UK International Development
Partnership | Progress | Prosperity

I-SAPS Institute of Social and Policy Sciences
Informing Policies, Reforming Practices

AI-Enabled Adaptive Learning

Khyber Pakhtunkhwa

Bridging Technical Assistance for Governments (B-TAG)

AI-Enabled Adaptive Learning

Khyber Pakhtunkhwa

Bridging Technical Assistance for Governments (B-TAG)

Table of Contents

1. Executive Summary

Khyber Pakhtunkhwa continues to face persistent learning gaps among school-age children despite sustained investment in access, infrastructure and teacher development. Conventional classroom instruction, delivered to large and academically heterogeneous groups, struggles to respond to the wide variation in learning levels that exists within a single classroom. To address this challenge, the Bridging Technical Assistance for Governments (B-TAG) programme provided technical assistance to the Elementary and Secondary Education Department (E&SED) and the Directorate of Professional Development (DPD), Government of Khyber Pakhtunkhwa, to design, develop and pilot an Artificial Intelligence (AI)-enabled adaptive learning solution for public schools in the province. The pilot was implemented in 50 public high and higher secondary schools in the districts of Peshawar and Nowshera, and targeted Grade 8 students in English, Mathematics and Science.

The intervention sought to improve Student Learning Outcomes (SLOs) through personalised instruction. Using each student's baseline assessment results, the platform generated adaptive learning pathways tailored to individual learning needs and continuously tracked progress over the implementation period. The pilot combined technology-enabled learning with structured teacher support to ensure effective delivery within existing school systems and timetables.

This report documents the design, implementation and evaluation of the pilot. Key components included the development of curriculum-aligned assessment items linked to prioritised SLOs, deployment of a web-based AI learning platform, and capacity-building support for teachers. Assessment instruments and learning content were reviewed and validated by the Directorate of Curriculum and Teacher Education (DCTE) to ensure alignment with the curriculum in use in Khyber Pakhtunkhwa and with quality standards for grade-level appropriateness.

Participating schools were selected using predefined eligibility criteria that ensured hardware availability and connectivity. Teachers received targeted training on platform utilisation, student registration, assessment administration and strategies for supporting student engagement during scheduled learning sessions. Students completed baseline assessments prior to using the platform, and endline assessments were administered at the conclusion of the intervention period using the same standardised instruments and protocols.

The pilot was evaluated through a quasi-experimental design incorporating treatment and control groups, pre- and post-assessments, and complementary qualitative evidence from implementation monitoring. A total of **2,282 Grade 8 students** were assessed at both baseline and endline, comprising **1,234 students in treatment schools** and **1,048 students in control schools**.

The findings demonstrate that students who participated in the AI-enabled adaptive learning programme achieved consistently greater learning gains than their peers in control schools across all three subjects. Average scores in treatment schools increased by:

- **3.59 marks** in English;
- **3.24 marks** in Mathematics; and
- **4.29 marks** in science.

In comparison, students in control schools recorded gains of **1.64 marks** in English, **1.92 marks** in Mathematics and **1.40 marks** in science.

Difference-in-Differences (DiD) analysis indicates that, after accounting for improvements observed in control schools, the intervention was associated with additional gains of **1.95 marks in English**, **1.32 marks in Mathematics** and **2.88 marks in Science**. Expressed in percentage terms, treatment group students outperformed control group trends by **10.2 percentage points in English**, **7.7 percentage points in Mathematics** and **13.4 percentage points in Science**. Formal DiD regression models confirm that these effects are statistically significant in all three subjects, with the strongest evidence of programme impact observed in Science.

Disaggregated analysis by gender points to an important pattern. Female students, who constituted approximately **79 percent of the assessed sample**, recorded substantial gains across all three subjects, with average increases of 4.6 marks in English, 3.6 marks in Mathematics and 5.1 marks in Science within treatment schools. Male students in treatment schools, by contrast, showed minimal change over the same period. This divergence is discussed in the results and limitations sections and warrants further investigation before firm conclusions are drawn.

Beyond learning gains, the pilot demonstrated the feasibility of integrating AI-enabled adaptive learning technologies within the public education system of Khyber Pakhtunkhwa. The intervention highlighted the potential of personalised learning pathways, real-time progress monitoring and data-driven instructional support to address diverse learning needs within existing school structures.

The report concludes with key lessons emerging from implementation and provides recommendations for strengthening, refining and scaling AI-enabled adaptive learning solutions within the province's education system. These recommendations focus on improving implementation effectiveness, enhancing teacher support mechanisms, strengthening technology infrastructure, and leveraging assessment and learning data to inform future education reforms.

Scope and Reach

Parameter	Treatment group	Control group
Districts	Peshawar and Nowshera	Peshawar and Nowshera
Schools	50 (35 female, 15 male)	50 (35 female, 15 male)
Grade 8 students assessed	1,234 (934 girls, 300 boys)	1,048 (872 girls, 176 boys)
Teachers trained	44 (31 female, 13 male)	Not applicable
Subjects covered	English, Mathematics, Science	English, Mathematics, Science

Table 1: Scope and Reach of the Intervention

2. Introduction

2.1 Programme Background and Rationale

Despite significant progress in expanding access to education, public education systems in Pakistan continue to face persistent challenges in improving student learning outcomes. Evidence from national assessments and large-scale learning data consistently indicates that many students progress through the education system without acquiring the expected competencies in core subjects. These learning gaps often reflect cumulative deficits from earlier grades, variations in instructional quality, overcrowded classrooms, and limited opportunities for teachers to provide differentiated support to students with diverse learning needs.

Khyber Pakhtunkhwa presents a particularly instructive setting for testing technology-supported approaches to this problem. The province combines dense urban school networks in districts such as Peshawar with more dispersed provision in adjoining districts such as Nowshera, and digital infrastructure is distributed unevenly across the system. A subset of high and higher secondary schools maintains functional computer laboratories and basic connectivity, creating a realistic entry point for digital learning without requiring new capital investment.

The selection of Grade 8 is strategic from several standpoints. As the final year of elementary education, Grade 8 represents a critical transition point before students enter secondary education and begin studying more advanced subject content. Learning gaps that remain unaddressed at this stage often become increasingly difficult to remediate and may adversely affect students' performance in secondary school and subsequent board examinations. Grade 8 students are also likely to possess the digital skills necessary to engage independently with technology-enabled learning platforms, making them well suited to pilot adaptive learning approaches. From an implementation perspective, many high and higher secondary schools serving Grade 8 already have functional computer laboratories and basic internet connectivity, providing the minimum infrastructure required for digital learning.

However, classroom-level assessment practices often provide limited insight into individual learning needs, making it difficult for teachers to identify specific learning gaps and provide targeted support. Consequently, substantial variation in student learning levels within the same classroom frequently remains unaddressed, contributing to widening achievement gaps over time.

Globally, digital learning technologies and artificial intelligence have emerged as promising tools for supporting personalised learning and improving educational outcomes. AI-enabled learning systems can utilise diagnostic assessments, adaptive content delivery and real-time performance data to tailor learning experiences to individual students. When integrated within formal schooling systems, such technologies can complement teacher-led instruction by providing actionable insights into student learning and enabling more targeted and responsive teaching practices. Recognising this potential, the B-TAG programme has been exploring the use of AI-enabled solutions to strengthen teaching and learning within public education systems in Pakistan.

Building on earlier experience with AI-supported learning platforms in the public sector, B-TAG, in collaboration with the E&SED and the DPD, designed and piloted an AI-enabled adaptive learning intervention for public schools in Khyber Pakhtunkhwa. The intervention was designed to provide students with personalised learning pathways directly through an adaptive digital platform. Based on students' performance in baseline assessments, the platform automatically adjusted learning content and activities to match individual learning needs while continuously tracking progress over time. The focus on Grade 8 reflected both the educational significance of this transition year and the practical feasibility of implementing technology-supported learning in schools with access to computer laboratories and basic digital infrastructure.

The rationale for the intervention was grounded in three interconnected considerations.

i) There was a need for more granular diagnostic information to identify individual learning gaps among students approaching the transition to secondary education.

ii) Public schools required scalable approaches for providing differentiated learning support within classrooms characterised by diverse learning levels.

iii) There was a need to generate provincially relevant evidence on the feasibility, effectiveness and operational requirements of implementing AI-enabled adaptive learning solutions within government school systems in Khyber Pakhtunkhwa.

This report documents the design, implementation and evaluation of the AI-enabled adaptive learning pilot. It presents key implementation processes, including intervention design, school selection, assessment and content development, teacher orientation and training, platform deployment, and the administration of baseline and endline assessments. The report also examines programme outcomes and implementation experiences to inform future efforts aimed at integrating adaptive learning technologies within public education systems in Khyber Pakhtunkhwa and beyond.

3. Methodology

3.1 Overall Evaluation and Documentation Approach

The evaluation framework adopts a quasi-experimental design incorporating treatment and control groups, with student learning measured through standardised assessments administered at baseline and endline. This design allows for comparison of changes over time between participating and non-participating schools. Quantitative assessment data are complemented by qualitative inputs gathered during implementation monitoring, collected to contextualise observed patterns and to document experiences related to platform use, assessment administration and classroom facilitation.

3.2 Study Design

The study design is based on a quasi-experimental framework intended to assess changes in student learning outcomes associated with exposure to the AI-enabled adaptive learning platform. The design incorporates treatment and control groups and applies a pre- and post-intervention comparison to support attribution while remaining compatible with operational and administrative constraints in public sector school settings.

Participating schools were assigned either to a treatment group, which implemented the intervention, or to a control group, which continued with standard instructional practices during the study period. Both groups were assessed at two points in time using the same standardised assessment instruments: a baseline assessment administered before the start of the intervention and an endline assessment administered following completion of the implementation period. This structure allows for the measurement of changes in learning outcomes over time within each group and the comparison of those changes between groups.

The primary analytical approach underpinning the study design is the Difference-in-Differences (DiD) framework. This approach estimates the intervention effect by comparing the change in outcomes over time in treatment schools relative to the change observed in control schools. By focusing on differences in changes rather than absolute levels, the design seeks to account for baseline differences between groups and for time-related factors that may affect all schools simultaneously. The validity of this approach rests on the assumption that, in the absence of the intervention, learning trajectories in treatment and control schools would have moved in parallel.

3.3 Sampling Strategy

The sampling approach was designed to ensure adequate representation across districts, school gender and school categories, while remaining consistent with the operational requirements of the intervention and the planned rollout arrangements. Given that the intervention relied on a web-based AI learning platform and computer-based assessments, the **school** was adopted as the primary unit of assignment.

