premium Students: While the data indicated a positive trend in NWEA scores among Pupil Premium (PP) students using Eedi, this difference was not statistically significant. This may be attributed to differing levels of engagement across students.

Engagement: CACE analysis suggests that the intervention effect increases substantially with higher levels of engagement with Eedi. Among students who complied with the intervention at different engagement thresholds, there was a clear dose-response relationship, with higher levels of engagement associated with progressively larger effect sizes. These CACE estimates, which account for non-compliance and estimate the effect among those who would engage with the intervention, demonstrate that adequate engagement with Eedi leads to meaningful improvements in NWEA assessment outcomes in both PP and non-PP students. Analysis of usage patterns demonstrates a notable effect of Eedi on PP students, particularly when usage exceeds 120 check-in questions per year. PP students who engaged with the platform at this level showed significant improvements in their NWEA scores.

Efficacy in Low Pre-test Scorers: In the subgroup of students whose pre-test scores fell below the median, Eedi demonstrated a positive and significant impact on performance outcomes. This suggests that even among lower-performing students, engagement with Eedi can contribute to improved academic performance.

Implementation: The process evaluation demonstrated that a high proportion of teachers used Eedi assignment data to inform their teaching, allowing them to address specific areas where students exhibit misconceptions or difficulties, thereby tailoring their instruction to address these gaps in understanding. By proactively incorporating insights from Eedi, these teachers were able to make their lessons more targeted and responsive to student needs. The process evaluation also identified that control schools were using similar interventions to Eedi, which could be considered close substitutes. The data shows that 89% of control group teachers used an online system for providing homework, and these platforms often included features that were comparable to those offered by Eedi.

Implications and Future Directions

Check-In Questions: The study suggests that Eedi demonstrates a unique and especially large effect. Students answering increased number of questions show a marked increase in attainment on NWEA standardised assessment; suggesting that higher engagement, measured by the number of questions attempted, is associated with greater attainment scores in both PP and non-PP students.

These findings highlight that Eedi can be an effective tool to improve maths attainment, particularly when students are adequately engaged.

Findings demonstrate that pupils with higher engagement levels show significantly greater NWEA improvements compared to control schools.

The intervention also showed a positive effect among students with below-median pre-test scores, reinforcing Eedi's potential to support a broad range of learners.

What Worked Education

Eedi

Introduction

In recent years, there has been a significant increase in the quantity of educational technology marketed to schools worldwide. However, few of these technologies are methodically evaluated to determine if they improve outcomes. As a result, many may not be beneficial, potentially wasting resources and overshadowing effective solutions with unsubstantiated claims (UNESCO, 2023). This study aims to contribute to the literature on evaluating educational technology using appropriately robust methods.

Background

Computer-Aided Instruction in High School Mathematics

It has long been anticipated that computer programs would play a significant role in teaching mathematics to middle and high school students. Early pioneers in educational technology, such as Suppes (1969), predicted that Computer-Aided Instruction (CAI) would play a critical role in teaching mathematics, especially in secondary education. Suppes believed that CAI would revolutionise mathematics education by using individualised instruction to provide a tailored learning experience, thereby improving student outcomes.

Much enthusiasm for CAI in mathematics may stem from its lower cost compared to human tuition, offering the potential to teach more for less and expand educational access globally. Advancements are expected in maths CAIs, as they are widely regarded as particularly suitable for the capabilities of computer programming (Schmidt & Tang, 2020). Existing studies, however, suggest that the actual impact of CAI in middle and high school mathematics has been modest.

Examining the average impact across numerous studies, and focusing only on those with fewer misleading influences on results, Cheung and Slavin (2013) conducted a meta-analysis examining the effectiveness of CAI in maths education. While their analysis covered various age groups, including middle and high school students, they found that the average effect size for CAI interventions was approximately 18% of one standard deviation (ES = 0.18). This translates to less than a full grade increase in most grading systems worldwide, and much less than a point in a grade point average from 0 to 4, failing to represent the significant contribution to tuition that was widely expected.

Limitations of Computer-Aided Instruction in Mathematics

While the promise of CAI is compelling, its impact on the mathematics attainment of middle and high school students has been limited compared to younger age groups. A meta-analysis by Outhwaite et al. (2023) highlights this trend, showing that CAI programs tend to have much larger effect sizes (above 0.50 standard deviations) for students under the age of 11. However, these effects diminish considerably in studies focused on middle and high school students.

Several studies have reported diminishing effects of computer-aided instruction (CAI) as students progress to higher grade levels. A meta-analysis of educational technologies conducted by Tamim et al., (2011) reported that educational technologies, including CAI, tend to have a greater effect on younger students than older ones, with the impact reducing in more advanced educational stages. Cheung & Slavin (2013) also observed that CAI had a larger impact on elementary students compared to middle and high school students, with the effect size diminishing as the age group increased. Similarly, Dynarski et al., (2007) reported that the effectiveness of CAI products in mathematics appeared to decrease in later grades, with high school students showing less significant improvements compared to younger students. Another example is the meta-analysis conducted by Hillmayr et al., (2020), which focussed on secondary school education (including high school). Hillmayr et al., found that the effectiveness of digital tools, such as CAI, in mathematics learning decreases with age, confirming that older students tend to benefit less from these technologies compared to younger learners.

Wang & Hannafin (2021) hypothesise that the gradual diminishing returns of CAI, particularly in secondary education, may be due to factors such as curriculum complexity and decreased engagement from older students. More advanced topics in the mathematics curriculum require higher-order thinking skills. The increased cognitive demands placed on older students may explain why CAI has not had the same transformative effect at this level of education.

One exception to this general trend is the use of Dynamic Geometry Software (DGS) in high school geometry instruction. DGS allows students to manipulate graphical representations of geometric concepts, offering a visual-spatial approach that can be particularly effective for understanding the relationships between geometric shapes and theorems.

A meta-analysis of nine quasi-experimental studies on DGS in high schools found an average effect size of one standard deviation - significantly higher than most other CAI programs (Schmidt & Tang, 2020). However, the effectiveness of DGS is limited to geometry, a single component of the broader high school mathematics curriculum. These results suggest that while CAIs can be effective for specific topics like geometry, a comprehensive system that improves performance across the entire high school mathematics curriculum has yet to be developed.

CAI Case Study

In the UK, one of the most widely used CAI platforms in high schools is Sparx, which has been rigorously evaluated through randomised control trials (RCTs). Sparx offers tailored homework and instructional videos that adapt to the learner's level of proficiency, aiming to support mastery learning in algebra and other mathematics topics. Unlike other CAIs that rely solely on digital instruction, Sparx combines digital interaction with traditional handwritten exercises, encouraging students to write out solutions as part of their learning process.

The effectiveness of Sparx has been studied in a number of trials. One of the most notable evaluations, conducted by Nawaz and Welbourne (2019), found that Sparx contributed to significant improvements in national mathematics exams among Year 7 secondary school students (11-12 years of age). The study used an RCT design, widely regarded as one of the most reliable methods for determining the causal effects of educational interventions. Even after controlling for differences in students' abilities between the treatment and control groups, the study demonstrated a meaningful improvement in mathematics performance for students using Sparx, although the overall effect size was still modest (ES = 0.10).

Another evaluation by Brown and Culora (2021) supported these findings, confirming that Sparx had a positive impact on high school mathematics attainment, but the effect size was again modest at 0.10 standard deviations. These findings align with the broader literature on CAI in mathematics, which suggests that while such programs can contribute to improved performance, the gains are often modest and insufficient to radically transform student achievement in high school.

Targeting Misconceptions to Improve Effectiveness

This shortfall raises the question: how can maths CAIs be made sufficiently effective to substantially improve school outcomes? One possibility, explored by the software investigated in this study, is that CAIs could be more effective if they supplemented teaching by specifically targeting the common errors children make in their developing reasoning. Arguably, the maths curriculum is more linear than most others, with access to advanced topics dependent upon understanding prerequisite concepts and skills, maths curriculums are designed to guide children through this progression.

The approach behind the software in this study suggests that learners often rely on their existing assumptions and familiar heuristics, using techniques that have worked on simpler concepts. If this is the case, children are likely to experience misdirection, where certain features of a problem trigger assumptions that a solution can be found using a simpler, previously learned concept.

These common misconceptions or misapplied techniques may significantly hinder progress. If children are operating under such misapprehensions, it is crucial to identify and address them. This presents a significant challenge, as children will misinterpret in various ways and hold many different kinds of misconceptions (Ay, 2017).

Maths teachers have long recognised the challenge of common misconceptions, where there are patterns but diversity in children's naive interpretations of the curriculum (Barton, 2018; Wang et al., 2021). Multiple-choice diagnostic questions are a reliable vehicle for identifying such misconceptions, with each distractor designed to identify the specific nature of a student's misunderstanding (Barton, 2018; Briggs et al., 2006; Kuo et al., 2016; Wang et al., 2020, 2021). As an example, see Figure 1, a diagnostic question written by Eedi's co-founder, Craig Barton.

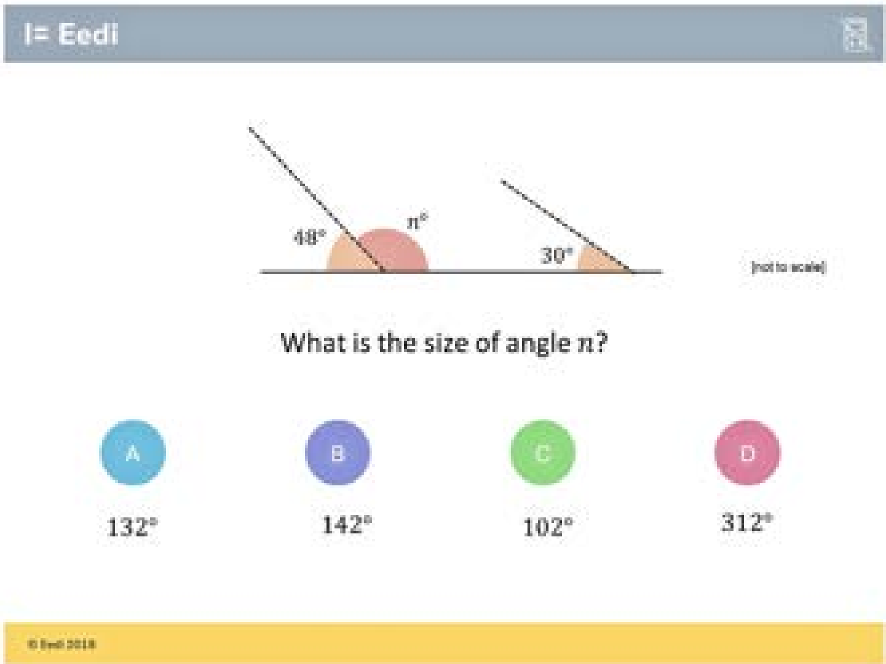

Figure 1. Example of a diagnostic question

In this question, each incorrect option maps to a distinct misconception. A student choosing B may have a misconception with subtraction, a student choosing C may not understand which angles need to be included in the calculator, whereas a student choosing D may have selected the wrong angle relationship.

To maintain the highest quality standards for its question bank, Eedi employs a three-pronged approach:

1. Selection of exemplary questions from diagnosticquestions.com, which houses over 66,000 teacher-created questions;
2. Curation and creation of questions by Eedi's content author team of former mathematics teachers, all trained by Craig Barton. Thus far, Eedi has developed 11,342 expert-generated diagnostic questions;
3. Continuous refinement based on student response data, with underperforming questions undergoing analysis and improvement.

This methodology combines classroom practitioners' observations with data-driven insights and expert analysis, ensuring Eedi's diagnostic questions provide meaningful insights into students' understanding of mathematics. Of course, a teacher could simply ask such questions to their class and identify the specific nature of their misconceptions. But what do they do next? Within a class of 30 students, there is likely to be a wide variety of misconceptions for any given topic. Eedi can identify misconceptions for smaller subgroups of students, and effectively help teachers deliver targeted support.

Summary

The use of computer-aided instruction in high school mathematics has shown some positive effects. Some programs have demonstrated statistically significant improvements in student achievement, with effect sizes generally below 0.20 standard deviations.

To substantially improve high school mathematics outcomes with CAI, future developments must focus on creating more comprehensive systems that address the entire curriculum and are capable of significantly boosting performance across a range of mathematical topics. Additionally, integrating tools that target student misconceptions and provide real-time feedback may help CAIs become more effective for high school learners. Further research is needed to refine these approaches and explore how CAI can better meet the complex needs of high school students in mathematics education.

Eedi Maths Tuition Software

Eedi maths tuition software facilitates teacher-directed or independent work. At its core are diagnostic multiple-choice questions, developed using the framework of Craig Barton (2018). Each question presents students with one correct answer alongside three carefully designed incorrect options that map to common mathematical misconceptions. Through this design, both teachers and Eedi's machine learning systems can precisely identify where students' understanding may have gaps.

The platform adapts to different learning contexts through two primary pathways:

1. **In classrooms**, teachers use Eedi to assign diagnostic quizzes and gather insights into student understanding. The system responds to each student interaction - when answers are incorrect, it provides targeted support through video tutorials and practice exercises; when answers are correct, it offers additional problem-solving materials that extend learning opportunities.
2. **Independent learners** (i.e. on quizzes not assigned by teachers) work with Eedi's Dynamic Quiz technology, a system developed by Wang et al. (2020) that draws on variational autoencoders, a form of factor analysis, to condense historical data on children's correct answers into the strongest relationships between student ability, question difficulty, and question reliability at distinguishing student ability (Wang et al., 2020, 2021). This analysis is used by Eedi to select questions that most effectively distinguish levels of ability. When students study independently on Eedi, they are always given content to match their current level of understanding. When misconceptions emerge through incorrect answers, students receive targeted support materials followed by equivalent diagnostic questions, ensuring concepts are fully grasped before moving forward.

The technical architecture supporting these learning experiences consists of several components. A Partial Variational Auto-Encoder (PVAE) analyses student response patterns and manages sparse data matrices (Wang et al., 2020). Bayesian optimization guides the dynamic quiz system, continuously selecting optimal questions for each student's learning journey (Ma et al., 2018). Additionally, an algorithmic approach based on the Ebbinghaus forgetting curve generates retrieval quizzes, strengthening long-term retention of mathematical concepts.

This synthesis of pedagogical methods and technology creates a responsive learning environment that serves both individual students and their teachers. By identifying specific misconceptions and matching content to student understanding, Eedi supports effective learning in both classroom and independent settings. The result is a structured approach to mathematics education that maintains appropriate challenge levels while fostering sustained engagement with mathematical concepts.

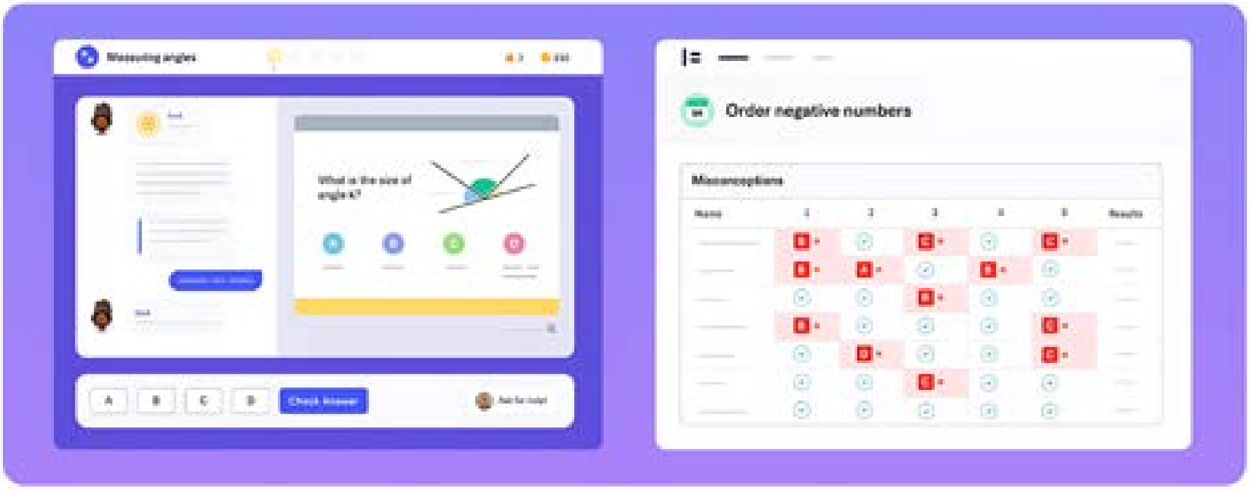

Figure 2. The Eedi chatbot interface for delivering digital lessons and 1:1 human tutoring (left) and an example of class analytics (right)

Eedi’s Theory of Change

Eedi’s goal is to improve student outcomes and empower teachers, creating a dynamic educational ecosystem. Its theory of change outlines a framework illustrating how initial steps and features integrate to achieve long-term educational impacts. It emphasises the interconnectedness of student engagement, personalised learning, teacher support, and data-driven instruction. When students actively engage in diagnostic quizzes, they resolve misconceptions and achieve proficiency in maths topics. Robust teacher support is crucial. Educators equipped with misconception analytics can better facilitate learning, adapt instructional methods, and provide timely feedback. Data-driven instruction allows teachers to make informed decisions based on real-time insights into student performance. This feedback loop continuously refines teaching and learning processes. Data informs students' learning pathways, generating personalised questions that enhance recall and proficiency.

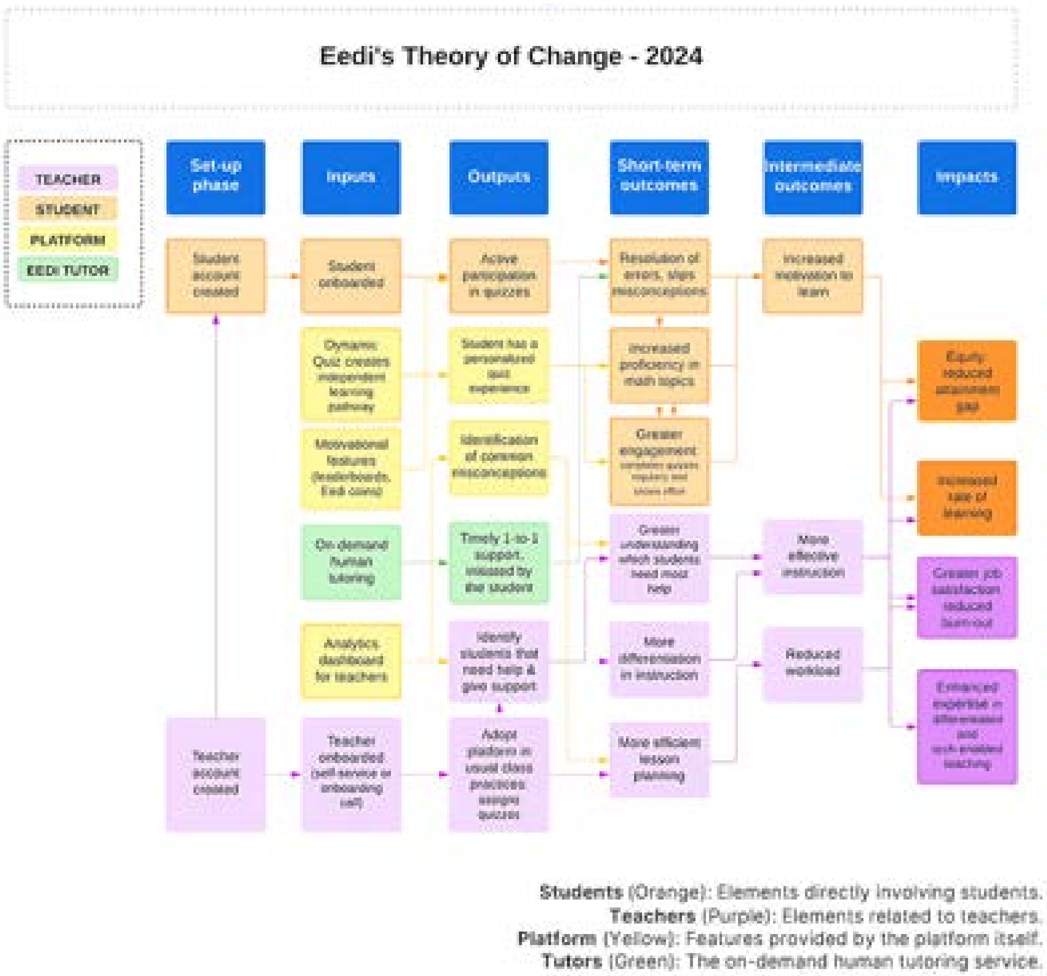

Figure 3. Eedi’s Theory of Change, developed by Eedi (2024)

Set-Up Phase

- **Teacher account created**: The process begins with teachers creating accounts.
- **Teacher onboarded**: Teachers undergo onboarding, either through self-service or an onboarding call, ensuring they are familiar with the platform's functionalities and can effectively integrate it into their teaching practices.
- **Student account created**: Following the teacher setup, student accounts are created.
- **Student logs in**: Students log in to the system, initiating their journey on the platform.

Inputs

- **Student completes diagnostic quizzes**: Students take diagnostic quizzes on topics assigned by their teacher. They can also learn independently by completing quizzes in their personalised learning pathway.
- **Dynamic quiz creation**: Based on the dynamic quiz, the platform creates personalised learning pathways, tailoring the learning experience to students' proficiency levels.
- **Motivational features**: The platform incorporates motivational elements such as leaderboards and Eedi coins to enhance student engagement and motivation.
- **On-demand human tutoring**: Students have access to on-demand human tutoring, providing timely support and personalised assistance when needed.
- **Analytics dashboard for teachers**: An analytics dashboard is provided for teachers, enabling them to monitor student progress, identify areas needing attention, and tailor their instruction accordingly.

Outputs

For students

- **Active participation in quizzes**: Students actively engage with the quizzes, driven by the personalised and motivational features of the platform.
- **Personalised quiz experience**: Incorrect answers trigger lesson flows to address misconceptions. Lessons include fluency practice and help videos.
- **Identification of common misconceptions**: Through student responses, the platform identifies common misconceptions, enabling targeted interventions.
- **Timely 1-to-1 support**: Students can initiate one-on-one support sessions with on-demand tutors, ensuring they receive help when needed most.

For teachers

- **Adoption of platform in class practices**: Teachers integrate the platform into their regular classroom practices, including assigning quizzes and using data to inform their teaching strategies.
- **Identifying students needing help**: The platform helps teachers identify students who need additional support, facilitating timely and effective assistance.

Short-Term Outcomes

For students

- **Resolution of errors and misconceptions**: Targeted interventions help resolve errors and misconceptions, improving student understanding.
- **Increased proficiency in maths topics**: Personalised learning and timely support lead to increased proficiency in assigned maths topics.
- **Greater engagement**: Students show greater engagement with their maths homework, as evidenced by regularly completing quizzes and putting in consistent effort.

For teachers

- **Better understanding of student needs**: Teachers gain a deeper understanding of which students need help and in what areas, allowing for more effective support.
- **More differentiation in instruction**: Understanding which topics specific students require further help allows teachers to offer targeted instruction or intervention.
- **More efficient lesson planning**: Analytics make lesson planning more efficient, reducing workload.

Intermediate Outcomes

For students

- **Increased motivation to learn**: As students' understanding and proficiency improves, students are increasingly motivated to learn.

For teachers

- **More effective instruction**: Instruction targets misconceptions directly to improve student understanding in areas most in need of support.
- **Reduced workload**: Reduced workload as planning becomes more streamlined.

Impacts

For students

- **Equity and reduced attainment gap**: The personalised learning approach helps reduce the attainment gap, promoting educational equity.
- **Increased rate of learning:** The overall rate of learning increases as students receive tailored support and instruction, enhancing their academic performance.

For teachers

- **Greater job satisfaction and reduced burnout**: Improved workload management and student outcomes enhance job satisfaction.
- **Reputation**: Teachers gain a reputation for prioritising student-centred learning and innovative approaches.

Study Rationale

Study Significance

Eedi were successful in receiving a grant from LEVI (Learning Engineering Virtual Institute) to build on their current platform, developing causal discovery methods and applying them to their award-winning dataset to identify causal relationships between misconceptions. As part of the grant, Eedi is testing the effectiveness of their product over a 5-year period with the ambition of doubling the rate of maths progress for middle school (KS3) low-income students.

This study follows on from an initial pilot in Year 1 for a pilot impact study to test the effectiveness of Eedi in KS3 mathematics in UK schools (Harrison, Brown & Higgins, 2023). The study used small scale randomised controlled trials to test the impact of Eedi on specific mathematics topics, using class level randomisation with middle ability classes assigned to receive Eedi or act as a control.

During the short time of the delivery of the intervention for a Key Stage 3 topic in mathematics, students who used Eedi made significantly greater progress in this topic than students who did not. On average, students using Eedi scored 14% higher in tests after using Eedi than in tests prior, compared to 6% for students not using Eedi. This shows that using Eedi has a positive impact improving students' performance in maths tests used in this study.

School	Sample Size	Effect Size (ES)	Standard Error (SE)
School A	62	+0.76	0.26
School D	58	-0.16	0.27
School Y	132	+0.54	0.18
All Schools	252	+0.44	0.13

Table 1. Eedi cumulative evidence base for KS3 mathematics

There are some indications here that Eedi is more effective with PP students (ES 0.50) than with non-PP (ES 0.41), however these indications are tentative due to the relatively small number of Pupil Premium students in the sample.

Analysis of usage indicates higher engagement in School A and Y compared to School D, this appears to convincingly explain the absence of effect in School D. However, the small sample size, short timescales and the use of non-standardised assessments as a proxy for learning were limitations in the design of the study.

This study increases the sample size, implements Eedi over an academic year and uses the standardised NWEA Map Growth Assessments as the primary outcome measure.

Research Questions

Impact Evaluation

The impact evaluation was designed to answer the following questions:

- What is the effect of Eedi on NWEA MAP Growth Maths Attainment in KS3 Year 7?
- Does the effect of Eedi on mathematics vary according to whether pupils are Pupil Premium?
- Does the effect of Eedi on mathematics vary by the time spent on the platform?
- How much does the impact of Eedi vary between schools and classes?

Process Evaluation

The process evaluation component will address the following questions:

- How is the intervention implemented? What are the enablers and barriers to implementation of the programme?
- What constitutes 'usual practice' in intervention and control schools, and does this change over the duration of the trial? Are control schools using similar interventions to Eedi that might be considered close substitutes for it?
- How and why does the implementation of the programme vary? To what extent does any variability affect the achievement of expected outcomes?
- To what extent do teachers link Eedi content to taught lessons?

Methods

Study Design

This study was designed as a randomised controlled trial with whole schools (Year 7) assigned at random to either intervention (exposure to Eedi) or control conditions (non-exposure to Eedi) on a 1:1 basis.

Randomisation

Schools were allocated to intervention and control conditions using random allocation on a 1:1 basis for access to Eedi for the Year 7 cohort for the academic year 2023-24. Randomisation was carried out by the WhatWorked team who were 'blind' as to the identity of schools.

Schools were anonymised to comply with data sharing agreements. Each school was allocated an anonymous ID and Eedi used the DfE performance tables and a WhatWorked School look up to record the current progress 8 score, attainment 8 score, percentage of schools eligible for Pupil Premium (PP), percentage of pupils with English not as their first language (EAL) and school type.

Once the randomisation was complete, the sample was compared based on these characteristics to check that both the intervention and control groups were similar.

Condition	Progress 8 Score	Attainment 8	Pupil Premium (%)	English not First Language (%)
Intervention (n=16)	-0.0058	48.51	35.98	18.5
Control (n=16)	-0.0673	48.88	29.04	18.6

Table 2. School Level Characteristics

Participants

Potential schools were identified using Eedi's own database of teachers and approached to determine their interest in participating in the study. Schools were only recruited if they did not currently use Eedi in Year 7 and if they agreed to participate and withhold the release of Eedi to Year 7 if they were selected as the control. Schools were able to use Eedi in other year groups.

Eedi examined their data records to ensure that the participating schools did not have active pupil accounts for Year 7 pupils and did not create accounts for the control schools before the post-assessment had been completed at the end of the academic year.

In each participating school, the school agreed to use Eedi with all pupils in Year 7 if they were randomly selected as an intervention school. Pupils in the intervention received Eedi support for the remainder of the academic year and the control group could receive access to the Eedi programme once the trial was complete.

Sample

The initial aim was to recruit twenty schools for the evaluation. We assumed that each school had approximately 200 Year 7 pupils, with ten schools allocated as the control (n= 2,000) and ten schools as the intervention (n= 2,000). Therefore, the intervention and control groups sample size would be 2,000 pupils in each arm of the trial and a total sample size of 4,000 pupils.

Thirty-three schools initially agreed to participate. One school was excluded from the evaluation as this was an alternative provision provider with Pupil Premium at 97% and low Progress 8 and attainment scores. The initial sample was therefore of 32 schools.

From this sample, 16 schools were randomly allocated to the intervention and 16 to the control condition. Due to NWEA implementation issues, the sample was reduced to 23 schools. Unfortunately, three additional schools dropped out of the study, leaving eight control and twelve intervention schools. In the intervention, one school was unable to complete the midpoint assessments due to staff illnesses but agreed to complete the post-assessment.

Please refer to Appendix B to see a table of school level characteristics.

Intervention

The Eedi intervention includes targeted quizzes designed to supplement KS3 mathematics. These online quizzes provide students with feedback, short video explanations, and the option to request on-demand 1:1 support from a qualified teacher via a chat-based feature.

The program, intended to complement classroom teaching, can be used during school-directed time or as homework. Eedi offers teachers insights into common misconceptions through diagnostic questions, a presentation function for sharing with the class, and a reporting feature to track student progress.

Schools had flexibility in implementing the intervention, which allowed for measuring its effectiveness in realistic settings. While this approach increases the ecological validity of the study by reflecting how schools are likely to use Eedi, it may also have led to variations in implementation, potentially making it harder to identify an overall effect.