The study sample was restricted to the districts of **Peshawar and Nowshera**, which were identified as the designated intervention districts for the pilot. The sampling frame comprised only **public high and higher secondary schools with functional computer laboratories**, as these schools possessed the minimum infrastructure necessary to administer digital assessments and facilitate platform-based learning activities. Schools without operational computer laboratories were excluded from the sampling frame because they could not adequately support the technical requirements of the intervention.

To identify the study sample, schools were selected using a combination of census and random sampling approaches. All **55 female schools in Peshawar** with functional computer laboratories were included in the study, reflecting the limited number of eligible female schools in the district. In addition, **30 male schools in Peshawar** and **15 female schools in Nowshera** were selected through random sampling from the pool of eligible schools. This approach ensured that the final sample adequately reflected the available distribution of eligible schools while maintaining operational feasibility.

Figure 1 presents the sampling approach adopted for the study.

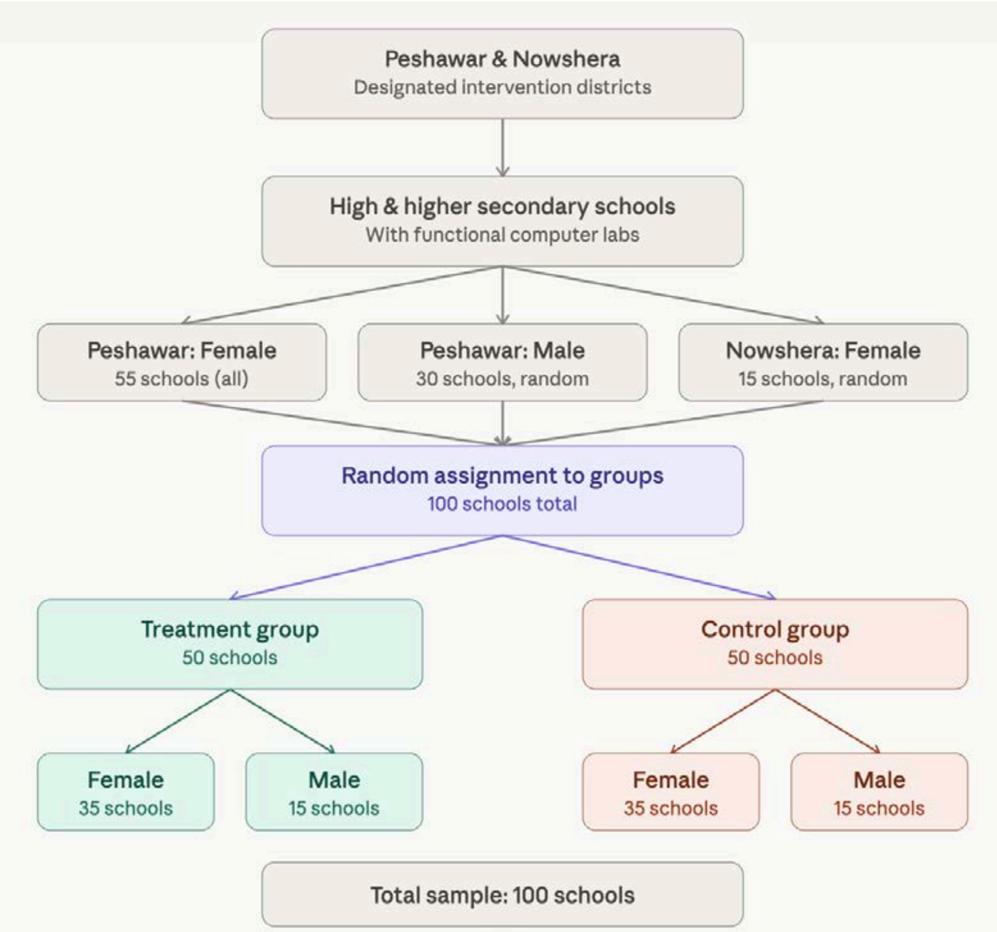

Figure 1: Sampling Approach

The identified sample of 100 schools was subsequently assigned randomly to either treatment or control groups. Random assignment at the school level was intended to minimise selection bias and facilitate credible comparison between intervention and non-intervention schools within the quasi-experimental evaluation design.

The final sample comprised **100 public high and higher secondary schools**, evenly divided between treatment and control groups. Of these:

- **50 schools** were assigned to the treatment group;
- **50 schools** were assigned to the control group; and

- each group comprised **35 female schools (70 percent)** and **15 male schools (30 percent)**.

The composition of the final sample therefore maintained an identical gender distribution across both study arms, ensuring balance between treatment and control groups. The composition of the final sample is presented in **Figure 2**.

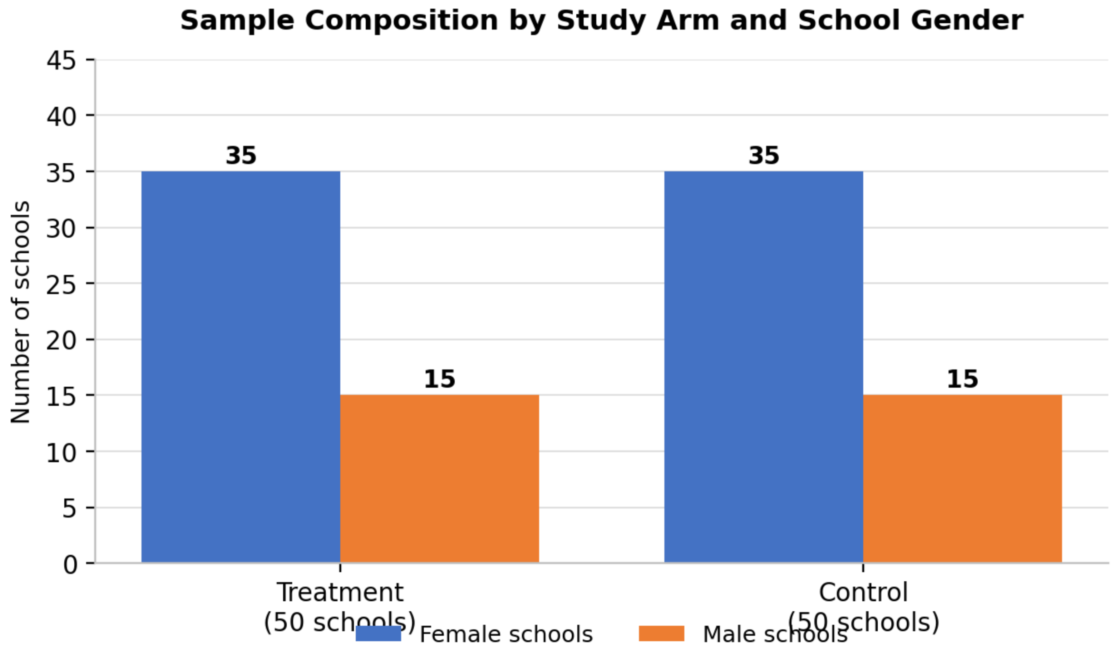

Figure 2: Sample Composition by Study Arm and School Gender

At the student level, a total of **2,282 Grade 8 students** were assessed at both baseline and endline: **1,234 students in treatment schools** (934 girls and 300 boys) and **1,048 students in control schools** (872 girls and 176 boys). Reflecting the composition of the eligible school population, female students accounted for **1,806 students, or approximately 79 percent of the assessed sample**. This distribution is important for the interpretation of aggregate results and is discussed further in Sections 6.7 and 7.

3.4 Data Sources

The primary quantitative data source comprises standardised student learning assessments administered at baseline and end line. Assessment instruments were developed for Grade 8 English, Mathematics and Science and aligned with approved curricula and SLOs. Each subject was assessed out of a total of 50 marks. The same instruments were used at both time points to ensure comparability over time. Assessments were administered through the AI-enabled platform under standardized protocols, with each student accessing the assessment using a unique identifier.

Platform-generated administrative data constitute a second key data source. These data include records of student logins, assessment completion, interaction with learning content and system usage patterns during the implementation period. Platform data were used to document exposure to the intervention and to support monitoring of implementation fidelity, rather than to generate independent outcome measures.

Qualitative inputs collected during monitoring visits were used to capture perceptions of platform usability, assessment administration processes and practical challenges encountered during delivery.

These inputs inform the interpretation of quantitative findings and the lessons presented in Sections 7 and 8.

4. Literature and Evidence Review

4.1 Overview of AI-Enabled Learning in Public Education

AI-enabled learning has been increasingly explored within public education systems as a means of addressing persistent challenges related to heterogeneous learning levels, limited instructional time and constraints on teacher capacity. In policy and programme contexts, such approaches are typically framed as tools to support diagnostic assessment, personalised practice and data-informed instruction, rather than as substitutes for classroom teaching.

Within public sector settings, AI-enabled learning systems commonly integrate three functional elements: structured assessment to identify learning gaps, adaptive sequencing of content aligned with curricular standards, and monitoring mechanisms that generate actionable information for teachers and administrators. These systems rely on predefined learning objectives and rule-based or algorithmic decision processes to adjust content difficulty and pacing in response to student performance. The emphasis is generally on formative support and remediation, particularly in contexts, where large class sizes and diverse learner profiles limit opportunities for individualized attention.

Evidence from prior programmes and implementation reports indicates that AI-enabled tools are most often deployed as complementary interventions within existing school structures. Their effectiveness and feasibility are closely linked to contextual factors, including infrastructure availability, alignment with national and provincial curricula, teacher preparedness and administrative support. In public schools, where digital access and technical capacity vary widely, AI-enabled interventions have typically been implemented in settings with minimum thresholds of connectivity and device availability to reduce operational risk.

From an institutional perspective, the use of AI-enabled learning in public education has also been associated with efforts to strengthen monitoring and evaluation. Platform-generated data can provide timely information on student participation, assessment completion and engagement with learning materials, supporting more systematic documentation of implementation processes. However, the interpretability and use of such data depend on clearly defined assessment frameworks and standardized administration protocols.

Overall, the literature and programme documentation suggest that AI-enabled learning in public education is best understood as an enabling mechanism within broader instructional and system-level reforms. Its contribution lies primarily in supporting more granular assessment, structured differentiation and improved documentation of learning processes, subject to the constraints and capacities of public sector delivery environments.

4.2 Evidence from Comparable Interventions

Evidence from comparable education interventions indicates that technology-supported and AI-enabled learning approaches have been used across a range of public sector contexts to address learning gaps, particularly where traditional instructional models face capacity constraints. These interventions vary in design and scope but commonly combine structured assessments, adaptive or guided practice, and teacher-facing data to support instructional decision-making.