Outcome Measures

The primary outcome for this study was attainment in KS3 Year 7 mathematics as defined by the NWEA MAP Growth Standardised Assessment. The online assessments are completed in the Autumn, Spring and Summer term and are standardised to US measures to allow for international comparisons for the grant funders.

All participating Year 7 pupils completed the 45-minute online pre-test under examination conditions in the initial assessment window in September 2023, mid-point assessment in February 2024 and post-assessment in July 2024.

NWEA Implementation Issues

Eedi had encountered numerous challenges in the set-up and implementation of the MAP Growth Baseline Assessments at the start of the Autumn term, delaying the implementation of the assessments from September 2023 to late October 2023. The main challenge that could not be overcome when implementing the NWEA test online with schools was that the duration of the test exceeded single teaching periods allotted for the test. As a result, some children did not complete the test in a single session, increasing administrative workload for schools, and reducing their willingness to participate.

The project lost 9 schools from the sample of 32 due to these issues and a formal complaint was sent to NWEA (see Appendix A for an overview). Eedi created additional guidance for schools for the mid-point assessment and this did run more smoothly in schools.

Analysis

Data Preparation

This evaluation investigated the impact of the NWEA program on academic performance using a clustered randomised controlled trial (RCT) design. The dataset included pre- and post-test scores, treatment assignment, demographic information, and program usage metrics.

Data Sources and Cleaning Process

Three datasets were merged:

1. **Usage Data**: Records of student usage of the NWEA platform.
2. **Assessment and Demographics**: Containing test scores, Pupil Premium (PP) status, gender, and intervention assignment (with 2871 students at baseline and 2901 students at post-test).
3. **Class Assignment Data**: Used to validate intervention receipt via class-level information.

Data Cleaning Process
After merging the usage and NWEA datasets and incorporating Class ID information, a total of 3,448 unique students were identified. Of these, 547 students were excluded due to missing post-test scores. Missing pre-test scores were imputed using multiple imputation based on school ID, treatment group, and post-test scores. The final analytical sample included 2,901 students.

Estimating Intervention Effects via Multilevel Models (Effect size)

Model Specification: Given the clustered randomised trial (CRT) design, a two-level multilevel model (MLM) was employed to evaluate the impact of the EEDI intervention, with students nested within schools. A random intercept was included at the school level to account for the clustering that resulted from randomisation occurring at the school level (i.e., whole schools were assigned to either the intervention or control group). The model was adjusted for baseline attainment using pre-test scores to ensure the observed differences between treatment and control groups reflect the intervention effect.

To properly adjust for baseline attainment, pre-test scores were partitioned into:

- Pre-test within-school: Student deviation from their own school's mean pre-test score
- Pre-test between-school: Each school's mean pre-test score

This approach controls for individual differences and baseline school-level variation, ensuring an unbiased estimate of the intervention effect.

Effect Size (ES) Calculation: In addition to the MLM analysis, Hedge's g was calculated to quantify the intervention effect. The effect size was derived from the treatment effect estimated by the fixed effects in the MLM, divided by variance computed as the sum of the within-school variance (σ^2) and the between-school closed variance (σ_b^2). This ES accounts for small sample sizes to provide a more accurate estimate of the intervention effect. It is defined as:

$$g = \frac{\beta_2}{\sqrt{\sigma^2 + \sigma_b^2}} J, \qquad \text{and} \qquad CI_g = g \pm Z_{\alpha/2}\left(\sqrt{Var(g)}\right)$$

where $J = 1 - \frac{3}{4(n_1 + n_0 - 2) - 1}$ is the correction factor for small samples and $Z_{\alpha/2}$ = 1.96 and the $Var(g)$ is estimated according to Hedges (2007).

Threshold Analysis for Engagement

The analysis aimed to identify the minimum number of check-in questions required to observe progress in gain scores. Using the threshold-response framework (see Muggeo et al., 2014; Laan, Zhang & Gilbert, 2023; Haynes, Fearnhead & Eckley, 2017), the mean improvement was computed as $E[Y \mid X > v]$ where Y represents the gain score, X is the number of questions answered, and v is a specific threshold. For each value of v, the average gain score was calculated for all students who answered at least v questions. To formally determine the minimum effective threshold, a statistical method was deployed that does not assume independence. This is because the computation of $E[Y \mid X > v]$ includes all observations used to compute $E[Y \mid X > (v+1)]$ along with additional observations corresponding to values of X between v and $v+1$. To account for this dependence, a distribution-free approach was used based on a **Nonparametric Cost Function** to estimate the minimum threshold value.

Sensitivity Analysis

Sensitivity analysis was conducted to assess the effect of the intervention on students with low pre-test scores. The cut-off for low performance was defined using the overall median pre-test score. Median pre-test score was calculated for Pupil Premium and non-Pupil Premium separately. The analysis included only those students whose pre-test scores were below the median, and the effect size was estimated using the same method as in the primary analysis. This analysis aimed to evaluate the impact of the intervention specifically on lower-achieving students.

We performed sensitivity analysis to assess the robustness of our findings by examining the intervention effect among students who completed at least 95 check-in questions compared to control subjects. This threshold-based approach complemented the CACE analysis by providing an alternative method to evaluate intervention effectiveness among engaged participants.

Subgroup analysis was conducted separately for Pupil Premium (PP) and non-Pupil Premium students to examine whether the intervention impact among highly engaged students differed across socioeconomic groups. This analysis helped determine if the platform's effectiveness varied by student background when engagement levels were adequate.

Results from this analysis were used to validate and triangulate findings from the CACE framework, providing converging evidence for the intervention's effectiveness among engaged participants and ensuring the robustness of our conclusions regarding dose-response relationships and equity effects.

CACE Analysis

We employed the Complier Average Causal Effect framework using two-stage least squares estimation with multilevel modelling to account for the clustered design. Random assignment to treatment served as an instrumental variable for treatment compliance, satisfying the key assumptions of relevance, exogeneity, and exclusion restriction.

In the first stage, we modelled compliance as a function of treatment assignment, controlling for baseline pre-test scores (partitioned into within-school and between-school components). In the second stage, we estimated the effect of predicted compliance on post-test outcomes using a multilevel model with random intercepts at the school level.

We defined compliance using multiple binary thresholds calculated using non-parametric function cost function for Pupil Premium and non-Pupil Premium separately and together. This approach allows examination of dose-response relationships and provides estimates of the intervention effect specifically among students who would comply with treatment if assigned. The CACE estimates represent the causal effect for compliers, addressing potential bias from non-compliance that may attenuate intention-to-treat estimates. Effect sizes were calculated as Hedges' g, adjusted for clustering using intracluster correlation coefficients and design effects to account for the nested structure of the data.

Process Evaluation

A light touch process evaluation was conducted by the evaluation team to assess the implementation and feasibility of the evaluation strategy. The process evaluation provided an opportunity to understand any variation in the intervention across the trial schools and how any variation may affect the primary outcome measure.

The process evaluation explored:

- **Fidelity** - Has the programme been delivered as intended by the developer?
- **Dosage** - How much of the intervention has been delivered?
- **Reach** - What is the reach and scope of participation?
- **Alignment to current teaching** - Are schools aligning the intervention with lessons in their classes?
- **Monitoring of control groups** - What other interventions are utilised in intervention schools?

The process evaluation adopted the following methods for data collection:

- Online teacher surveys for intervention and control schools
- Analysis of data from the Eedi data system

To ensure anonymity of schools and teachers, as outlined by Eedi in their Memorandum of Understanding (MOU) agreement with participating schools, the evaluators provided online survey questions and these were sent to participating schools by the Eedi team. The responses were saved in a data folder that was shared between Eedi and the evaluation team with identifiable information anonymised. Teachers were asked to complete the online process evaluation survey at both the midpoint and final NWEA assessment windows for the evaluation. Table 3 provides an overview of the process evaluation timeline.

Date	Activity
September 2023	Online survey designed
February 2024	Online survey completed by intervention and control school teachers
March 2024	Analysis of usage data from the Eedi system (midpoint)
July 2024	Online survey completed by intervention and control school teachers
July 2024	Analysis of usage data from the Eedi system

Table 3. Process evaluation timeline

Ethics and Data Protection

Any risk of harm was mitigated by the design (a wait list approach where all students have the opportunity to experience the intervention). Moreover, learners only participated in teaching and learning activities that were typical for participating schools. As all data was anonymised and the data collected from learners was limited to attainment scores or Pupil Premium eligibility, head teacher consent for the evaluation was deemed to be sufficient. Data was collected on students' use of the Eedi platform. This included some socio-demographic information (for example Pupil Premium eligibility, which schools provided to Eedi). Student data was not been linked to the National Pupil Database.

Legal basis for processing

The impact evaluation complied with the GDPR legal basis for processing personal data as this research project is in the public interest. Data sharing was necessary for the parties to undertake a research project into the effectiveness of the Eedi platform aimed at students in KS3.

The project was in the public's interest as results will help assess the performance of Eedi on student achievement. The Eedi platform is designed to deliver online maths homework assignments and provide one to one targeted academic support. The collection and sharing of data from students participating in the research project was necessary for the parties to evaluate the effectiveness and impact of Eedi on students' attainment and achievement in maths. In addition, the collection and sharing of student data and surveys from teachers participating in the project was necessary to assess the process evaluation. All pupil and school-level data was anonymised before sharing with the WhatWorked evaluation team.

Summary

Impact Evaluation

The impact evaluation aimed to assess the impact of Eedi on attainment in mathematics. Overall, the impact evaluation found that students using Eedi showed significant improvements in NWEA scores compared to those not using the platform. Higher engagement, particularly answering more check-in questions, was linked to better outcomes for both Pupil Premium and non-Pupil Premium students as shown in CACE analysis.

Key Findings

Impact

Eedi had an effect size of 0.17, meaning students in Eedi schools scored, on average, 0.17 standard deviations higher than those in control schools using other or no maths platforms. This suggests that Eedi was more effective than alternative computer-assisted learning platforms.

Pupil Premium Students

While the data showed a positive trend in NWEA scores among Pupil Premium (PP) students using Eedi, this improvement was not statistically significant. Moreover, the interaction between intervention and PP status was not significant, indicating that the effect of Eedi did not differ between PP and non-PP students. In other words, Eedi was equally effective for both Pupil Premium and non-Pupil Premium students

Usage

A positive relationship was found between the higher engagement based on CACE analysis and improvement in NWEA scores. Students who engaged with the platform by answering more check-in questions showed greater improvements in maths scores as shown in the CACE analysis. Higher engagement with the platform led to better outcomes among PP and non-PP students.

School-Level Variation

There is variation in improvement across schools, with some Eedi-using schools achieving larger gains (e.g., Schools 1 and 10). This suggests that the success of Eedi may depend on school-specific factors, such as implementation and integration into the curriculum.

Low Pre-Test Scores

The intervention showed a positive and significant effect among students with pre-test scores below the median. These results suggest that the intervention can benefit lower-achieving students when appropriately targeted.

Summary

The findings from this study suggest that on average, students using Eedi showed a greater improvement in their NWEA scores than their peers in the control group. This difference was statistically significant. The consistent positive trend, after accounting for prior attainment, indicates encouraging early evidence of Eedi's potential to support maths learning.

This indicates that Eedi can be an effective supplementary tool for enhancing student performance in mathematics, though the magnitude of improvement remains modest, as reflected in the effect size (Cohen's d = 0.17), which suggests a small but meaningful impact.

These findings align with previous research on the effectiveness of computer-aided instruction (CAI), which has shown that CAI can provide benefits in mathematics education, particularly when students are given the opportunity to engage with targeted, interactive learning tools. Whilst the effect size reported in this study (d = 0.17) is modest, this is consistent with broader research (Cheung & Slavin, 2013).

Subgroup analysis of PP and non-PP students indicated that Eedi had a positive impact on both groups. However, the effect size was smaller for PP students compared to non-PP students, which may be explained by lower engagement levels among the PP group. However, 48.15% of Pupil Premium students reached the 95-question threshold, a proportion that surpasses the 5% benchmark of typical usage reported in literature reviewing student engagement with similar platforms (Holt, 2024).

The analysis of check-in questions answered suggests that the number of check-ins completed by PP students may influence their outcomes. The CACE analysis demonstrated a clear relationship between average improvement in NWEA and amount of Eedi check-in questions answered for PP and non-PP students. Sensitivity analysis showed that students who answered at least 95 questions had higher scores compared to control group students. This difference was statistically significant. This suggests that higher engagement (measured by the number of questions attempted) is associated with greater attainment scores.

Analysis of students with low pre-test scores showed a positive impact of the intervention compared to the control group, suggesting that Eedi can be beneficial for students with lower academic performance.

The trend observed in Figure 5, which shows a positive relationship between the number of check-in questions answered and improvement in NWEA scores, highlights the importance of student engagement with Eedi. PP students who answered more check-in questions showed greater improvement compared to those who answered fewer questions (Figure 6). This finding is consistent with existing literature that emphasises the role of active engagement in CAI (Hillmayr et al., 2020). Students who are more engaged with the platform, particularly in completing check-in questions that reinforce learning, are likely to benefit more from its features. The challenge is in encouraging consistent use of the platform, particularly among PP students who may face additional barriers to regular engagement. Addressing this gap in usage could be crucial in leveraging Eedi to close the performance gap between disadvantaged and non-disadvantaged students.

The data also reveals differences in outcomes between schools, with some schools showing greater improvement in NWEA scores than others. This variability suggests that the effectiveness of Eedi may also depend on school-specific factors, such as how well the platform is integrated into the broader curriculum, the level of teacher support provided, and the overall school environment. Schools that achieved the highest gains may have implemented Eedi more effectively, or they may have other support structures in place that complement the platform's use. Future research could explore these school-level differences in more detail to identify best practices for maximising the impact of Eedi.

Results

Impact Evaluation

Key Characteristics of the Students in the Study

Baseline and final analytic samples are presented in Tables 2 and 3. Table 4 summarises baseline sample characteristics (n = 3448), while Table 5 presents the final analytic sample (n = 2901), excluding all cases with missing post-test data.

Descriptive Statistics at Baseline

At baseline, pre-test data was available for 2,871 students. Among these, 547 students did not have corresponding post-test data. Additionally, 577 students had missing pre-test data. In total, data were collected on 3,448 students. Additional baseline characteristics, including summaries stratified by Pupil Premium status, are presented in Table 4.

Characteristics		Control	Treatment
School (N)		8	12
All Pupils (N)		1226	2222
Gender	Male	412	618
	Female	290	841
	Unspecified	524	763
Pre-test Mean (SD)		210.39 (14.4)	211.46 (15.6)
Adjusted Post-test Mean (SD)		215.79 (15.78)	216.9 (16.78)
Pupil Premium (PP)	PP (N)	434	617
	Non-PP (N)	790	1592
	Pupil-level (%)	35.4%	27.77%

Table 4: Distribution of students who participated in the trial by Pupil Premium (PP) Status

Descriptive Statistics using Final Sample

Not all students with pre-test scores had corresponding post-test data. After excluding those with missing post-test scores, the final analytic sample consisted of 2,901 students across 20 schools - comprising 8 control schools and 12 intervention schools. The number of students in the intervention group (n = 1,869) was higher than in the control group (n = 1,032). Additional baseline characteristics, including summaries stratified by Pupil Premium status, are presented in Table 5.

		Control	Treatment
School (N)		8	12
All Pupils (N)		1032	1869
Gender	Male	342	514
	Female	255	702
	Unspecified	435	653
Pre-test Mean (SD)		210.92 (13.72)	211.6 (15.26)
Adjusted Post-test Mean (SD)		215.79 (15.78)	216.9 (16.78)
Pupil Premium (PP)	PP (N)	349	513
	Non-PP (N)	681	1344
	Pupil-level (%)	33.82%	27.45%

Table 5: Distribution of students who participated in the trial by Pupil Premium (PP) Status

Descriptive Statistics Grouped by Pupil Premium Status

Table 6 presents descriptive statistics stratified by Pupil Premium (PP) status. On average, students eligible for Pupil Premium had a pre-test score of 206.81, while non-PP students scored 213.26. The post-test scores were 210.69 for PP students and 219.07 for non-PP students. This corresponds to a mean gain of 3.87 points in the PP group and 5.82 points in the non-PP group.

	Descriptive Statistics	Pupil Premium	Non-Pupil Premium
Pre-test	Mean	206.81	213.26
	SD	14.78	14.34
	Median	207.5	214
Post-test	Mean	210.69	219.07
	SD	16.60	15.68
	Median	212	220
Mean Gain	Mean	3.87	5.82
	SD	10.07	8.56
	Median	5	6

Table 6: Descriptive statistics grouped by Pupil Premium status

Primary Analysis

Students who received the intervention scored, on average, 1.45 (0.07, 2.83) points higher on the post-test compared to the control group, after adjusting for baseline performance (Table 7). The corresponding effect size was 0.17 (0.01, 0.32), calculated using the estimated treatment effect 1.45 (0.07, 2.83) and the square root of the total variance from the multilevel model. As shown in Table 8, the intervention group performed better than the control group with an adjusted percentage difference of 0.67%.

Parameters	Estimate	95% LB	95% UB
Pre Score (between-cluster)	1.10	1.01	1.20
Pre Score (within-cluster)	0.88	0.85	0.90
Treatment	1.45	0.07	2.83

Table 7: Parameter estimates (Primary Analysis)

Condition	Sample Size	Adjusted Mean	Raw Standard Deviation	Effect Size	% Difference
Treatment	1869	216.96	16.78	0.17 (0.01, 0.32)	0.67%
Control	1032	215.52	15.78		

Table 8: Effect size and percentage difference

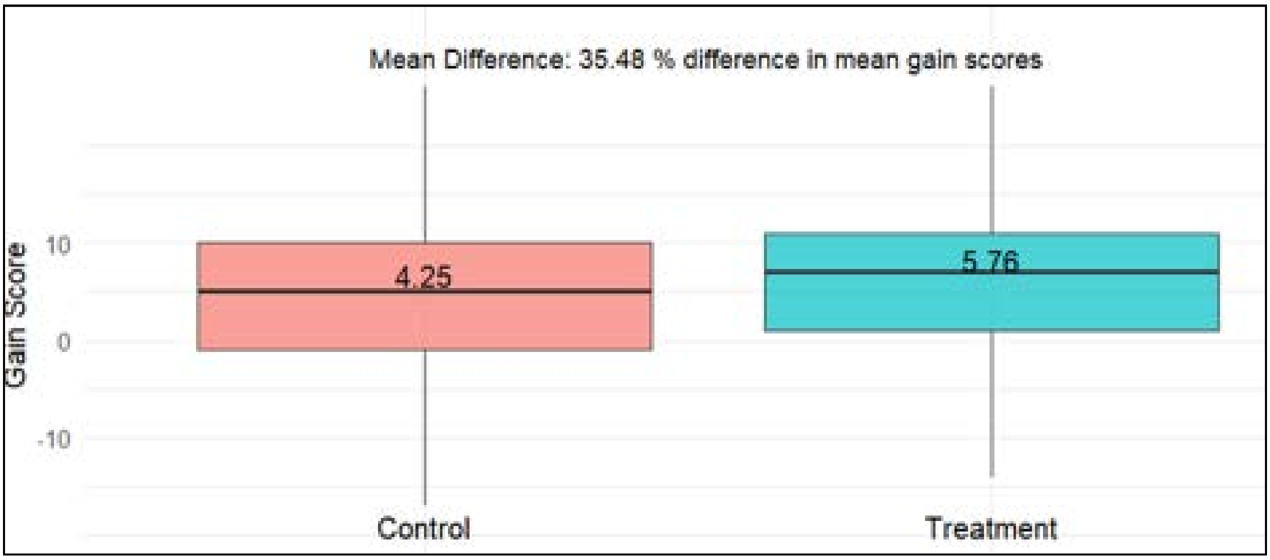

Figure 4 . Boxplot of percentage in mean gain scores

Secondary Analysis

Pupil Premium: Interaction Analysis

To test whether the intervention had differential effects based on Pupil Premium status, an interaction term was added to the model:

$$Y_{ij} = \beta_0 + \beta_1 Y0ij + \beta_2 Tij + \beta_3 PPij + \beta_4 Tij \cdot PPij + b_{0i} + \epsilon_{ij},$$

where PP_{ij} represents the Pupil-Premium status of pupil *j* from school *i* .

The analysis examined whether the intervention had a differential effect on students receiving Pupil Premium. There was some indication that the intervention may have been slightly more effective for non-Pupil Premium students; however, this difference was not statistically significant (see Table 9). Therefore, there was insufficient evidence to conclude that the intervention had a different impact on the two groups (-0.69, 95% CI: -2.18 to 0.81)

Pupil Premium: Subgroup Analysis

Although the interaction analysis did not demonstrate statistically significant differential effects, a subgroup analysis was conducted for exploratory purposes (Table 10). The same model used for primary analysis was applied separately to each Pupil Premium (PP) subgroup. It is important to note that while the raw standard deviations are presented in the table, the effect sizes were derived using the treatment effects from Table 9 divided by the square root of the total variances from the multilevel model. Subgroup analysis showed that effect of intervention is positive for both premium and non-premium subjects.

Parameters	Estimate	95% LB	95% UB
Pre Score (between-cluster)	1.09	1.00	1.18
Pre Score (within-cluster)	0.86	0.84	0.89
Treatment	1,52	0.12	2,91
PP	-1.82	-2.99	-0.65
Treatment TRUE:PP	-0.69	-2.18	0.81

Table 9. Parameter estimates (Interaction Model)

PP Status	Condition	Sample Size	Adjusted Mean	Raw Standard Deviation	Effect Size	% Difference
Non-PP	Treatment	1344	219.46	16.01	0.18 (0.02, 0.35)	0.70%
	Control	681	217.94	15.00		
PP	Treatment	513	211.21	16.96	0.12 (-0.09, 0.32)	0.54%
	Control	349	210.08	16.09		

Table 10. Conditional effect size across Pupil-premium groups

School Level Analysis

As the randomisation was conducted at the school level, with each school assigned entirely to either the intervention or control condition, it was not possible to conduct comparative analyses between groups at the school level.

Consequently, the analysis was limited to descriptive statistics, including raw means, sample sizes, and standard deviations for pre-test, post-test, and gain scores. The data reveals variation in pupil progress across both intervention and control schools (Table 11). Within the intervention group, several schools demonstrated strong gains.

For example, School 1 achieved an average improvement of 7.49 points and School 10 recorded a gain of 7.42 points. In contrast, more modest progress was observed in schools such as School 4, with a gain of 2.64 points, and School 2, with 4.34 points.

These differences suggest that although the intervention was generally associated with positive outcomes, its effectiveness varied across schools. This variation may reflect differences in implementation fidelity, pupil engagement, or broader contextual factors. In the control group, pupil progress also showed variability, with gain scores ranging from 1.94 points in School 30 to 6.71 points in School 11. Table 11 provides scores for each participating school in the intervention and control conditions.

Condition	School	Sample Size	Raw Pre-Test Mean (SD)	Raw Post-Test Mean (SD)	Mean Gain Score (SD)
Control	6	122	211.78 (13.57)	214.66 (14.07)	2.88 (8.83)
	7	103	212.87 (14.25)	219.46 (13.37)	6.58 (8.21)
	11	75	212.16 (12.35)	218.87 (12.9)	6.71 (7.69)
	15	89	212.51 (11.53)	216.1 (13.74)	3.6 (9.83)
	16	186	219.94 (11.17)	226.64 (12.47)	6.7 (8.42)
	24	210	206.67 (12.77)	210.33 (15.76)	3.66 (11.32)
	30	132	206.3 (15)	208.24 (17.23)	1.94 (10.55)
	31	115	210.25 (13.47)	212.57 (15.42)	2.32 (9.11)
Treatment	1	298	214.11 (13.11)	221.59 (13.44)	7.49 (7.28)
	2	128	204.41 (15.98)	208.76 (17.97)	4.34 (10.24)
	4	168	208.5 (14.41)	211.14 (14.67)	2.64 (9.63)

Condition	School	Sample Size	Raw Pre-Test Mean (SD)	Raw Post-Test Mean (SD)	Mean Gain Score (SD)
Treatment	10	130	210.75 (15.99)	218.17 (19.14)	7.42 (9.34)
	12	161	210.99 (14.44)	215.69 (16.15)	4.7 (9.07)
	13	171	201.92 (10.21)	207.22 (11.9)	5.29 (7.96)
	17	137	233.13 (7.79)	239.93 (8.19)	6.8 (6.27)
	19	146	219.15 (12.87)	225.34 (13.88)	6.19 (7.88)
	20	94	207.3 (12.53)	212.51 (13.39)	5.21 (8.28)
	28	138	200.36 (11.14)	205.67 (12.35)	5.31 (8.78)
	29	130	211.58 (13.76)	217.02 (16.43)	5.45 (9.27)
	32	168	209.17 (13.51)	216.01 (14.16)	6.84 (9.3)

Table 11. Impact at school-level

Threshold Estimation

A nonparametric cost function approach was used to estimate the threshold for Pupil Premium (PP) and non-Pupil Premium students based on the number of questions attempted, denoted as v, ranging from 0 to 370. Since fewer than 30 PP students answered more than 370 questions, this value was set as the upper limit for threshold estimation. The analysis identified 95 check-in questions as the estimated threshold for all students, as indicated by the blue vertical line in Figure 5.

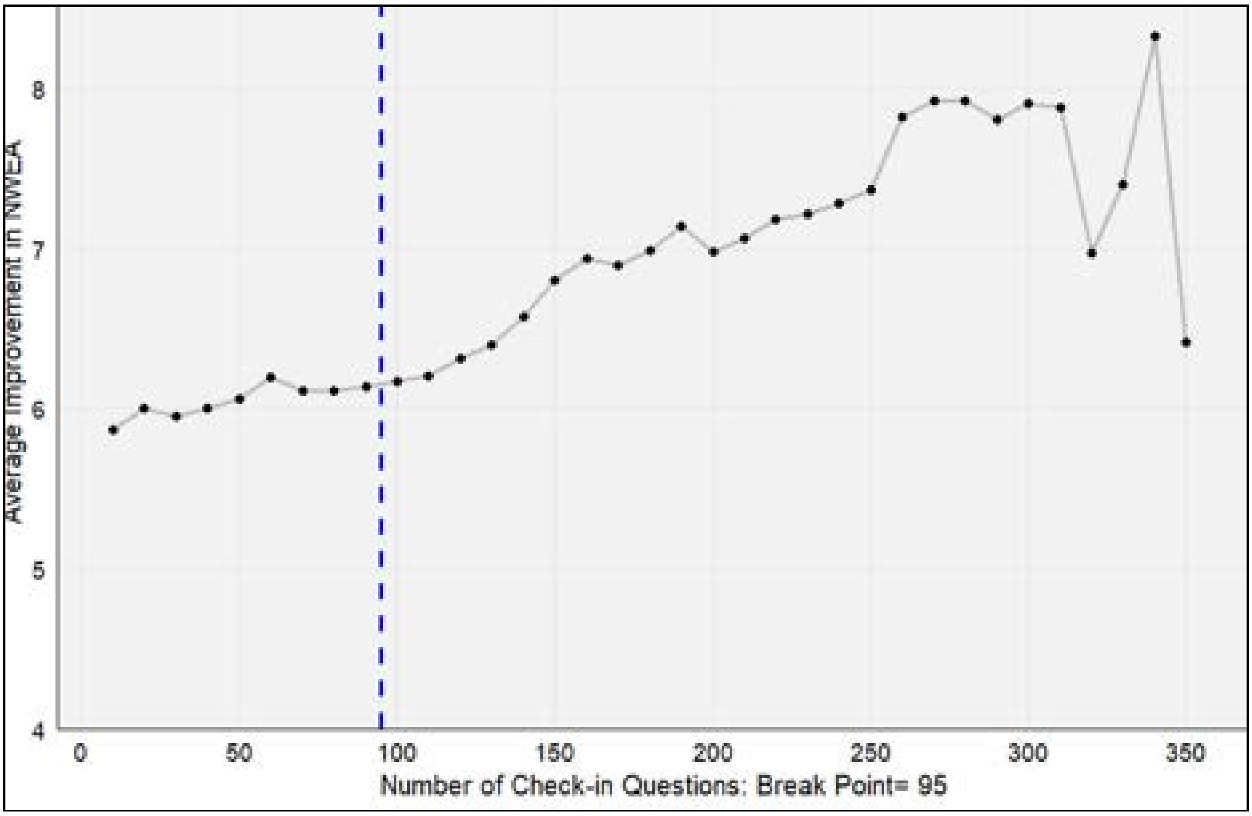

Figure 5. Mean gain score vs. number of Check-In questions for PP and non-PP students. Vertical line indicates threshold at 95 questions

A nonparametric cost function approach was used to estimate the threshold for Pupil Premium (PP) students based on the number of questions attempted, denoted as v, ranging from 0 to 310. Since fewer than 30 PP students answered more than 310 questions, this value was set as the upper limit for threshold estimation. The analysis identified 120 check-in questions as the estimated threshold for PP students, as indicated by the blue vertical line in Figure 6.