Documentation from prior public sector pilots and donor-supported programmes suggests that learning technologies tend to show more consistent effects when they are tightly aligned with national curricula and embedded within existing school routines. Interventions that integrate assessment and learning content within a single platform have been used to support formative feedback and targeted practice, especially in environments characterised by large class sizes and heterogeneous learning levels. Reported outcomes from such programmes often highlight improvements in diagnostic visibility and instructional focus, rather than uniform gains across all learning domains.

Comparable interventions have also underscored the importance of implementation conditions in shaping observed outcomes. Programme records frequently note that teacher training, clarity of facilitation roles and administrative support at the school level are critical to consistent platform use. Where these conditions are weak or uneven, reported benefits tend to be limited or highly variable across schools. As a result, many evaluations place emphasis on documenting implementation fidelity alongside learning outcomes.

In low- and middle-income country settings, evidence from pilots employing adaptive or data-driven learning tools points to mixed but instructive results. Some interventions report positive associations between platform use and student learning gains, while others identify modest or uneven effects, often attributed to infrastructure limitations, disruptions to instructional time, or challenges in sustaining teacher engagement. These findings highlight the need for cautious interpretation and for situating results within documented delivery contexts.

Overall, evidence from comparable interventions suggests that AI-enabled learning tools can contribute to improved instructional targeting and assessment practices when implemented within well-defined parameters. At the same time, the literature emphasises that such tools are not self-executing solutions and that their performance depends on alignment with curriculum, capacity-building for implementers, and realistic expectations regarding scale and scope within public education systems.

4.3 Relevance to the Pakistani Public-School Context

The relevance of AI-enabled learning approaches to the Pakistani public-school context is shaped by a combination of systemic constraints and institutional priorities. Public schools operate within environments characterised by large class sizes, variability in student preparedness and limited capacity for individualised instruction. At the upper elementary level, these challenges are compounded by increased subject complexity and greater expectations for student independence in learning.

Within this context, assessment practices in public schools are often oriented toward summative evaluation, with limited mechanisms for systematically diagnosing learning gaps during the academic year. As a result, instructional responses to student underperformance are frequently generalised rather than targeted. AI-enabled platforms that integrate diagnostic assessment and structured learning pathways align with identified needs for more granular information on student learning without requiring substantial changes to existing curricula or teaching roles.

Infrastructure and access considerations are central to determining relevance and feasibility. While digital access is uneven across the public school system, a subset of schools, particularly at the high and higher secondary levels, have functional computer laboratories and basic connectivity. In Khyber

Pakhtunkhwa, this subset is concentrated in and around district headquarters and larger settlements, which shaped the district selection for this pilot. Interventions designed to operate within these constraints, using shared devices and standard web browsers, are more readily applicable to the Pakistani public-school context than models that require high-end hardware or continuous connectivity.

Teacher capacity and workload are additional considerations. Public school teachers often manage large classes with limited time for individual follow-up. AI-enabled tools that provide structured assessment workflows and curated learning resources may support teachers by reducing the cognitive and administrative burden associated with diagnosing and addressing diverse learning needs. However, evidence from prior programmes also indicates that such tools require clear facilitation protocols and targeted training to be used consistently and effectively.

From a policy and administrative perspective, relevance is also linked to alignment with approved curricula, Student Learning Outcomes and institutional mandates. Interventions that are perceived as parallel or external to the formal system face greater barriers to adoption and sustainability. In contrast, AI-enabled learning models that are framed as system-support tools and implemented in coordination with government counterparts are more consistent with prevailing governance arrangements in Pakistan.

Overall, the applicability of AI-enabled learning in Pakistani public schools lies in its potential to support existing instructional processes through structured assessment and data-informed facilitation, within clearly defined operational boundaries. Its relevance is therefore contingent on realistic alignment with infrastructure capacity, teacher roles and institutional frameworks rather than on assumptions of system-wide digital transformation.

4.4 Implications for Programme Design

The evidence and contextual considerations outlined in the preceding sections have direct implications for the design of AI-enabled learning interventions within public sector education systems. Programme documentation and comparable experiences suggest that design choices must prioritize alignment with institutional structures, operational feasibility and clarity of roles over technological sophistication alone.

First, alignment with approved curricula and Student Learning Outcomes is a foundational design requirement. Interventions that embed assessment and learning content within existing curricular frameworks are more readily integrated into school routines and are less likely to be perceived as parallel or additional burdens. This alignment also supports comparability of assessment data and facilitates interpretation by government stakeholders.

Second, programme design must account for infrastructure constraints and variability across schools. Defining minimum hardware, connectivity and access requirements in advance helps manage implementation risk and supports more consistent delivery. Designs that rely on web-based platforms accessible through standard browsers and shared devices are better suited to public school environments than models requiring specialized equipment or continuous high-bandwidth connectivity.

Third, the role of teachers should be clearly defined within the intervention logic. Evidence from comparable programmes indicates that AI-enabled tools are most effective when positioned as facilitative supports rather than replacements for instruction. Design features should therefore emphasize usability, structured workflows and clarity in how teachers are expected to engage with assessments, learning resources and platform outputs.

Fourth, training and ongoing support mechanisms are integral to programme design. Short-term orientation is often insufficient for sustained and consistent use of digital tools. Design considerations should include structured capacity-building activities, practical guidance on classroom facilitation and accessible support channels to address technical or procedural issues during implementation.

Finally, implications for evaluation and documentation should be embedded at the design stage. Clearly defined assessment protocols, data collection processes and documentation tools enable systematic monitoring of both outcomes and implementation fidelity. This supports proportionate interpretation of results and provides an evidence base for institutional learning without extending claims beyond what the data can support.

5. Implementation Design and Processes

Implementation of the pilot proceeded through four sequenced phases: baseline assessment, teacher orientation and training, platform-based learning in treatment schools, and endline assessment. The implementation timeline is summarised in **Table 2**.

Activity	Period	Coverage
Baseline assessment	18 – 31 January 2026	Treatment and control schools, Peshawar and Nowshera
Teacher orientation and training	27 February 2026	44 IT teachers from 50 treatment schools, RPDC (Female) Peshawar
Platform-based learning facilitation	March – April 2026	Treatment schools only
Endline assessment	13 April – 5 May 2026	Treatment and control schools, Peshawar and Nowshera

Table 2: Implementation Timeline

Baseline assessments were administered ahead of teacher training in order to generate the diagnostic profiles required to configure individual learning pathways. Teacher orientation was completed before treatment schools commenced platform-based learning, so that facilitation arrangements and student credentials were in place at the point at which students first accessed the platform.

5.1 Adaptation of the AI-Enabled Learning Platform

B-TAG supported the implementation of an AI-enabled learning platform for Grade 8 students in Khyber Pakhtunkhwa. The platform was customised using curriculum-aligned instructional content for English, Mathematics and Science across Grades 6 to 8 to provide targeted remedial support based on individual student learning needs. Using students' baseline assessment results, the platform generated adaptive learning pathways tailored to each learner and continuously tracked progress throughout the intervention.

The platform incorporated several features to enhance accessibility and student engagement, including bilingual support, voice-enabled navigation and a user-friendly interface. Students could access the platform using unique login credentials, enabling personalised learning experiences and continuous monitoring of their progress. The platform could be accessed both through school computer laboratories and remotely using mobile devices, allowing students to continue learning beyond the classroom where feasible.

As a web-based application, the platform required no local software installation and could be accessed through commonly available internet browsers. Browser settings enabling audio input, cookies and local storage were required to support authentication and maintain user sessions. This approach minimised technical complexity and reduced dependence on specialised IT support within schools. Reliable internet connectivity was essential for accessing the platform and its learning resources.

Separate access privileges were established for teachers and administrators responsible for student registration, platform administration and implementation oversight. Teachers and authorised users could enrol students, monitor usage and learning progress, and access implementation reports using standard computing devices.

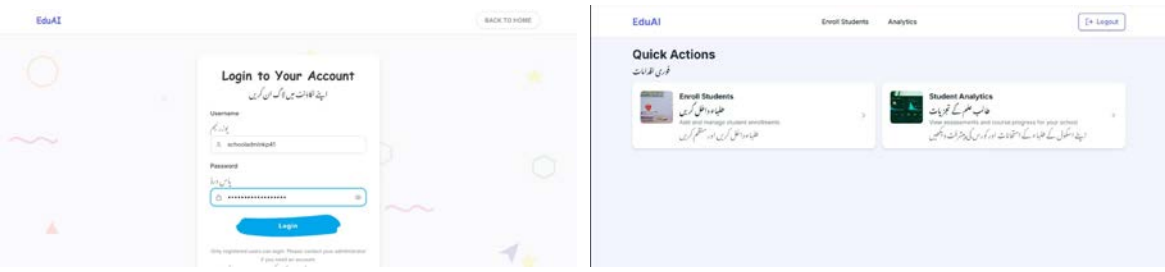

Figure 3: Platform screenshots – learning interface and adaptive content delivery