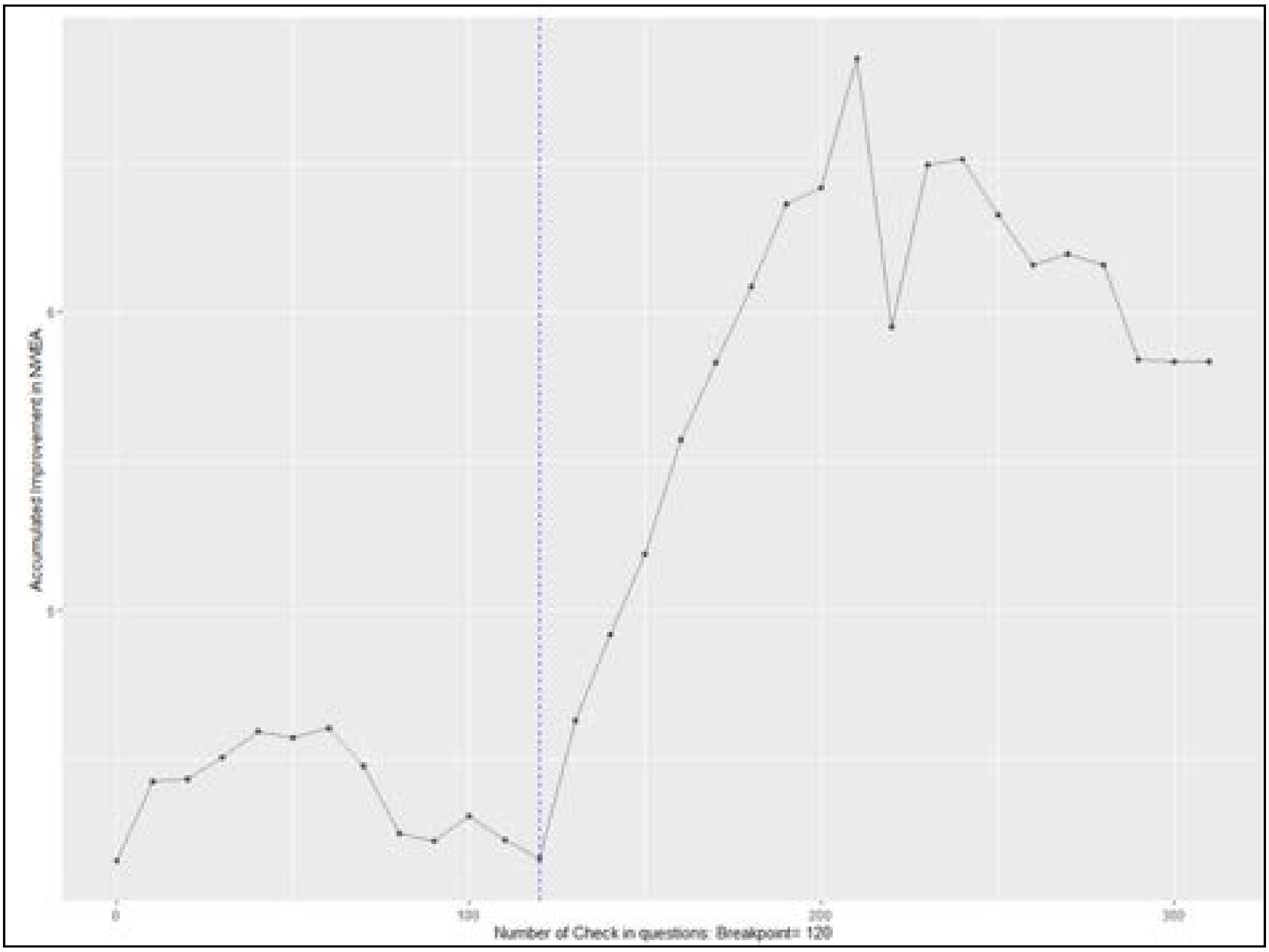

Figure 6. Mean gain score vs. number of Check-In questions for PP students. Vertical line indicates threshold at 120 questions

Check-In Questions and Responses by Pupil Premium Status

Table 12 presents a breakdown of student engagement with check-in questions by all students. The analysis evaluated the number of students who answered at least 95 check-in questions, considering both all responses and only those that were correct.

Questions	PPFSM Indicator	Total Questions Answered	Pupils (N)	Pupil answering less than 95 questions, N (%)	Pupil answering more than 95 questions, N (%)
All questions	Non-PP	185326	1344	480 (35.71%)	855 (63.62%)
	PP	54371	513	247 (48.15%)	247 (48.15%)
	Missing	2774	12	0 (0%)	12 (100%)
Only those with correct answers	Non-PP	153990	1331	622 (46.73%)	709 (53.27%)
	PP	42816	490	305 (62.24%)	185 (37.76%)
	Missing	2104	12	2 (16.67%)	10 (83.33%)

Table 12. Student Engagement with Check-in Questions: Threshold of 95 questions.

All questions: 513 PP students answered a total of 54371 check-in questions. Among them, 247 (48.15%) answered 95 or more questions. 1344 non-PP students answered a total of 185326 check-in questions. Among them, 855 (63.62%) answered 95 or more questions.

Correctly answered questions: 490 PP students correctly answered a total of 42816 check-in questions. Among them, 185 (37.76%) answered 95 or more questions. A total of 1,331 non-PP students correctly answered 153990 check-in questions. Among them, 709 students (53.27%) answered 95 or more questions.

CACE Analysis

Table 13 presents the findings of the CACE analysis using multiple cut-off for compliance. At all cut-off, the effect of the intervention is positive among subjects who comply with the intervention compared to the control group. The CACE estimates are consistently larger than the ITT estimates, reflecting the concentrated effect among compliant participants. As the compliance threshold increases, the effect size grows, suggesting that higher levels of engagement with the intervention yield greater benefits.

Tables 14 and 15 present the CACE analysis findings for PP and non-PP students, respectively. Across all engagement cut-off points, the effect of Eedi was positive in both groups. Moreover, the effect size increased with higher levels of engagement, indicating that greater engagement with Eedi was associated with improved student outcomes.

These results demonstrate that while the ITT analysis provides a conservative estimate of the intervention's effectiveness in a real-world implementation scenario, the intervention shows substantial efficacy among those who actively engage with it.

Compliance cut-off used	Compliers (N)	Effect Size
≥ 78	1243 (66.5%)	0.247(-0.065,0.399)
≥ 95	1114 (59.6%)	0.276(0.121,0.431)
≥ 120	917 (49.1%)	0.336(0.174,0.498)

Table 13: CACE Analysis findings

Compliance cut-off used	Compliers (N)	Effect Size
≥ 78	282 (55%)	0.171 (-0.027, 0.37)
≥ 95	247 (48.1%)	0.195 (-0.011, 0.402)
≥ 120	194 (37.8%)	0.249 (0.025, 0.472)

Table 14: CACE Analysis findings (PP Students)

Compliance cut-off used	Compliers (N)	Effect Size
≥ 78	949 (70.6%)	0.268 (0.112, 0.425)
≥ 95	855 (63.6%)	0.298 (0.14, 0.456)
≥ 120	712 (53%)	0.359 (0.195, 0.523)

Table 15: CACE Analysis findings (Non-PP Students)

Sensitivity Analysis

Inclusion of Students with Low Pre-Test Scores

We performed sensitivity analysis by including subjects scored less than median pre-score. After adjusting for baseline performance, students in the intervention group scored, on average, 2.17 points higher on the post-test compared to those in the control group (95% CI: 0.54 to 3.79; see Table 16). The corresponding effect size was 0.23 (95% CI: 0.07 to 0.39). As shown in Table 17, the intervention group also outperformed the control group with an adjusted percentage difference of 1.06%.

Parameters	Estimate	95% LB	95% UB
Pre-test score (between cluster)	1.63	1.30	1.97
Pre-test score (within cluster)	0.89	0.84	0.94
Treatment	2.17	0.54	3.79

Table 16: Parameter estimates (Sensitivity Analysis)

Condition	Sample Size	Adjusted Mean	Raw Standard Deviation	Effect Size	% Difference
Treatment	958	206.96	13.06	0.23 (0.07, 0.39)	1.06%
Control	517	204.79	13.51		

Table 17: Effect size and percentage difference (Sensitivity Analysis)

Summary

Implementation & Process Evaluation

The process evaluation aimed to assess the implementation and feasibility of Eedi across participating schools. Overall, the platform was well-received by teachers and students, with high levels of engagement reported. However, variability in teacher engagement, content suitability, and technical issues posed challenges that may have impacted the platform's effectiveness.

Key Findings

Ease of Use

- 64% of teachers attended the training sessions, which were well-received and helped facilitate the use of Eedi.
- 78% of teachers found the platform easy to set up and navigate, aiding in its integration into teaching practices.

Teacher Engagement

- 43 out of 64 teachers used Eedi data to inform their teaching, with 36% using it for lesson planning and 23% for whole-class feedback.
- 21 teachers did not use Eedi results at all, indicating a range of engagement levels.

Student Engagement

- Students were engaged with Eedi due to its interactive and competitive features, such as leaderboards and avatars, which increased participation.

Barriers to Implementation

- Technical issues, such as platform stability, hindered the smooth implementation of Eedi in some cases.
- The rapid increase in difficulty within the content was challenging for lower-ability students, limiting the platform's accessibility for this group.

Platform Usage in Control Schools

- Control schools frequently used platforms like MathsWatch and Sparx, which served as close substitutes for Eedi, potentially affecting trial outcomes.

Summary

The process evaluation of the Eedi intervention reveals key insights into how the programme was implemented and the factors that influenced its success. The evaluation was minimally intrusive, relying on online surveys completed by both control and intervention groups. With a 91% response rate from teachers, the findings provide a comprehensive overview of the enablers and barriers to Eedi's implementation and its feasibility as a tool in classroom settings.

The successful implementation of Eedi was supported by several factors. A majority of teachers (64%) attended the training sessions, which were well-received and helped facilitate the onboarding process. The platform's ease of use was also a major enabler, with 78% of teachers reporting that it was easy to set up and navigate. Teachers who actively integrated Eedi into their teaching found it helpful, with many using the platform's data on misconceptions to inform their lesson planning. For example, 36% of teachers used Eedi to adjust their future lessons, demonstrating a direct link between the platform's insights and classroom practice.

However, not all teachers engaged with Eedi in the same way. While many saw value in its data-driven feedback, others did not fully integrate the platform into their teaching routines, with 21 out of 64 respondents reporting that they did not use Eedi's data at all. This variability in engagement highlights differing approaches to implementation, suggesting that Eedi was used as a supplementary tool rather than a core component of teaching for some educators.

Several enablers supported the successful use of Eedi. In addition to the platform's user-friendly design and the positive reception of the training sessions, student engagement played a crucial role. Teachers reported that students enjoyed Eedi's interactive and competitive elements, such as leaderboards and avatar customisation, which likely enhanced their willingness to participate in the platform's activities.

Technical issues, though not widespread, posed challenges for some teachers, particularly in maintaining platform stability and ensuring smooth access for students. Moreover, the content's difficulty level was noted as a concern, especially for lower-ability students.

Some teachers expressed frustration that the questions became too difficult too quickly, limiting the platform's effectiveness for a subset of students. Additionally, time constraints in already packed curriculum schedules made it difficult for some teachers to fully integrate Eedi into their teaching.

The process evaluation also sheds light on the 'business as usual' practices in control schools, which were similar to intervention schools in terms of homework frequency and duration. Both groups typically set maths homework once a week, with an average expected completion time of 30 minutes. However, while control schools did not use Eedi, they employed a variety of other online platforms, such as MathsWatch and Sparx, which offer similar features to Eedi. This suggests that control schools were using interventions that may have served as close substitutes, potentially influencing the trial's ability to detect significant differences between the groups.

.

Results

Implementation & Process Evaluation

The process evaluation for this research was designed to be minimally intrusive, employing only online surveys as the primary means of data collection for the implementation and process evaluation.

Both control and intervention classes were requested to complete a brief survey at the mid-point and upon the completion of the programme. The mid-point surveys were used to inform how Eedi was used and to understand what constituted 'business as usual' in control schools.

The process evaluation in this report focuses on the responses at the completion of the programme. The response rate of teachers completing the online survey was 91%, with one hundred and twenty-nine responses from the one hundred and forty-one surveys sent. The online surveys for the control and intervention classes can be found in Appendix E.

How is the intervention implemented? What are the enablers and barriers to implementation of the programme? Is implementing the programme feasible?

Implementation of the intervention

The implementation of the Eedi intervention was carried out by providing teachers in the intervention group with access to the Eedi platform, accompanied by training sessions designed to facilitate its use. A majority of teachers (64%) participated in these training sessions, which were generally well-received. One teacher noted, "Overall, the training session was helpful," with another agreeing that it was "pitched at the right level." This positive reception likely contributed to the successful onboarding of teachers. Moreover, most teachers (78%) found the platform easy to set up, with one commenting, "The Eedi platform was generally easy to set up, navigate, and use for assigning tasks and reviewing results."

However, the extent to which teachers integrated Eedi into their teaching varied. While 43 out of 64 respondents used the results from Eedi assignments to inform their teaching, 21 teachers did not use these results at all.

Among those who did use Eedi data, 36% used the misconceptions identified by the platform to inform future planning, and 23% provided whole-class feedback by going through problematic questions with students. One teacher mentioned, “I used the misconceptions Eedi uncovered to inform my future planning,” indicating how the platform’s insights were directly linked to lesson planning. Yet, the varying degrees of engagement highlight a divergence in how central Eedi was to different teachers' instructional approaches.

Enablers and barriers to implementation

Several factors enabled the effective implementation of the Eedi intervention. First and foremost, the platform’s ease of use was a significant enabler. Teachers generally found Eedi straightforward to navigate and integrate into their existing practices, which reduced the potential disruption to their teaching routines. Additionally, the positive reception of the training sessions provided a solid foundation for teachers to start using the platform effectively. As one teacher noted, “The training session was helpful and made it easier to get started with Eedi.”

Student engagement with Eedi also served as an enabler. Many teachers reported that students enjoyed using the platform, particularly its interactive and competitive elements. One teacher observed, “Students enjoyed the interactive and competitive elements like leaderboards and avatar/character customisation,” which likely increased students’ willingness to engage with the content provided through Eedi.

However, barriers were also noted. Technical issues, though not widespread, did present challenges when they occurred. Some teachers reported encountering IT glitches and difficulties with platform stability, which hindered the smooth implementation of the intervention. A teacher pointed out, “There were some IT glitches, slow updates, and difficulties with platform stability, which affected student access from home.” Additionally, the content’s suitability was potentially another barrier, particularly for lower-ability students. Some teachers expressed concerns that the difficulty level of the questions increased too rapidly and that there was a lack of customisation options to tailor the content to different student needs. One teacher remarked, “As a teacher of low-ability classes, the questions became very hard very quickly.”

Moreover, time constraints and challenges in integrating Eedi into existing curriculum schedules were highlighted as barriers. Some teachers found it difficult to balance the use of Eedi with other teaching responsibilities, making it harder to fully utilise the platform. A teacher explained, "It was challenging to integrate Eedi into our already full curriculum schedules."

Feasibility of implementation

The feasibility of implementing the Eedi programme appears promising, particularly given the overall positive feedback regarding its ease of use and the supportive training provided. However, the feasibility is not without challenges. Addressing the technical issues and ensuring the content is accessible and effective for students of varying abilities are crucial steps that need to be taken to enhance the programme's usability. One teacher's comment encapsulates this need: "The platform was a good idea in general, but it needs to be refined, particularly in terms of content customization for different student levels."