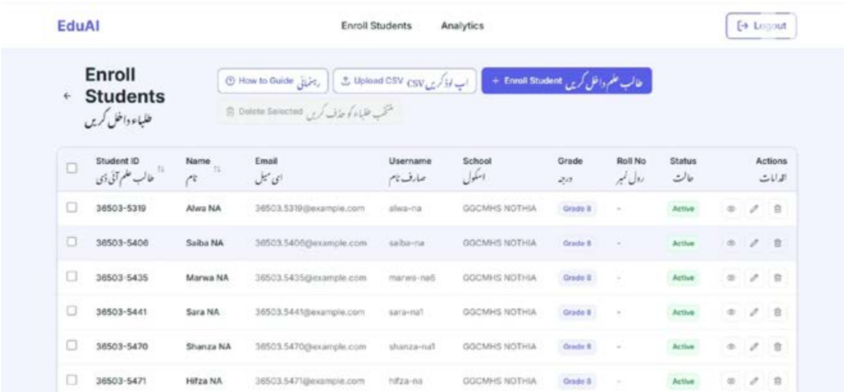

Figure 4: Platform screenshots – student learning pathway

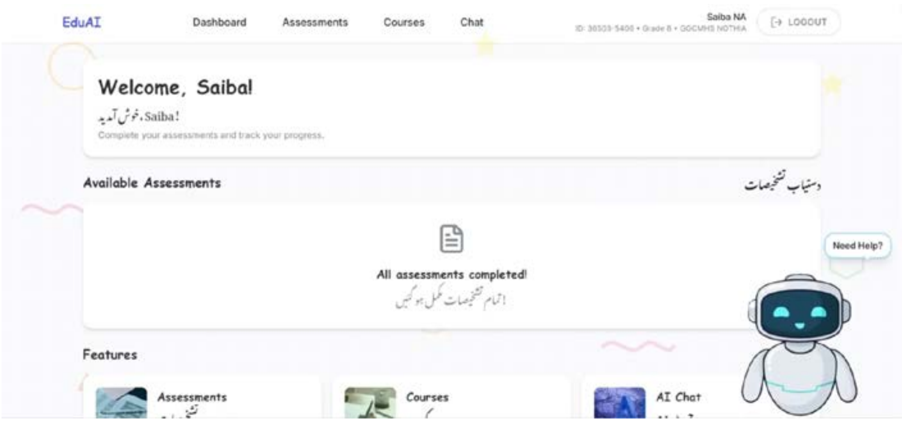

Figure 5: Platform screenshots – learning checks and progress tracking

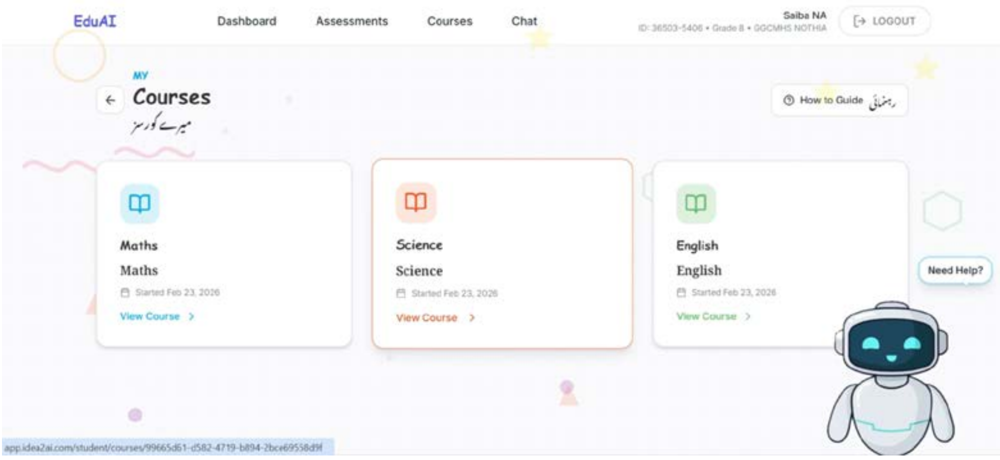

Figure 6: Platform screenshots – interactive chatbot support

5.2 Content Development and Review

To support the effective implementation of the platform, baseline and end line assessment items were developed for prioritised SLOs across Grades 6 to 8 in English, Mathematics and Science. Alongside the assessments, the platform integrated curated learning resources aligned with the identified SLOs to support personalised remediation. These resources included existing textbooks, teacher guides and carefully selected open-source materials, all mapped to the relevant learning outcomes. Following the diagnostic assessment, the platform utilised these resources to generate structured, adaptive learning pathways tailored to individual student needs, ensuring that targeted remedial support was available across the selected grades and subjects.

A comprehensive quality assurance process was undertaken to verify the accuracy, quality and curricular alignment of both the assessment items and the learning resources. Assessment and curriculum specialists reviewed the materials to ensure their alignment with approved SLOs, curriculum standards, grade-level appropriateness and technical quality. Feedback from the review was systematically incorporated into subsequent revisions, ensuring that the content met quality standards and was suitable for implementation in public sector classrooms in Khyber Pakhtunkhwa.

5.3 Teacher Training

Following the development of student learning content, including curated instructional resources for Grades 6 to 8 and baseline and endline assessment items, a series of instructional videos was developed to support implementation. The videos introduced the objectives and design of the AI-enabled learning intervention, demonstrated platform navigation for teachers and headteachers, and provided guidance on student registration, platform usage, assessment administration and classroom facilitation.

A blended training programme was designed to ensure that participating teachers were equipped to implement the intervention effectively. Training covered platform navigation, student registration and login procedures, administration of baseline and endline assessments, facilitation of scheduled computer laboratory sessions, and monitoring of student progress. Teachers were also oriented to the purpose of the intervention, the appropriate use of AI-enabled learning within the school timetable, and the distinction between platform-supported learning activities and routine classroom instruction.

The orientation and training session was conducted on **27 February 2026** at the Regional Professional Development Centre (Female), Peshawar. A total of **44 IT teachers (31 female and 13 male)** from the 50 participating treatment schools were trained to support the implementation of the AI-enabled adaptive learning platform. During the session, detailed orientation was provided on the platform, and school administrator credentials together with individual student login identifiers were shared with the participating IT teachers. The Director, Directorate of Professional Development, also attended the orientation session, signalling institutional support for the intervention.

Figure 7: Teacher orientation and training session, RPDC (Female), Peshawar – 27 February 2026

5.4 Conduct of Pre-Assessment

Pre-assessments were conducted between **18 and 31 January 2026** in both treatment and control schools across Peshawar and Nowshera, to establish baseline measures of student learning in English, Mathematics and Science and to generate diagnostic information to inform subsequent platform-based learning pathways.

Assessments were administered through school computer systems using each student's unique user identifier. Facilitators were trained on the use of standardised protocols and monitoring tools to ensure assessment integrity, appropriate testing conditions and readiness of required hardware and software. They facilitated the assessment sessions, supporting students with basic navigation where required while adhering to guidelines intended to preserve assessment integrity. Where necessary, assessments were conducted in multiple sessions to accommodate shared device access and to ensure that all participating students had an opportunity to participate.

A total of **2,282 Grade 8 students** completed the baseline assessment across the 100 sampled schools, forming the analytical sample used for the evaluation reported in Section 6.

5.5 Platform-Based Learning Facilitation

Following the completion of baseline assessments and teacher orientation, students in treatment schools engaged with the AI-enabled adaptive learning platform during scheduled learning sessions facilitated by teachers, primarily in school computer laboratories. Teachers first oriented students on the use of the platform and distributed individual login credentials, after which each student accessed the platform using their own identifier.

Based on each student's baseline assessment results, the platform generated a personalised learning pathway and delivered subject- and topic-specific learning resources in English, Mathematics and Science. Learning content was sequenced according to individual learning needs and aligned with relevant SLOs, ensuring that students received targeted remedial support. Students were required to work through the learning deficiencies identified for them, and progress was verified through embedded learning checks at defined points along each pathway. Where students required further explanation of a concept, they could interact with the platform's chatbot for additional support. Student interaction with the platform was primarily text-based, with voice-enabled features available where technical conditions permitted.

Teachers and designated school staff supported implementation by facilitating student access to the platform, monitoring participation and assisting students as required during learning sessions. Using their own login credentials, teachers were able to monitor individual student progress and engagement with the platform, enabling them to provide timely support where needed.

In schools with a limited number of functional computers, students accessed the platform through shared devices using staggered schedules or small-group sessions, ensuring that all participating students had opportunities to engage with the learning activities. Using their unique login credentials, students were also able to access the platform remotely through mobile phones and laptops, enabling continued learning beyond the classroom where feasible.

Throughout the implementation period, the platform captured data on student participation, learning progress, engagement with instructional content and completion of assigned activities. These data supported routine implementation monitoring and provided evidence on student learning patterns and platform utilisation.

District education authorities in Peshawar and Nowshera monitored the implementation of the intervention to ensure smooth execution and to address operational issues as they arose. The B-TAG team also conducted periodic monitoring visits to participating schools, providing technical and implementation support where required.

Monitoring visits indicated a generally high level of engagement among teachers and students with the AI-enabled learning platform. While implementation was largely successful, schools reported challenges related to the limited availability of functional computers and inconsistent internet connectivity. These constraints were more pronounced in some schools than others and are relevant to the interpretation of the differential results reported in Section 6.7.

5.6 Conduct of Post-Assessment

Post-assessment was conducted between **13 April and 5 May 2026** in both treatment and control schools in Peshawar and Nowshera, using the same standardised assessment instruments administered at baseline. Administration protocols applied during the pre-assessment phase were retained to support comparability and minimise procedural variation. Assessment conduct was facilitated by trained facilitators, and where necessary assessments were conducted in multiple sessions to accommodate shared device access.

Completion of post-assessments concluded the field implementation phase of the intervention. Endline data generated through this process form the basis for the analysis presented in Section 6 under the approved evaluation framework.

6. Results and Analysis

The baseline and endline assessments administered to Grade 8 students in English, Mathematics and Science were compared to evaluate the effect of the intervention. The results are presented progressively, beginning with overall trends in student performance, followed by comparisons between treatment and control groups, estimates of the intervention effect using a DiD approach, and disaggregated analysis by gender. Formal DiD regression results are presented at the end of the section to assess the statistical significance of observed changes.

6.1 Overall Performance Across Subjects

Figure 4 presents average student scores across all participating schools, combining treatment and control groups. Improvements were observed across all three subjects between baseline and endline assessments, indicating overall gains in student learning during the intervention period.

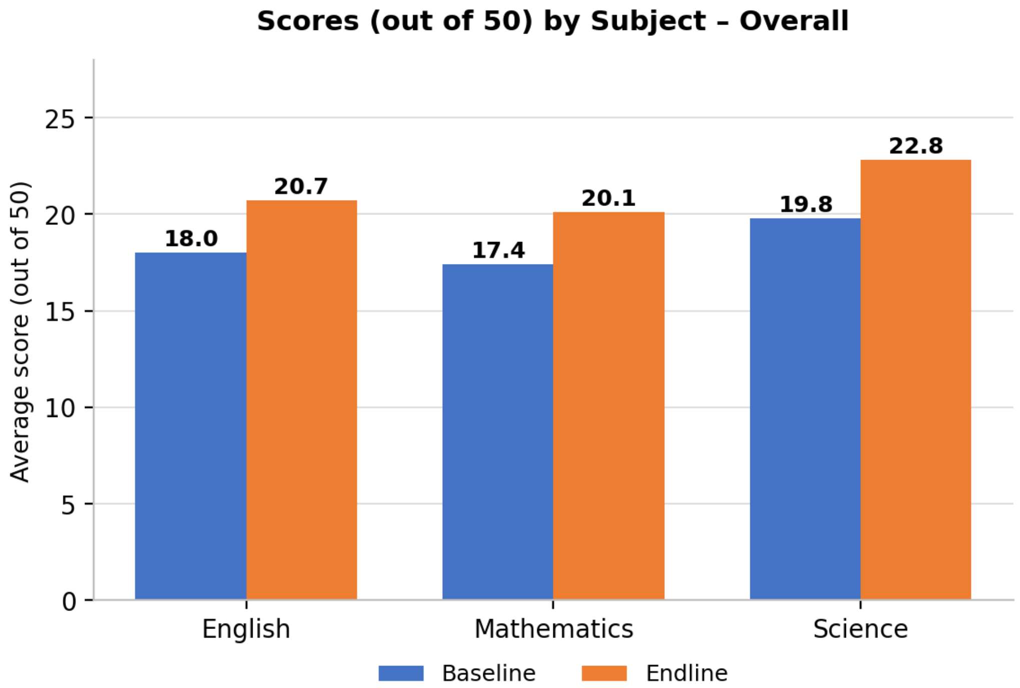

Figure 8: Scores (out of 50) by Subject – Overall

Average English scores increased from **18.0 marks** at baseline to **20.7 marks** at end line, representing an improvement of **2.7 marks**. Mathematics scores increased from **17.4 marks** to **20.1 marks**, also reflecting a gain of **2.7 marks**. The largest overall improvement was observed in science, where average scores increased from **19.8 marks** at baseline to **22.8 marks** at end line, an increase of **3.0 marks**.

The results suggest positive learning gains across all subjects over the implementation period. However, overall improvements do not distinguish between changes attributable to the intervention and those arising from other factors affecting all students. Subsequent sections therefore examine differences between treatment and control groups to better understand the contribution of the AI-enabled adaptive learning intervention.

6.2 Performance by Treatment and Control Groups

Figure 5 compares average scores by subject and study group at both baseline and end line.

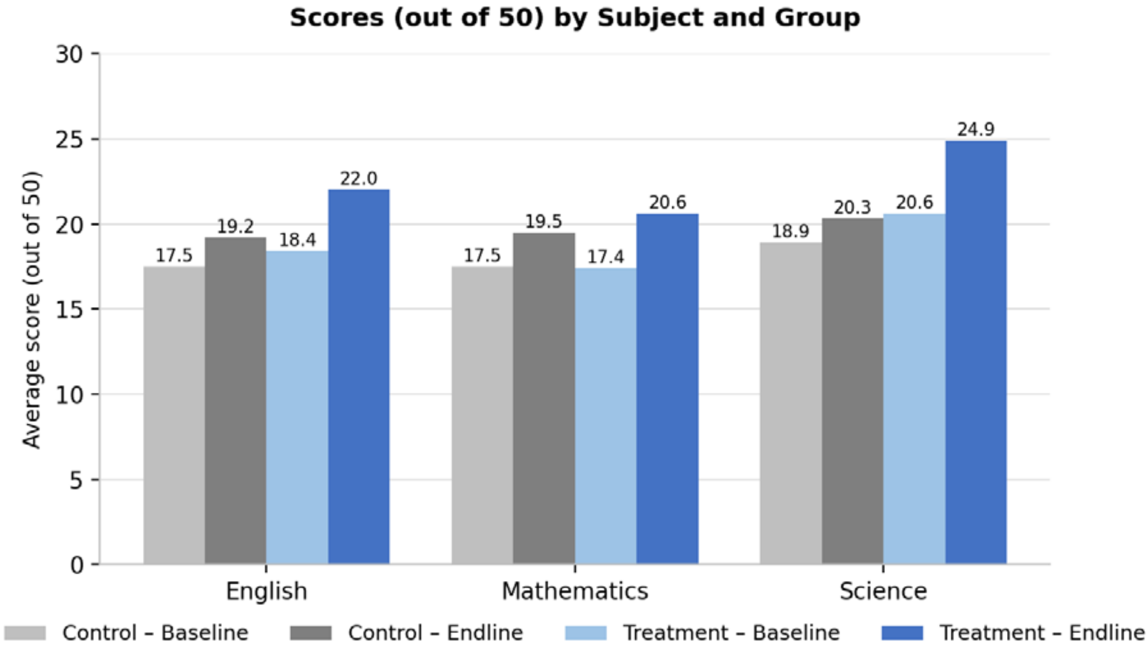

Figure 9: Scores (out of 50) by Subject and Group

- In **English**, students in the control group improved from **17.5 marks** to **19.2 marks**, a gain of **1.7 marks**, while students in the treatment group improved from **18.4 marks** to **22.0 marks**, a gain of approximately **3.6 marks**.
- In **Mathematics**, control group scores increased from **17.5 marks** to **19.5 marks**, representing an improvement of **2.0 marks**, whereas treatment group scores increased from **17.4 marks** to **20.6 marks**, an improvement of **3.2 marks**.
- The largest divergence between groups was observed in **Science**. Control schools recorded a modest increase from **18.9 marks** to **20.3 marks**, equivalent to **1.4 marks**, while treatment schools improved from **20.6 marks** to **24.9 marks**, corresponding to a gain of **4.3 marks**.

Although students in both groups demonstrated improvements over time, the magnitude of gains was consistently larger among students exposed to the AI-enabled adaptive learning intervention. It should be noted that treatment schools began the period with slightly higher average scores in English and Science; the DiD framework applied in the following sections is intended to account for such baseline differences in levels.

6.3 Absolute Changes in Scores

Figure 6 presents changes in average scores between baseline and end line for both treatment and control groups.

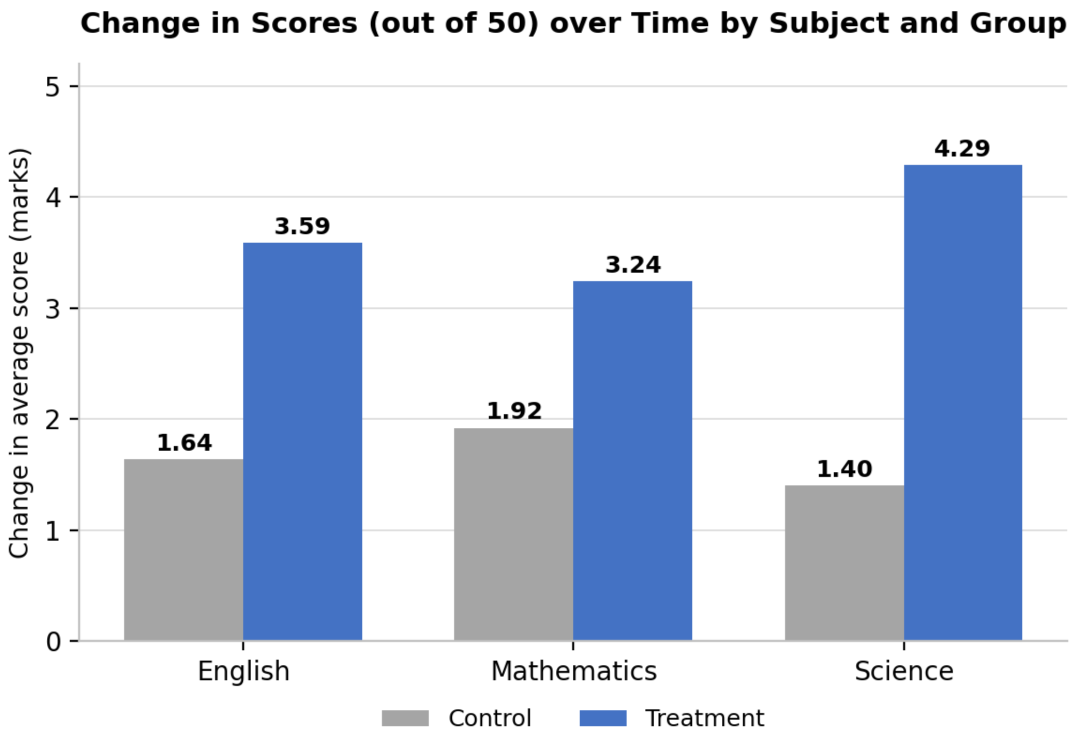

Figure 10: Change in Scores (out of 50) over Time by Subject and Group

Control group students recorded improvements of:

- **1.64 marks** in English;
- **1.92 marks** in Mathematics; and
- **1.40 marks** in science.

In comparison, students in treatment schools achieved substantially larger gains:

- **3.59 marks** in English;
- **3.24 marks** in Mathematics; and
- **4.29 marks** in science.

The findings indicate that treatment schools outperformed control schools in all three subjects. The largest absolute gain among treatment schools was observed in science, where average scores increased by more than four marks over the implementation period.

6.4 Difference-in-Differences Estimates

Figure 7 presents the Difference-in-Differences estimates, which compare changes in learning outcomes in treatment schools relative to changes observed in control schools.

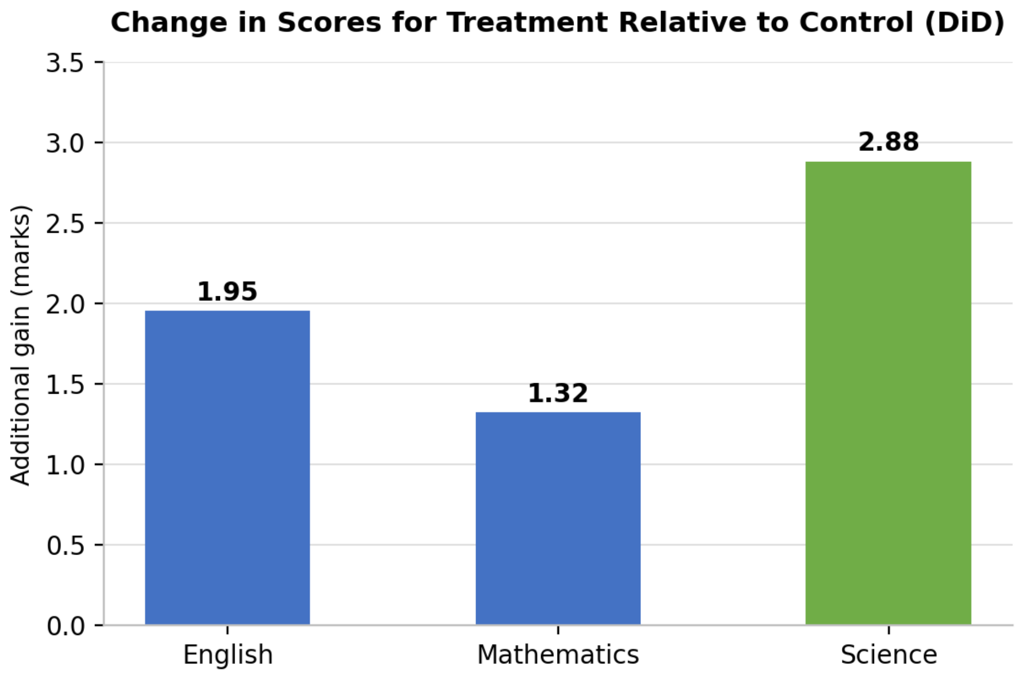

Figure 11: Change in Scores for Treatment Relative to Control (DiD)

The DiD estimates indicate additional gains associated with the intervention of:

- **1.95 marks** in English;
- **1.32 marks** in Mathematics; and
- **2.88 marks** in science.

These estimates suggest that, after accounting for improvements that would have occurred in the absence of the intervention, students exposed to the AI-enabled adaptive learning platform achieved the largest additional gains in science, followed by English and Mathematics.