While the programme has shown potential benefits, especially in terms of engaging students and helping teachers identify and address misconceptions, the success of broader implementation would depend on overcoming the identified barriers. Teachers' willingness to continue using Eedi, as expressed in several feedback comments, suggests that with the right adjustments, the programme could be effectively scaled. One teacher summarised this sentiment: "Eedi is a good platform, especially for higher-ability students, and it has potential for effective teaching if certain areas are improved."

What constitutes 'usual practice' in the intervention and control schools, and does this change over the duration of the trial? Are control classes using similar interventions to Eedi that might be considered close substitutes for it?

Usual practice in intervention and control schools

Maths Homework Frequency

Reported frequency of maths homework was similar across intervention and control schools, with responding teachers in both groups reporting that maths homework is typically set by class teachers once a week (intervention n=56, control n=50).

In both intervention and control schools, the usual practice for setting maths homework involves using online platforms, although the specific platforms and their usage frequency differ between the two groups. Teachers from both groups reported that maths homework was typically set once a week. In intervention schools, 89% of teachers set homework weekly, compared to 77% in control schools. This suggests a similar baseline practice in terms of homework frequency. Figure 6 shows the usual practice of setting maths homework in both control and intervention classes.

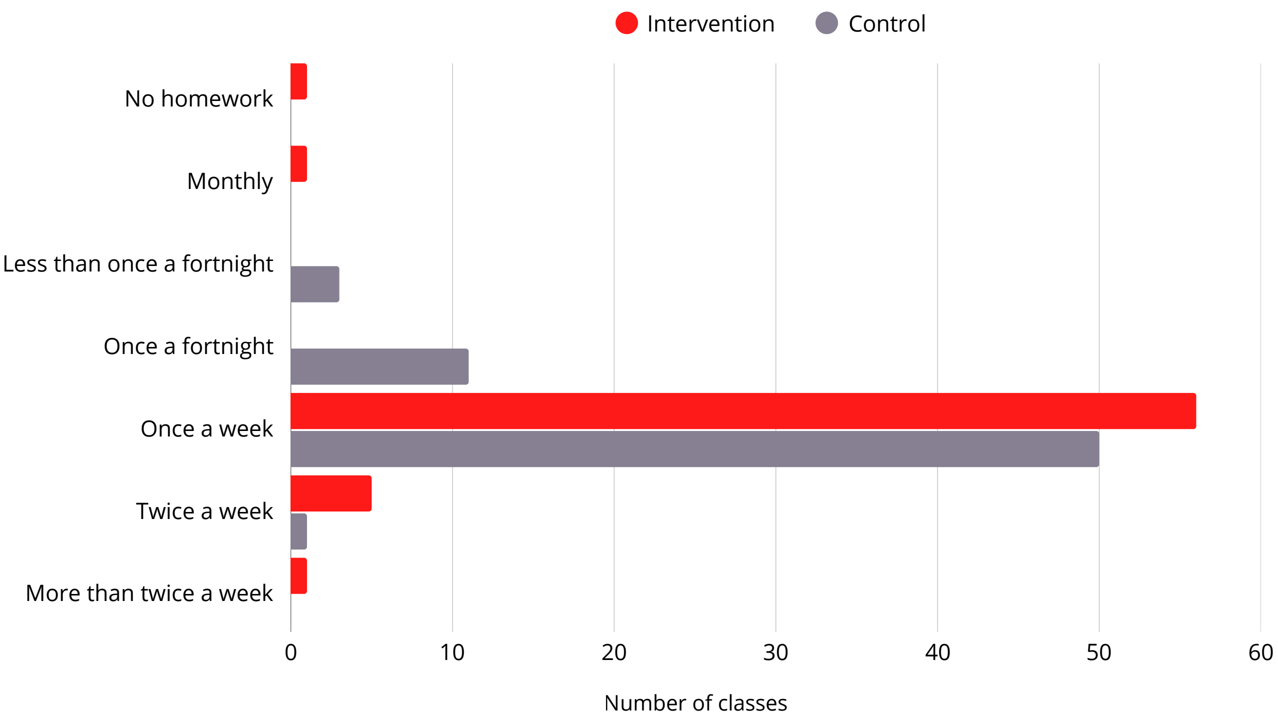

Figure 6. Frequency of setting maths homework in control and intervention conditions

Maths homework duration

Reported duration of maths homework was similar across intervention and control schools, with responding teachers in both groups reporting that maths homework should take students approximately 30 minutes to complete (intervention n=34, control n=35).

Students in intervention groups were more likely to receive homework of a shorter duration, whereas students in control groups were more likely to receive homework of a longer duration.

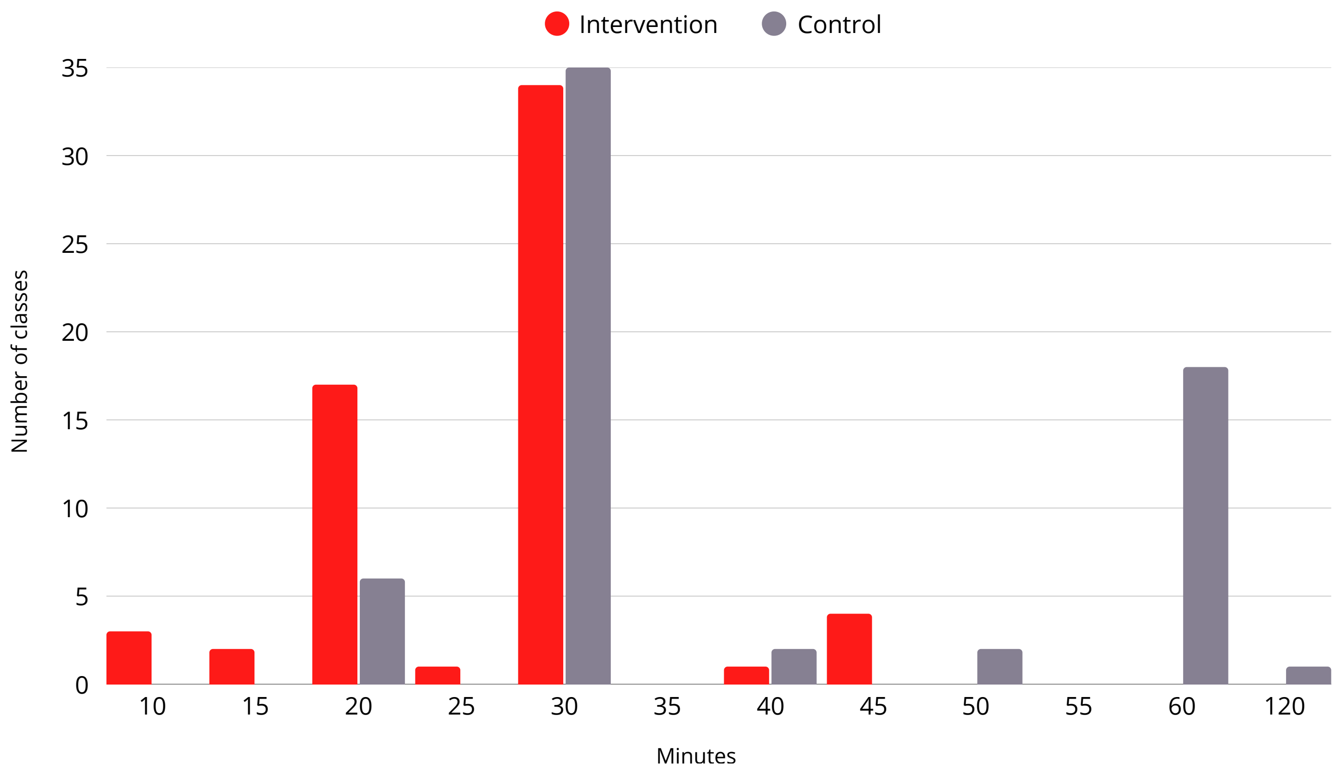

Figure 7. Average duration of weekly maths homework in minutes for control and intervention classes

Maths homework platforms

The most frequently reported platforms used to complete homework amongst intervention group students were Eedi and Mathswatch. Control schools, on the other hand, displayed a broader usage of various online platforms throughout the trial. The most commonly used platforms in control schools were MathsWatch and Sparx, both of which are considered robust tools for delivering maths homework. This suggests that while control schools were not using Eedi, they were employing similar online systems that might serve as close substitutes.

Similarity of Interventions in Control Classes

Control classes were using similar interventions to Eedi, which could be considered close substitutes.

The data shows that 89% of control group teachers used an online system for providing homework, and these platforms often included features that were comparable to those offered by Eedi. For example, MathsWatch and Sparx, the most frequently used platforms in control schools, provide structured homework tasks, interactive elements, and performance tracking, which are key features also found in Eedi.

Moreover, control schools used a wider range of platforms, including MyMaths, Century Tech, and other online tools, which suggests a diversified approach to homework that could mimic the varied and adaptive nature of Eedi. This variety might have reduced the distinction between the intervention and control groups, especially in schools where alternative platforms were particularly robust and engaging.

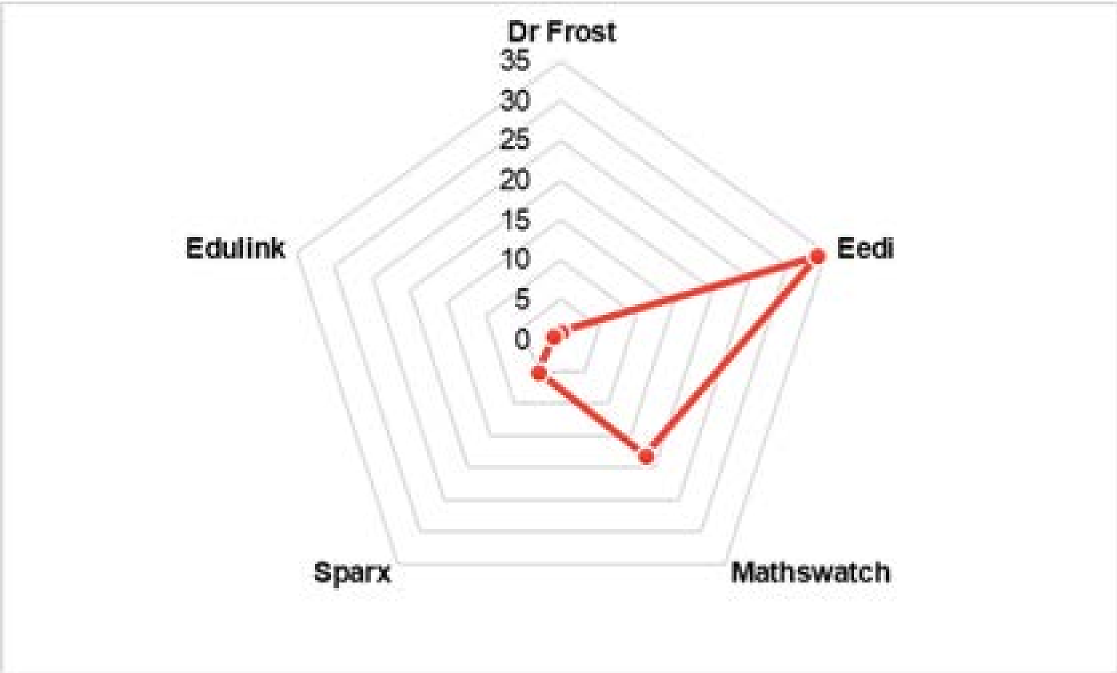

Figure 8. Intervention class homework platforms used during the trial

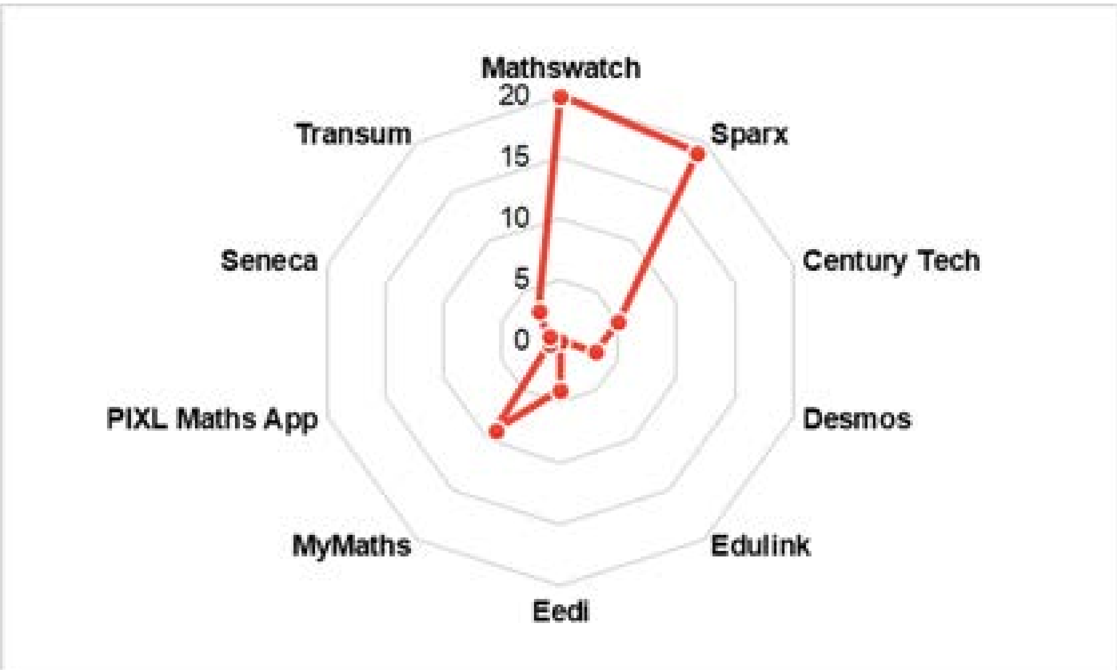

Figure 9. Control class homework platforms used during the trial

How and why does the implementation of the programme vary? To what extent does any variability affect the achievement of expected outcomes?

Variation in implementation of the programme

The implementation of the Eedi programme varied across different schools and classrooms within the intervention group. This variability stemmed from several factors, including differences in teacher engagement with the platform, the extent to which Eedi was integrated into existing teaching practices, and the diverse needs and abilities of students.

One of the primary factors contributing to this variability was the degree of teacher engagement with the Eedi platform. Some teachers fully embraced the platform, integrating it closely with their lesson planning. These teachers actively used the results from Eedi assignments to inform their teaching, as evidenced by one teacher's comment: "I used the misconceptions Eedi uncovered to inform my future planning." However, not all teachers engaged with the platform in the same way.

A significant number of teachers used Eedi more sporadically, or as a supplementary tool, rather than as an integral part of their teaching strategy. In fact, 21 respondents reported that they did not incorporate Eedi's results into their teaching at all. This variation in engagement suggests that the impact of Eedi could differ significantly depending on how deeply teachers integrate the platform into their instructional practices.

The extent to which Eedi was integrated into existing teaching practices also varied among teachers. Some teachers used Eedi as a complement to their regular teaching, particularly for assigning homework. However, others struggled to incorporate it into their already full curriculum schedules. One teacher highlighted this challenge by stating, "It was difficult to integrate Eedi into our existing teaching routines due to time constraints." This variation in how Eedi was used likely influenced its effectiveness in supporting student learning and addressing misconceptions.