The results indicate that the intervention was associated with positive learning effects across all subjects, with particularly strong effects in science.

6.5 Percentage Changes in Scores

Figure 8 presents percentage improvements in scores over time by treatment and control group.

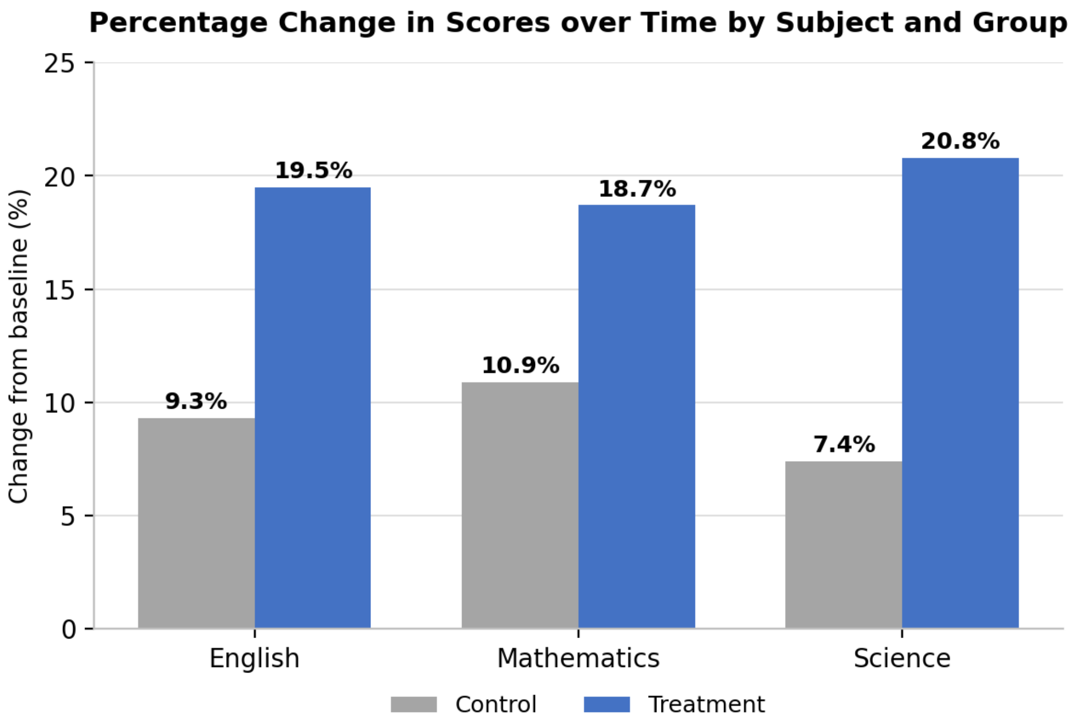

Figure 12: Percentage Change in Scores over Time by Subject and Group

Among control schools, percentage improvements were:

- **9.3 percent** in English;
- **10.9 percent** in Mathematics; and
- **7.4 percent** in science.

Students in treatment schools demonstrated substantially higher gains:

- **19.5 percent** in English;
- **18.7 percent** in Mathematics; and
- **20.8 percent** in science.

Figure 9 presents the Difference-in-Differences estimates in percentage terms. Relative to the control group, treatment students demonstrated additional gains of:

- **10.2 percentage points** in English;
- **7.7 percentage points** in Mathematics; and
- **13.4 percentage points** in science.

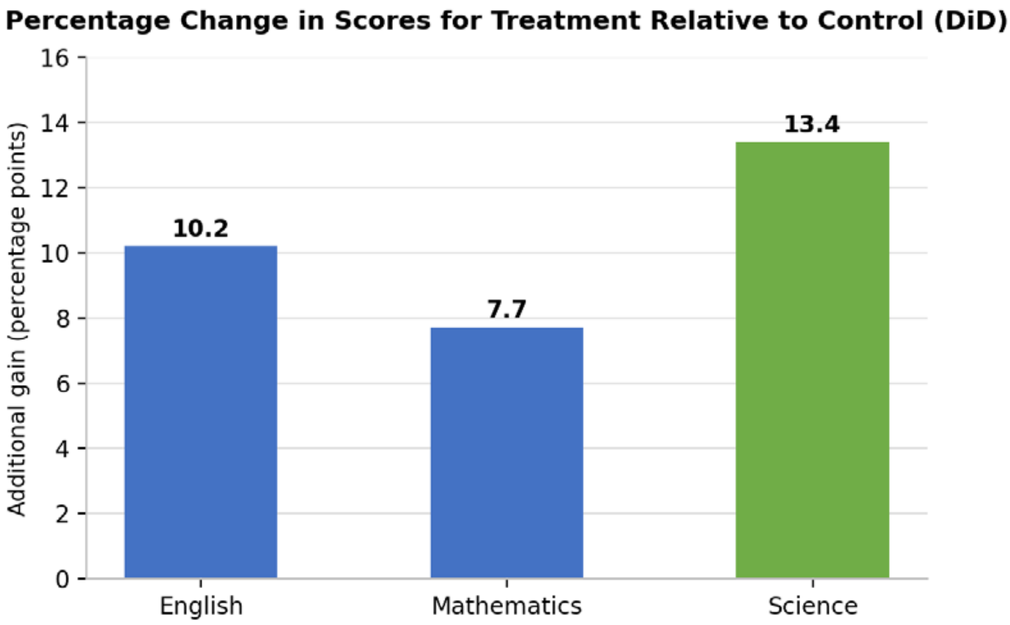

Figure 13: Percentage Change in Scores for Treatment Relative to Control (DiD)

The percentage-based analysis reinforces the findings from the absolute score analysis and highlights the particularly strong intervention effect observed in science.

6.6 Subject-Specific Trends

The subject-specific trend lines represent changes in performance over time and the divergence between treatment and control groups.

6.6.1 English

Figure 10 illustrates changes in English scores over time.

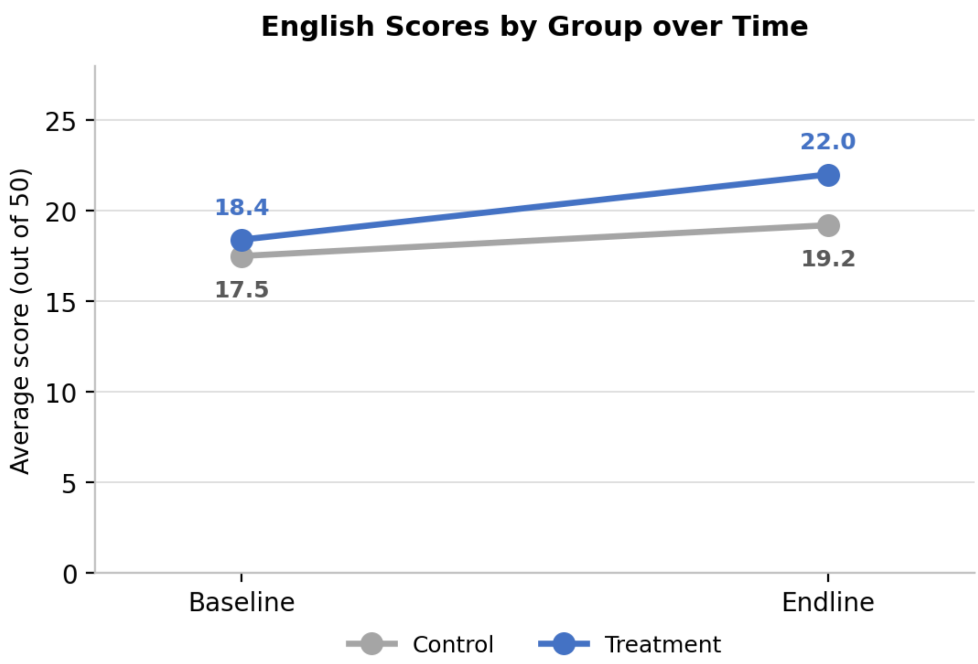

Figure 10: English Scores by Group over Time

At baseline, treatment students slightly outperformed control students, recording an average score of **18.4 marks** compared with **17.5 marks** in the control group. By end line, average scores had increased to **22.0 marks** in treatment schools and **19.2 marks** in control schools.

Although both groups improved over time, the slope of improvement was considerably steeper among treatment students, resulting in an end line gap of nearly three marks between the two groups. This pattern suggests that the intervention was associated with meaningful improvements in English learning outcomes.

6.6.2 Mathematics

Figure 11 presents trends in Mathematics performance.

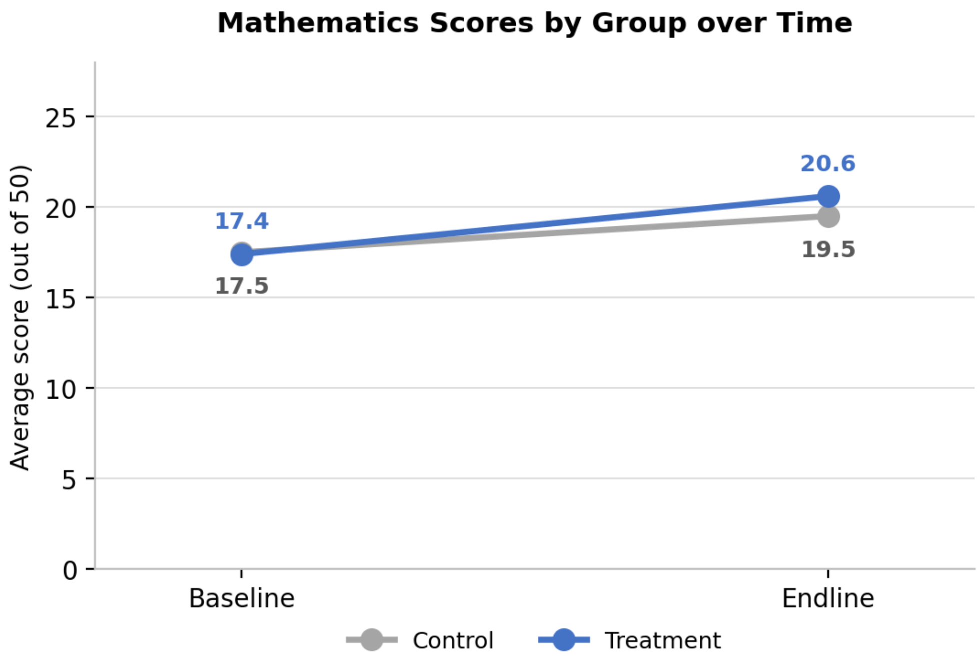

Figure 14: Mathematics Scores by Group over Time

At baseline, the two groups exhibited nearly identical performance levels, with average scores of **17.5 marks** in the control group and **17.4 marks** in the treatment group. By end line, scores increased to **19.5 marks** and **20.6 marks**, respectively.