Another source of variability in implementation was the diverse needs and abilities of students. Teachers working with lower-ability students often found that the rapid increase in difficulty within Eedi's content posed a significant challenge. One teacher remarked, *"The questions became very hard very quickly, which was tough for lower-ability students."* This mismatch between the platform's content and the students' abilities likely influenced how teachers used Eedi and may have limited its effectiveness in certain classrooms.

Impact of Variability on Expected Outcomes

The variability in the implementation of the Eedi programme likely had an impact on the achievement of expected outcomes. In classrooms where teachers actively engaged with the platform and integrated it into their teaching, Eedi was more likely to help students overcome misconceptions and improve their understanding of mathematical concepts. For example, teachers who used Eedi data to inform their lesson planning were able to tailor their lessons to address specific areas where students struggled, which could lead to better learning outcomes. One teacher's experience illustrates this potential: "Using Eedi helped me identify and address student misconceptions, which I believe improved their understanding."

However, in classrooms where Eedi was not as deeply integrated or where the content did not align well with student needs, the platform's effectiveness may have been diminished. Teachers who faced challenges in incorporating Eedi into their lessons or who found the content too difficult for their students may not have seen the same level of improvement. The variability in how Eedi was used likely contributed to inconsistencies in student outcomes across different classrooms.

Moreover, differences in student engagement with Eedi, influenced by factors such as the difficulty level of the content and technical issues, may have further affected outcomes. In some cases, students were highly engaged and benefited from the platform's interactive elements, while in others, technical glitches or a lack of appropriate support within the content led to frustration and disengagement. One teacher noted, "Students were frustrated by the difficulty of the questions and sometimes by technical issues, which affected their willingness to engage with the platform."

To what extent do teachers link Eedi content to taught lessons? How do teachers link Eedi content to taught lessons?

The degree to which teachers incorporated Eedi content into their taught lessons varies within the intervention group. While some teachers actively integrated Eedi into their lesson planning and teaching practices, others used it more as a supplementary tool rather than a core component of their instructional approach.

A considerable portion of teachers in the intervention group made deliberate efforts to connect Eedi content with their lessons. For example, many teachers utilised the results from Eedi assignments to inform their teaching. Of the 64 respondents, 43 reported that they used Eedi's results to guide their future lesson planning. This approach allowed them to address specific areas where students exhibited misconceptions or difficulties, thereby tailoring their instruction to address these gaps in understanding. By proactively incorporating the insights from Eedi, these teachers could make their lessons more targeted and responsive to student needs.

In contrast, some teachers used Eedi more as a supplementary resource, employing it primarily to reinforce or review content rather than as a central element of their lessons. For instance, 23% of the teachers indicated that they used Eedi to provide whole-class feedback, often by going through problematic questions on the board. While this method does link Eedi content to classroom instruction, it is more of an adjunct to the main teaching objectives rather than a driving force in lesson design.

There are also teachers who made only minimal connections between Eedi content and their taught lessons. Notably, 21 teachers reported that they did not use the results from Eedi assignments to inform their teaching at all. In these cases, Eedi seemed to be treated as a standalone platform rather than an integral part of the instructional process.

In terms of methods, teachers who actively linked Eedi content to their lessons do so in several ways. One common approach is to use Eedi's insights to inform future lesson planning. By analysing data on student misconceptions, teachers adjust their upcoming lessons to focus more on areas where students are struggling, which can enhance understanding and performance.

Another method involves using Eedi for whole-class feedback sessions, where teachers display common errors or challenging questions from Eedi on the board for class discussion. This not only reinforces the content but also clarifies misunderstandings in real-time, effectively linking Eedi content to classroom activities.

Some teachers also used Eedi to identify students who needed additional help, providing targeted support either in small groups or individually. This method ensures that the lessons are personalised based on data from Eedi, aligning the platform's content with the specific needs of individual students. Additionally, a smaller group of teachers used Eedi to spark classroom discussions by sharing examples of student responses, thereby linking Eedi content to broader classroom conversations and promoting critical thinking and peer learning.

Discussion

The analysis reveals that students using the Eedi platform achieved significantly higher NWEA scores compared to those in control schools, even after adjusting for baseline math abilities. Specifically, Eedi produced an effect size of 0.17, indicating that students using the platform scored on average 17% of a standard deviation higher than those using other platforms or none. This result suggests that Eedi was more effective than alternative platforms in supporting student learning outcomes.

The impact of intervention was similar among PP and non-PP students. The data indicated a positive trend in NWEA scores among PP and non-PP students using Eedi. With increasing engagement measured through number of check-in questions answered, the overall improvement increased in both groups..

The analysis of usage patterns demonstrated a notable effect of Eedi on PP students, particularly when usage exceeded 120 check-in questions per year. PP students who engaged with the platform at this level showed significant improvements in their NWEA scores.

The process evaluation demonstrated that a high proportion of teachers used Eedi assignment data to inform their teaching, allowing them to address specific areas where students exhibit misconceptions or difficulties, thereby tailoring their instruction to address these gaps in understanding. By proactively incorporating the insights from Eedi, these teachers were able to make their lessons more targeted and responsive to student needs.

The process evaluation also identified that control schools were using similar interventions to Eedi, which could be considered close substitutes. The data shows that 89% of control group teachers used an online system for providing homework, and these platforms often included features that were comparable to those offered by Eedi, likely reducing the observed effect size, due to the 'active control' nature of the control schools.

Study Limitations and Future Recommendations

One significant limitation of this study was the implementation challenges faced with the MAP Growth Baseline Assessments, which delayed their administration from September 2023 to late October 2023. These issues led to the dropout of 12 schools. As a result, 20 schools remained in the study from the original sample of 32 schools. Schools that remained in the study might have different characteristics compared to those that dropped out, introducing bias. This could affect the internal validity of the study, as the remaining sample may not be fully representative of the intended population.

The technical and administrative burdens placed on schools, mainly managing test durations exceeding single teaching periods, reduced their willingness to participate. An important lesson learned for future effectiveness studies involving Eedi involves selecting a standardised assessment which can be delivered in 30 – 40 minutes, allowing teachers sufficient time to allow learners to log into computers and access the assessment.

The process evaluation highlighted that control classes were using similar online maths interventions to Eedi, therefore the business as usual is technically an active control with students receiving similar support. Consequently, the positive effect size of 0.17 demonstrates that in the sample of students in this study, Eedi was more effective than the interventions used in the control.

A key lesson learned from the evaluation is the importance of ensuring learners engage with the Eedi platform and receive sufficient exposure to the Eedi questions and support. The challenge for Eedi is to develop the system to engage learners in completing a higher number of questions.

Conclusion

Students who used Eedi scored significantly higher than those in control schools on the NWEA mathematics assessment, with an average effect size of 0.17. The intervention effect was similar for both Pupil Premium (PP) and non-PP students. Furthermore, greater engagement with Eedi was associated with larger improvements in NWEA scores across both groups. This positive trend, observed even after adjusting for prior attainment, suggests that Eedi holds promising potential to enhance mathematics learning, including among disadvantaged students.

Attainment

Children using Eedi achieved higher scores in NWEA than children in control schools using other platforms or none.

Pupil Premium

While the data indicated a positive trend in NWEA scores among PP students using Eedi, the difference was not statistically significant.

Misconceptions

A high proportion of teachers used Eedi assignment data to inform their teaching, allowing them to address gaps in understanding.

CAI Platforms

Many control groups used CAIs comparable to Eedi. Despite this, Eedi schools had an effect size of 0.17, higher than control schools.

This study demonstrates that using Eedi has a positive impact on maths attainment, with improvements in pupil premium students' performance increasing with number of check-in questions completed.

References

Ay, Y. (2017). A review of research on the misconceptions in mathematics education. *Education Research Highlights in mathematics, Science and Technology, 12*(1), 21-31.

Barton, C. (2018). On Formative Assessment in maths: How Diagnostic Questions Can Help. *American Educator, 42*(2), 33.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic Assessment With Ordered Multiple-Choice Items. *Educational Assessment, 11*(1), 33–63. https://doi.org/10.1207/s15326977ea1101_2.

Brown & Culora. (2021). Independent analysis of the relationship between Sparx maths and maths outcomes. *RAND Europe*.

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., & Sussex, W. (2007). Effectiveness of reading and mathematics software products: Findings from the first student cohort.

Eedi (2024). Theory of Change documentation.

Harrison, W., Brown, J. & Higgins, S. (2023). A pilot impact study to evaluate the effectiveness of Eedi on raising attainment in mathematics at KS3 (11 years - 14 years old). Eedi. https://cdn.prod.website-files.com/5d11bcacde2cbd6f7d5d8955/649d83bd8b2660e3b5d0d575_Eedi%20Impact%20Pilot%20Evaluation%20May%202023.pdf

Haynes, Kaylea, Paul Fearnhead, and Idris A Eckley. 2017. “A Computationally Efficient Nonparametric Approach for Changepoint Detection.” Statistics and Computing 27: 1293–1305.

Hedges, Larry V. 2007. "Effect Sizes in Cluster-Randomized Designs."

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, 103897.

Holt, L. (2024). The 5 percent problem. Education Next. https://www.educationnext.org/5-percent-problem-online-mathematics-programs-may-benefit-most-kids-who-need-it-least/

Kitchener, K., & O'Donnell, J. (2001). White Rose Numeracy Project. Mathematics in School, 32-34.

Kuo, B.-C., Chen, C.-H., Yang, C.-W., & Mok, M. M. C. (2016). Cognitive diagnostic models for tests with multiple-choice and constructed-response items. Educational Psychology, 36(6), 1115–1133. https://doi.org/10.1080/01443410.2016.1166176.

Laan, Lars van der, Wenbo Zhang, and Peter B Gilbert. 2023. "Nonparametric Estimation of the Causal Effect of a Stochastic Threshold-Based Intervention." Biometrics 79 (2): 1014–28.

Muggeo, Vito MR, David C Atkins, Robert J Gallop, and Sona Dimidjian. 2014. "Segmented Mixed Models with Random Changepoints: A Maximum Likelihood Approach with Application to Treatment for Depression Study." Statistical Modelling 14 (4): 293–313.

Nawaz, S., & Welbourne, S. (2019). A digitally-delivered, double-blind RCT investigating the effect of homework length on progress: Why trial design matters. https://doi.org/10.13140/RG.2.2.23823.74403.

Outhwaite, L., Early, E., Herodotou, C., & Van Herwegen, J. (2023). Can maths Apps Add Value to Learning? A Systematic Review (CEPEO Working Paper Series 23–02). UCL Centre for Education Policy and Equalising Opportunities. https://EconPapers.repec.org/RePEc:ucl:cepeow:23-02.

Schmidt, J. T., & Tang, M. (2020). Digitalization in Education: Challenges, Trends and Transformative Potential. In M. Harwardt, P. F.-J. Niermann, A. M. Schmutte, & A. Steuernagel (Eds.), Führen und Managen in der digitalen Transformation: Trends, Best Practices und Herausforderungen (pp. 287–312). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28670-5_16.

Suppes, P. (1969). Stimulus-response theory of finite automata. *Journal of Mathematical Psychology, 6*(3), 327–355. https://doi.org/10.1016/0022-2496(69)90010-8.

UNESCO. (2023). Global Education Monitoring Report 2023: Technology in education: A tool on whose terms? (1st ed.). GEM Report UNESCO. https://doi.org/10.54676/UZQV8501.

Wang, Z., Lamb, A., Saveliev, E., Cameron, P., Zaykov, J., Hernandez-Lobato, J. M., Turner, R. E., Baraniuk, R. G., Barton, C., & Jones, S. P. (2021). Results and insights from diagnostic questions: The neurips 2020 education challenge. NeurIPS 2020 Competition and Demonstration Track, 191–205.

Wang, Z., Tschiatschek, S., Woodhead, S., Hernández-Lobato, J. M., Jones, S. P., & Zhang, C. (2020). Large-scale educational question analysis with partial variational auto-encoders. arXiv preprint arXiv:2003.05980.

Appendices

Appendix A

NWEA Implementation Issues

Incorrect Firewall Documentation: Eedi initially received a PDF document from NWEA, which they shared with all 32 schools to ensure compatibility with their devices. Subsequently, schools reported difficulties in accessing the tests. After further communication with NWEA, a different set of requirements was provided, requiring all schools to restart the process, causing significant delays. In UK schools, school firewall settings are strict and a priority for the project was to send the documentation to enable the settings to be configured. This can often take time due to priorities of the IT teams in schools at the start of the academic year. The incorrect documentation required all schools to restart this process.

Corrupt File Template: The file template provided for uploading student data did not work on several devices. Despite Eedi's collective efforts, the NWEA team were unable to identify the issue. Eedi had to use their development team to locate and resolve the encoding error, which resulted in considerable time and trial delays.

Browser Testing Discrepancy: The initial agreement included conducting the tests via a browser to simplify the process and reduce technical requirements. However, Eedi discovered that this approach was not feasible, and the NWEA team had to disable the lockdown browser, leading to further setup delays. UK schools are unlikely to download plug-ins and prefer browser based options due to the IT infrastructure in schools.

Lost Data on Suspended Tests: The NWEA team assured Eedi that students could resume suspended tests with their data intact. Regrettably, teachers reported that all data was lost when attempting to continue, causing frustration and disappointment.

24-Hour Delay and Confusion: The 24-hour wait time between test completion and results caused confusion among teachers, who had difficulty identifying which students had completed the assessments. This added complexity to the management of the process.

Delay in GDPR documentation: On 27th September, Eedi requested that the NWEA team fill out a security checklist, and following five follow-up contacts this was still not actioned, resulting in the school dropping out.

Appendix B

Memorandum of Understanding

Memorandum of Understanding (MoU)

Agreement to participate in the Eedi project
This MoU sets out the roles and responsibilities of schools participating in, and the parties involved in delivering and evaluating, the Eedi project. Schools should read this MoU in full, direct any questions to Eedi and sign and return it to Eedi as per the instructions at the end of this document.

The Eedi project is being delivered by Eedi in partnership with WhatWorked Education. Eedi are an education technology company (developers) and WhatWorked are a research evidence company (evaluators). The project is funded by Eedi with the aim of developing an evidence base for the impact of their programme in schools.

Aims of the project
The aim of this project is to evaluate the impact of the Eedi programme on mathematics attainment.

About Eedi
Eedi is an EdTech platform that provides personalised learning and AI-powered independent practice that's as effective as 1:1 tuition. The platform has more than 60,000 diagnostic maths questions and reaches a user base of more than 15,000 monthly users through their school and tutoring platforms. When a misconception is identified, Eedi provides targeted interventions using AI and video explanations made by qualified maths teachers to support learners to address their misconceptions, for free. Eedi resolves 90% of misconceptions reducing teacher's workload and saving valuable time. Eedi's premium model provides learners with an extra layer of support with access to on-demand, one-to-one support from a qualified, UK-based maths teacher via their built-in chat function. The Eedi platform allows teachers to use powerful class analytics to identify remaining misconceptions, so teachers can adjust their teaching to support their learners effectively. In addition to the academic support, the platform includes a coin-based rewards system allowing users to earn points for completing required activities to increase motivation, saving teachers time and effort chasing work.