The near-equivalence of baseline scores strengthens confidence in the comparability of the groups. Although both groups improved over time, treatment students demonstrated larger gains, suggesting a positive contribution of the intervention to Mathematics learning outcomes.

6.6.3 Science

Figure 12 shows trends in science scores.

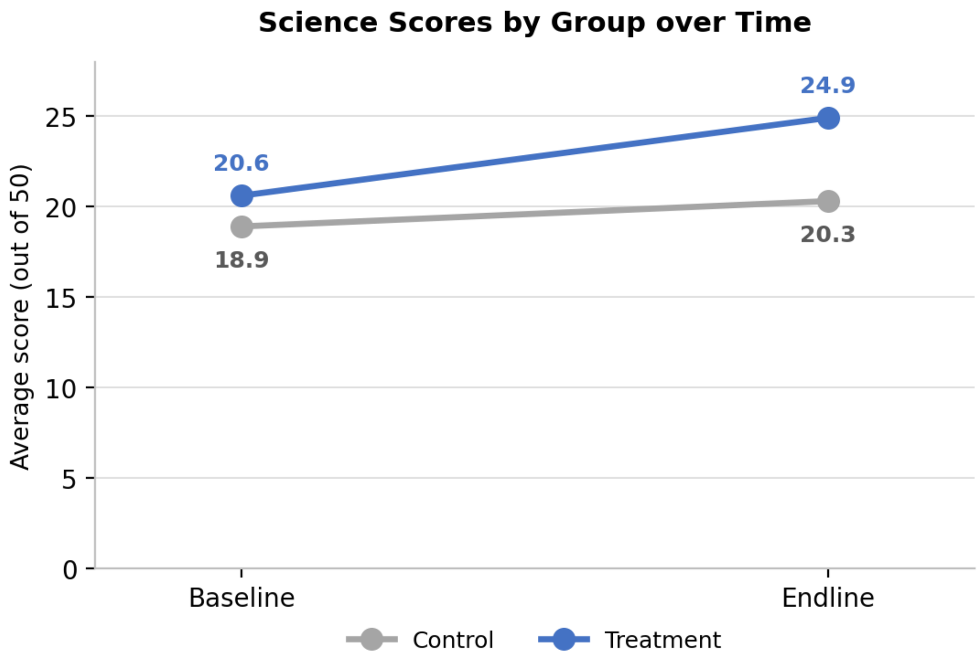

Figure 15: Science Scores by Group over Time

Treatment schools began with a somewhat higher baseline score (**20.6 marks**) than control schools (**18.9 marks**). By end line, treatment group scores had increased substantially to **24.9 marks**, while control group scores increased more modestly to **20.3 marks**.

The widening gap between the two groups over time indicates a strong intervention effect in Science, with treatment students achieving markedly larger improvements than their peers in control schools.

6.7 Gender-Based Performance

Given that female students constituted approximately 79 percent of the assessed sample, disaggregation by gender is essential to a proper interpretation of the aggregate results presented above.

6.7.1 Female Students – Overall

Figure 13 presents average scores for female students across all schools.

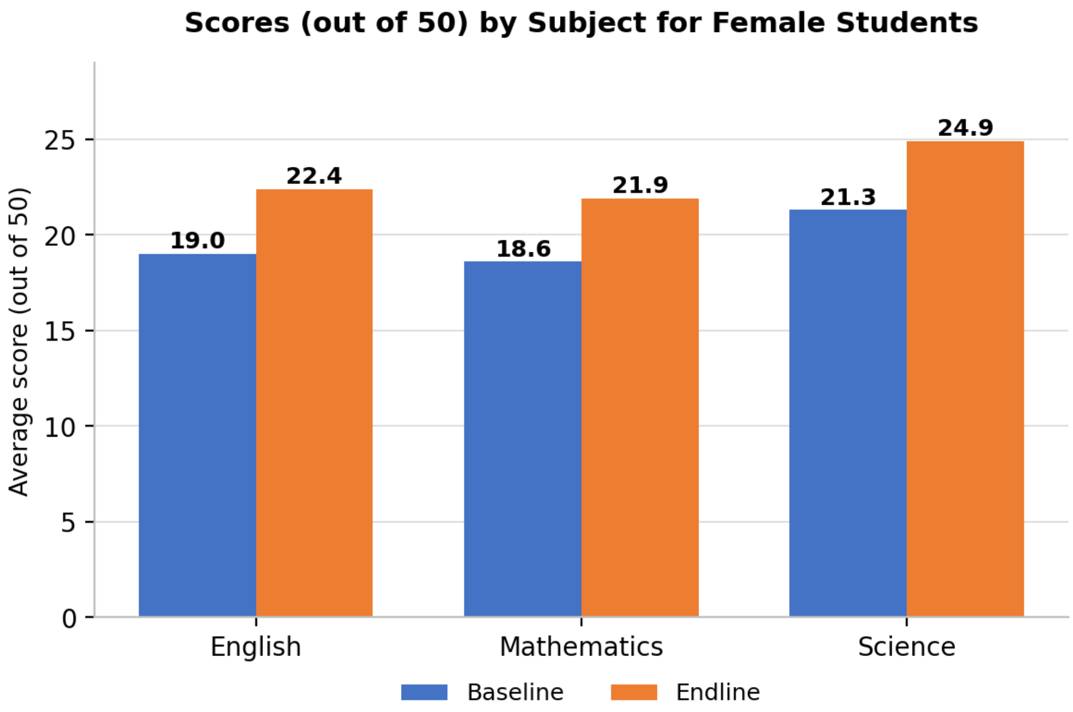

Figure 16: Scores (out of 50) by Subject for Female Students

Female students demonstrated improvements across all subjects. Average English scores increased from **19.0 marks** to **22.4 marks**, Mathematics scores improved from **18.6 marks** to **21.9 marks**, and Science scores increased from **21.3 marks** to **24.9 marks**.

The findings indicate substantial learning gains among female students, with the largest improvement observed in Science.

6.7.2 Male Students – Overall

Figure 14 presents average scores for male students.

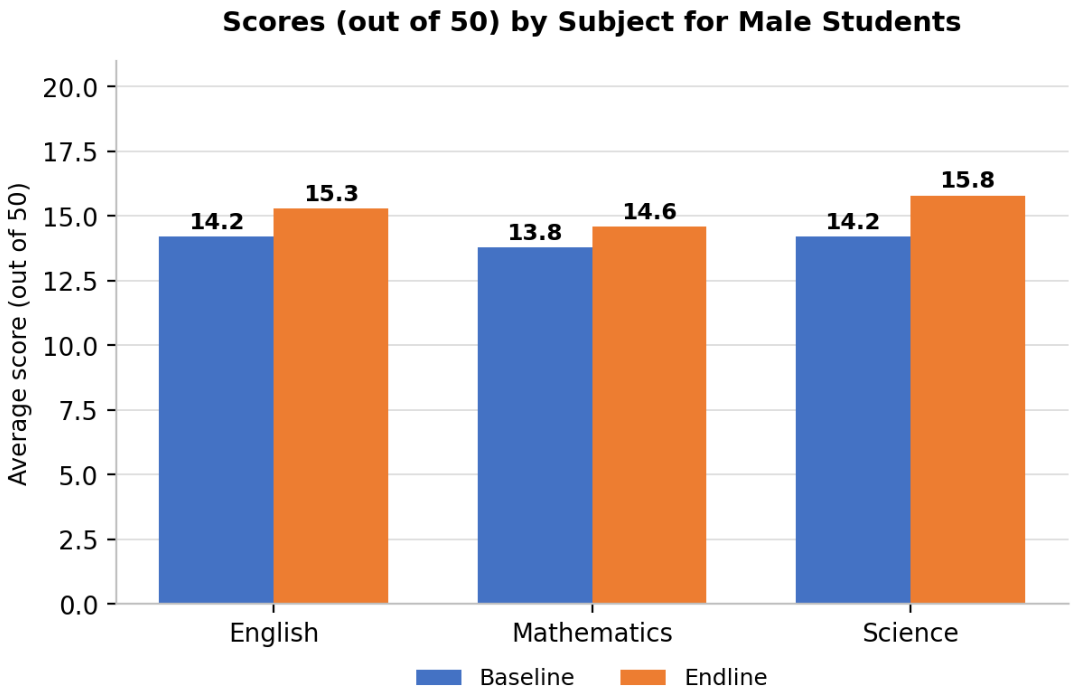

Figure 17: Scores (out of 50) by Subject for Male Students

Male students also demonstrated improvements across all subjects, although gains were smaller than those observed among female students. English scores increased from **14.2 marks** to **15.3 marks**, Mathematics scores improved from **13.8 marks** to **14.6 marks**, and Science scores increased from **14.2 marks** to **15.8 marks**.

Overall, **female students outperformed male students at both baseline and endline across all three subjects**.

6.7.3 Female Students – Treatment Group

Figure 15 presents learning outcomes among female students in treatment schools.

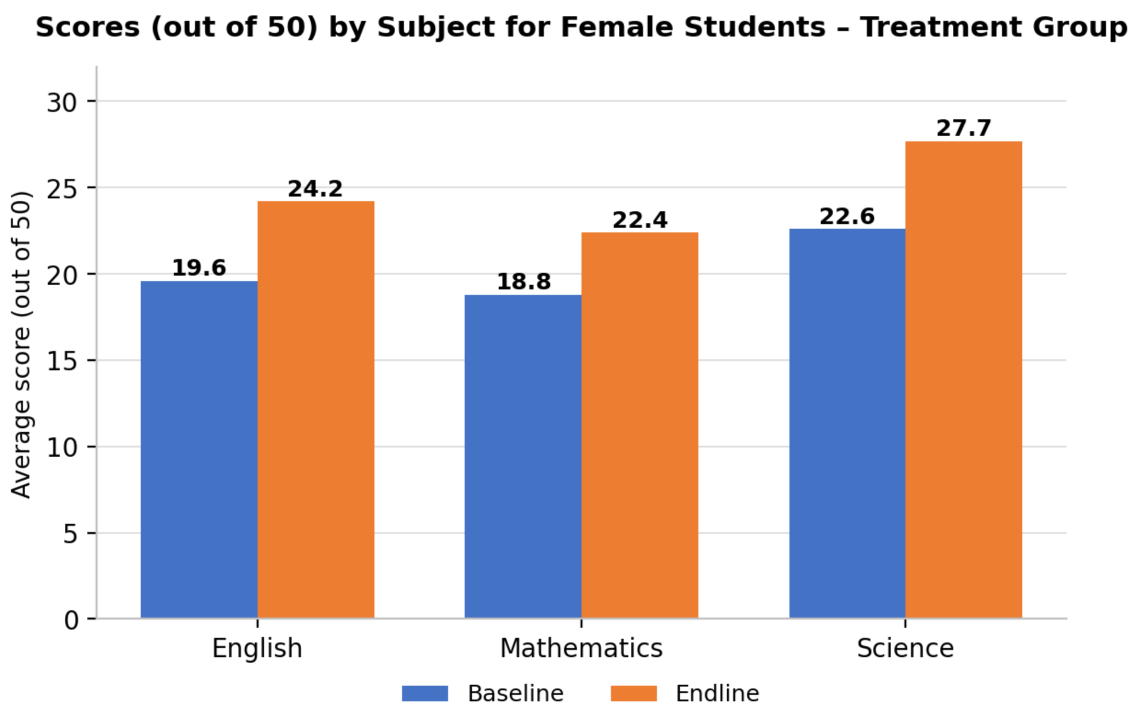

Figure 18: Scores (out of 50) by Subject for Female Students – Treatment Group

Average English scores increased from **19.6 marks** to **24.2 marks**, representing a gain of **4.6 marks**. Mathematics scores improved from **18.8 marks** to **22.4 marks**, a gain of **3.6 marks**, while Science scores increased from **22.6 marks** to **27.7 marks**, an improvement of **5.1 marks**.