The evaluation
WhatWorked Education will be the evaluators for the Eedi project.

Who is included
Eedi will recruit schools to start the evaluation to measure the impact over an academic year. The evaluation is conducted as a Randomised Controlled Trial (RCT) with randomisation at the school level. Schools will select KS3 Year 7 to participate in the evaluation. This means that the schools will be randomly chosen to receive Eedi to supplement classroom teaching and the other will not receive Eedi, only classroom teaching. After the evaluation, control schools will be able to provide Eedi in the next academic year. Control schools will be able to use Eedi in KS3 year groups not selected to participate in the evaluation.

Data collection
All data provided to WhatWorked education will be anonymised by Eedi. The evaluators will require the following data:

1. Eedi will collect data about each pupil through the Eedi platform covering their usage of the platform, their performances on quizzes and the pre and post-online assessments.
2. Eedi will collect the following student data from the school: First name, last name, email address, year group, class name, pupil premium status
3. Teachers will be asked to complete a short online survey as part of a process evaluation and this will be anonymised before sharing with the evaluators.

Each school that agrees to participate in the project will agree that one year group will be randomly allocated to the intervention and one to the control. The random assignment will be completed by WhatWorked.

Data protection
All pupil and teacher personal data will be treated with the strictest confidence by Eedi in accordance with the requirements of GDPR 2018. No pupil or teacher personal information will be shared with the evaluators, all data will be anonymised by Eedi prior to sharing with the evaluators. No school, teacher or pupil will be identified in any report arising from this evaluation. The information collected will be used for research purposes only and no information that can identify individuals will be used for any other purpose. Any personal data collected will be destroyed in accordance with the GDPR when this is no longer required.

Head teacher consent
The evaluation is useful for the school to help inform the most appropriate use of Eedi. The intervention is consistent with the kinds of approaches that schools would normally use and no personal data will be shared with the evaluators. As the research is part of usual practice within schools, the head teacher can give consent for the evaluation.

Responsibilities for schools, Eedi and evaluators

Responsibilities for schools, Eedi and the evaluators are set out below:

Responsibilities for schools

1. Identify KS3 Year 7 academic year group to participate in the Eedi project.
2. If required, allow Eedi to provide the training to the teachers involved in the project.
3. Provide the information required by Eedi to create the user accounts for both the intervention and control class.
4. Complete the pre and post online independent assessments (20 minutes in duration) under exam conditions.
5. Agree that Eedi can restrict access to the control year group for the duration of the evaluation. Eedi will then reactivate and grant access to Eedi in the next academic year.
6. The intervention class will complete one Eedi quiz per week for the duration of the academic year.
7. If Eedi usage is directed outside of school time, pupils without the required facilities at home to either complete the quizzes on-site after school or loan pupils access to devices to complete the quizzes.
8. Teachers complete the short online process evaluation survey.

Responsibilities for Eedi

1. Be the first point of contact for the delivery of the programme, via a designated account manager.
2. Provide school training if required to support the school to launch and implement Eedi.
3. Contact the school project lead to assist them setting up the pupil accounts and quizzes on the Eedi platform.
4. Assign access to the pre and post-assessment for the independent assessment.
5. Troubleshoot minor problems.
6. Monitor pupil usage and reach out if they appear to need support.
7. Provide access to their technical support team if required, available Monday to Friday 9am – 6pm.
8. Anonymise all pupil and teacher data prior to sharing with the evaluators.
9. Send the process evaluation survey to participating teachers and anonymise personal identifying data.

Responsibilities for the evaluators

1. Answer any queries about the evaluation.
2. Randomly allocate classes to the control or intervention group.
3. Analyse the anonymised data to measure the impact of Eedi.
4. Publish a report on the findings of the evaluation.

Timeline
Randomisation will be completed in July 2023 and school notified of their assignment to either a control or intervention year group. The independent standardised pre-test will be completed in September 2023 for both the control and intervention. The intervention group will use Eedi for the academic year with both the control and intervention completing the independent post assessment in July 2024.

We commit to the evaluation of Eedi as set out above.

Please electronically sign. Email it to hello@eedi.co.uk**.**

School name

....

Full name

....

Role

....

Signature

....

Date

....

Name of Account Manager, Eedi

....

Signature

....

Date

....

Appendix C

Sensitivity Analysis

Including Subjects Who Answered ≥ 95 Question

We conducted a sensitivity analysis to assess the impact of the intervention among students who answered at least 95 questions. After adjusting for baseline performance, students in the intervention group scored, on average, 1.58 points higher on the post-test compared to those in the control group (95% CI: 0.11 to 3.04; see Table 18). The corresponding effect size was 0.18 (95% CI: 0.02 to 0.34), calculated using the estimated treatment effect and the square root of the total variance from the multilevel model. As shown in Table 19, the intervention group also outperformed the control group with an adjusted percentage difference of 0.73%.

Parameters	Estimate	95% LB	95% UB
Pre-test score (between cluster)	1.11	1.01	1.21
Pre-test score (Within cluster)	0.84	0.81	0.87
Treatment	1.58	0.11	3.04

Table 18: Parameter estimates (Sensitivity Analysis)

<table>
<tr><th>Condition</th><th>Sample Size</th><th>Adjusted Mean</th><th>Raw Standard Deviation</th><th>Effect Size</th><th>% Difference</th></tr>
<tr><td>Treatment</td><td>1114</td><td>218.92</td><td>15.32</td><td rowspan="2">0.18 (0.02 to 0.34)</td><td rowspan="2">0.73%</td></tr>
<tr><td>Control</td><td>1032</td><td>217.34</td><td>15.78</td></tr>
</table>

Table 19: Effect size and percentage difference (Sensitivity Analysis)

Pupil Premium: Interaction Analysis

To examine whether the intervention's effectiveness varied by Pupil Premium status among students who answered at least 95 questions, we included an interaction term in the model. There was some indication that the intervention may have been slightly more effective for non-Pupil Premium students; however, this difference was not statistically significant (Table 20). Subgroup analysis showed that the intervention had a statistically significant effect among non-Pupil Premium students, whereas the effect for Pupil Premium students was not statistically significant. Nevertheless, the direction of the effect for Pupil Premium students was positive (Table 21).

Parameters	Estimate	95% LB	95% UB
Pre-test score (between cluster)	1.10	1	1.20
Pre-test score (within cluster)	0.83	0.8	0.86
Treatment	1.52	0.03	3.01
PP	-1.97	-3.12	-0.82
Treatment TRUE:PP	-0.49	-2.18	1.20

Table 20. Parameter estimates (Interaction Model)

PP Status	Condition	Sample Size	Adjusted Mean	Raw Standard Deviation	Effect Size	% Difference
Non-PP	Treatment	855	221.08	14.79	0.21 (0.05, 0.37)	0.78%
	Control	681	219.36	15,00		
PP	Treatment	247	213.40	15.55	0.09 (-0.014, 0.321)	0.40%
	Control	349	212.56	16.09		

Table 21 Conditional effect size across Pupil-premium groups

Appendix D

School Sample Document

School ID	Progress 8 Score	Attainment 8	% Pupil Premium	% English not First Language	School Type
1	0.05	50.3	18.6	29.4	Secondary
2	-0.85	39.8	53	7.5	Secondary
3	0.09	52.3	30.87	11.9	Secondary
4	-0.29	47.4	49.23	3	Secondary
5	-0.24	52.3	20.11	3.8	Secondary
6	0.19	47.5	47.72	31.9	Secondary
7	N/A	N/A	N/A	N/A	Secondary
8	0.17	52.3	9.46	11.3	Secondary
9	-0.13	43.4	30.06	1.6	Secondary
10	-0.06	45.9	24.07	3.6	Secondary
11	-0.19	52.6	49.34	41.5	Secondary
12	-0.02	50.3	24.8	4.7	Secondary
13	N/A	N/A	N/A	N/A	Secondary
14	0.04	47	52.54	25	Secondary
15	-0.19	44.2	34.76	0.6	Secondary
16	0.41	62.3	15	26.3	Secondary
17	0.69	79.5	6.16	16.7	Secondary

School ID	Progress 8 Score	Attainment 8	% Pupil Premium	% English not First Language	School Type
18	0.26	52.3	16.38	2.2	Secondary
19	N/A	N/A	N/A	N/A	Secondary
20	N/A	N/A	N/A	N/A	Secondary
21	-0.48	40.1	34.81	4.9	Secondary
22	-0.4	38.4	39.81	11.1	Secondary
23	0.19	49.1	12.74	32.8	Secondary
24	-0.83	39.6	30.04	8.8	Secondary
25	-0.16	52.4	24.76	1	Secondary
26	0.15	43.3	56.18	5.1	Secondary
27	0.67	52.1	28.3	58.l3	Secondary
28	N/A	N/A	N/A	N/A	Secondary
29	0.7	54.3	53.08	69.7	Secondary
30	-0.25	39.7	54.63	38.9	Secondary
31	-0.65	44.4	26.72	4.8	Secondary
32	0.05	42.5	24.17	44.6	Secondary

Appendix E

Process Evaluation - Intervention

Thank you for taking the time to complete this survey, your contribution is extremely valuable. This survey is confidential and is being collected for evaluation purposes only. Any identifiable information will be anonymised and this will not be shared beyond the evaluation team. Neither you nor your school will be identified in reports produced in relation to this study. By completing the survey, you consent to your data being used for the evaluation.

Q1 Do you set homework with your class?
Yes / No

Q2 How frequently do you set homework for your class?
We do not set homework
More than twice a week
Twice weekly
Weekly
Fortnightly
Monthly

Q3 Would you say you set homework more often at some times of the year than others or would you say you stuck to this consistently throughout the year?
Homework was set more often for parts of the year but less so others.
I always set homework this often throughout the year.

Q4 Approximately how many minutes would you expect the homework to take most children?
Number

Q5 If you used an online system to set homework, which system did you use?
I did not set homework using an online system
Eedi
Sparx
Mathswatch
Dr Frost
Other please specify

Q6 Does the school, your department or your colleagues tend to use/ or advocates using another online system other than Eedi to support maths progress.
No
Yes, Sparx
Yes, mathswatch
Yes, Dr Frost
Yes, other please specify

Q7 How often did you use Eedi in class time?
I did not use Eedi in class time
More than twice a week
Twice weekly
Weekly
Fortnightly
Monthly

Q8 Did you set any additional homework other than Eedi?
Yes / No

Q9 Can you describe any additional homework you have used to supplement your classroom teaching?
(open ended question)

Q10 In setting up Eedi with your class, did you attend a training session delivered by the Eedi team?
Yes, an Eedi led session (online or in person)
No
Don't know / can't remember

Q11 To what extent do you agree or disagree with the following statements about the implementation of Eedi.
Strongly agree
Agree
Disagree
Strongly disagree
Don't know / NA

Overall, the training session was helpful.
The training session was pitched at the right level for me.
The Eedi platform was easy to set up for my class.
I encountered minimal IT issues when implementing Eedi with my class.
The students in my class had minimal issues accessing the Eedi platform.
The Eedi platform is easy to use to monitor student performance.

Q12 Please use the space below to comment on any of the responses you have given above.
(open ended question)

Q13 How did your homework completion rate on Eedi compare to other homework completion rates?
a) Paper (open response)
b) Online (open response)

Q14 What comments (positive and negative) did your students make about Eedi?
(open ended question)

Q15 What comments (positive and negative) would you like to make about Eedi?
(open ended question)

Q16 What is the main barrier to using Eedi? What did you struggle with the most?
(open ended question)

Q17 How well do you think the Eedi platform supports student learning and engagement?
(open ended question)

Q18 Did you use any of the questions in Eedi or similar versions in your classroom teaching
(open ended question)

Q19 Did you use the present function to display Eedi to your class to discuss any class misconceptions in your lessons?
Yes / No

Q20How did you use the results from your class' Eedi assignments (tick all that apply):
I went through problematic questions with the whole class on the board (whole-class feedback)
I worked with small groups or individual students on specific misconceptions that Eedi uncovered.
I used the misconceptions Eedi uncovered to inform my future planning
I used Eedi to identify which students understood a concept and used them to help other students who were struggling
I shared examples of student responses with the rest of the class to instigate a discussion
I did not use the Eedi results to inform my teaching on this occasion

Q12 Please use the space below to comment on any of the responses you have given above.
(open ended question)

Q13 How did your homework completion rate on Eedi compare to other homework completion rates?
a) Paper (open response)
b) Online (open response)

Q14 What comments (positive and negative) did your students make about Eedi?
(open ended question)

Q15 What comments (positive and negative) would you like to make about Eedi?
(open ended question)

Q16 What is the main barrier to using Eedi? What did you struggle with the most?
(open ended question)

Q17 How well do you think the Eedi platform supports student learning and engagement?
(open ended question)

Q18 Did you use any of the questions in Eedi or similar versions in your classroom teaching
(open ended question)

Q19 Did you use the present function to display Eedi to your class to discuss any class misconceptions in your lessons?
Yes / No

Q20How did you use the results from your class' Eedi assignments (tick all that apply):
I went through problematic questions with the whole class on the board (whole-class feedback)
I worked with small groups or individual students on specific misconceptions that Eedi uncovered.
I used the misconceptions Eedi uncovered to inform my future planning
I used Eedi to identify which students understood a concept and used them to help other students who were struggling
I shared examples of student responses with the rest of the class to instigate a discussion
I did not use the Eedi results to inform my teaching on this occasion

Q21 Do you have any other comments or feedback that you would like to share?
(open ended question)

Process Evaluation - Control

Did you use an online system for providing homework this year?
Yes
No

If So, which was it?
Mathswatch
DrFrost
Sparx
Other- please specify

Also if you used an online system for homework, did you see improvements in children's attainment that you think are due to use of the system?
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree

Approximately, often did you set homework?
More than twice a week
twice a week?
Once a week
Once a fortnight
Less than Once a fortnight

Did you tend to stick to regularly setting homework this often throughout the year?
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree

How long would you expect it would take for most children to complete each homework?
Mins

Do you agree with the statement, I tend to set the homework at a challenging level.
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree

Do you agree with the statement, online homework systems substantially contribute to the progress children make in my class and I would not like to do without them.
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree

Did you use an online system in class time? If So, which was it?
Mathswatch
DrFrost
Sparx
Other- please specify

Typically, how often would children use a system in class time?
More than twice a week
twice a week?
Once a week
Once a fortnight
Less than Once a fortnight

Did you tend to stick to regularly using a system in class time throughout the year?
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree

Typically how long would children be asked to use online systems in class time in each session?
Mins

Does you school generally encourage/advocate the use of the system?
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree

What proportion of other teachers use the same system as you?
%

If you would like please add any other comments about the contribution of online systems to your teaching and the learning children achieve

Contact

Dr Wayne Harrison

Dr Emma Dobson

Prof Steve Higgins

WhatWorked Education
Harrison House
Hawthorn Terrace
Durham
DH1 4EL