These findings indicate particularly strong improvements among female students exposed to the intervention, especially in Science.

6.7.4 Male Students – Treatment Group

Figure 16 presents results for male students in treatment schools.

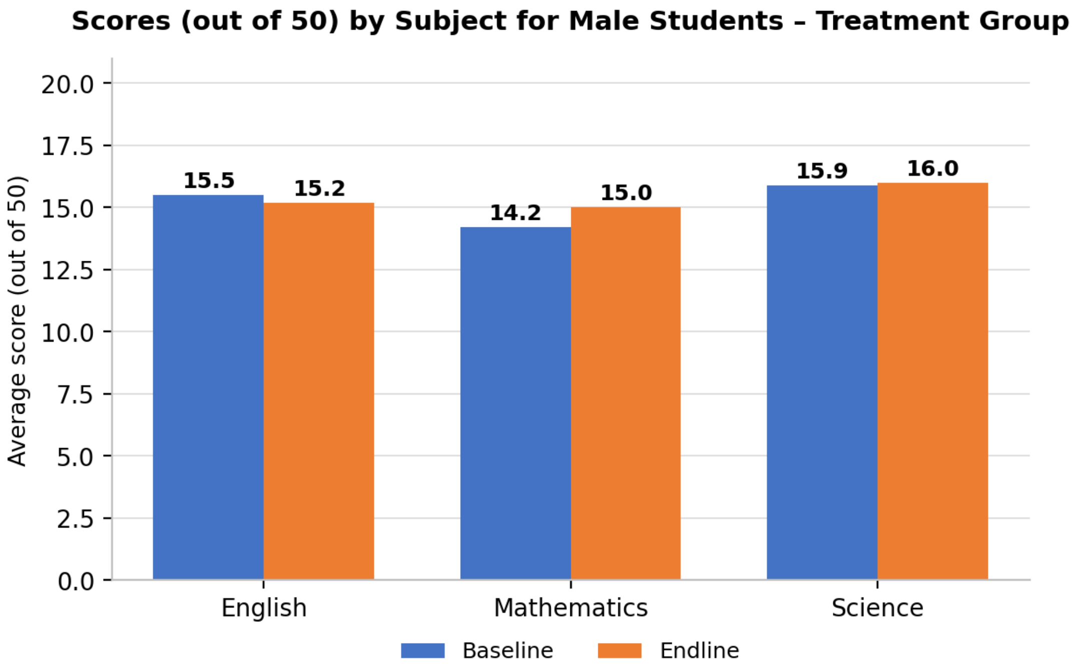

Figure 19: Scores (out of 50) by Subject for Male Students – Treatment Group

Average English scores declined marginally from **15.5 marks** at baseline to **15.2 marks** at endline. Mathematics scores improved modestly from **14.2 marks** to **15.0 marks**, while Science scores remained broadly unchanged, increasing only marginally from **15.9 marks** to **16.0 marks**.

Figure 17 compares the within-treatment change for female and male students directly.

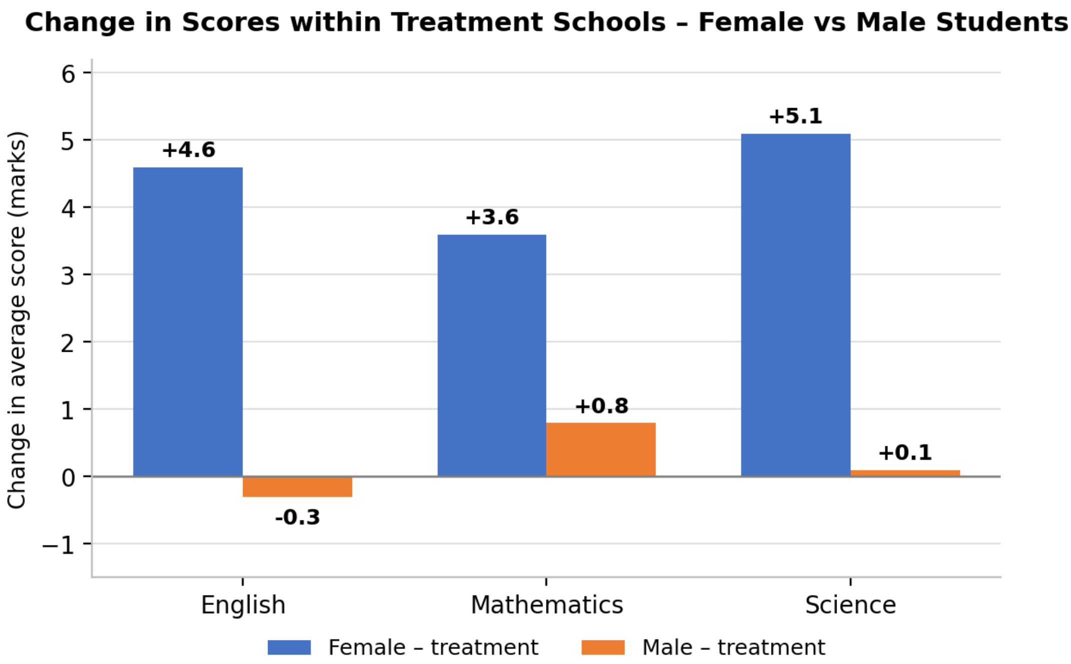

Figure 20: Change in Scores within Treatment Schools – Female vs Male Students

The results suggest that the benefits observed during the pilot were considerably stronger among female students than among male students. While male students demonstrated limited improvement, female students experienced substantial gains across all subjects. Several factors may plausibly contribute to this divergence, including the smaller size of the male treatment sub-sample (300 students across 15 schools), differences in baseline attainment, and variation in implementation intensity, computer laboratory availability and connectivity between male and female schools. The available data do not permit these explanations to be distinguished from one another, and this pattern should therefore be treated as a finding requiring further investigation rather than as a settled conclusion. It is discussed further in Sections 7 and 8.

6.8 Summary of Learning Gains

Table 3 consolidates the descriptive results presented in Sections 6.1 to 6.5.

Subject	Control: baseline → endline (change)	Treatment: baseline → endline (change)	DiD (marks)
English	17.5 → 19.2 (+1.64)	18.4 → 22.0 (+3.59)	+1.95
Mathematics	17.5 → 19.5 (+1.92)	17.4 → 20.6 (+3.24)	+1.32
Science	18.9 → 20.3 (+1.40)	20.6 → 24.9 (+4.29)	+2.88

Table 3: Summary of Learning Gains by Subject and Group

6.9 Difference-in-Differences Regression Results

To formally estimate the effect of the intervention, Difference-in-Differences regression models were estimated separately for English, Mathematics and Science. The interaction term between treatment status and time represents the estimated intervention effect after controlling for baseline differences and overall time trends.

The regression results indicate that the **intervention had a positive and statistically significant effect across all three subjects**.

The estimated treatment effect for **English** was **1.949 marks** and was statistically significant (**p = 0.001**), indicating that students in treatment schools improved by an additional 1.95 marks relative to students in control schools after controlling for baseline differences.

```
-> SUBJECT = ENGLISH

      Source |       SS           df       MS      Number of obs   =     5,482
-------------+----------------------------------   F(3, 5478)      =     47.15
       Model |  14995.3998         3  4998.46661   Prob > F        =    0.0000
    Residual |  580774.046     5,478  106.019358   R-squared       =    0.0252
-------------+----------------------------------   Adj R-squared   =    0.0246
       Total |  595769.446     5,481  108.697217   Root MSE        =    10.297

------------------------------------------------------------------------------
       Score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.treat |   .8642063   .3657553     2.36   0.018     .1471807    1.581232
      1.time |   1.635924   .4114977     3.98   0.000      .829225    2.442623
             |
  treat#time |
        1 1  |   1.949121   .5646118     3.45   0.001      .842258    3.055985
             |
       _cons |   17.52601   .2625514    66.75   0.000      17.0113    18.04071
------------------------------------------------------------------------------
```

Figure 21: DiD Regression Results – English

For **Mathematics**, the estimated treatment effect was **1.322 marks**, also statistically significant (**p = 0.011**). This represents the smallest of the three subject effects, though it remains statistically distinguishable from zero at conventional thresholds.

```
-> SUBJECT = MATH

      Source |       SS           df       MS      Number of obs   =     5,750
-------------+----------------------------------   F(3, 5746)      =     37.11
       Model |  10543.6926         3  3514.56419   Prob > F        =    0.0000
    Residual |  544128.013     5,746  94.6968349   R-squared       =    0.0190
-------------+----------------------------------   Adj R-squared   =    0.0185
       Total |  554671.706     5,749  96.4814239   Root MSE        =    9.7312

------------------------------------------------------------------------------
       Score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.treat |  -.1766688   .3385133    -0.52   0.602    -.8402825    .4869449
      1.time |   1.917394   .3790887     5.06   0.000     1.174237    2.660551
             |
  treat#time |
        1 1  |   1.322953   .5204426     2.54   0.011     .3026896    2.343217
             |
       _cons |   17.53511   .2425994    72.28   0.000     17.05953     18.0107
------------------------------------------------------------------------------
```

Figure 22: DiD Regression Results – Mathematics

The largest effect was observed in **science**, where the intervention was associated with an additional gain of **2.88 marks**, significant at the one percent level (**p < 0.001**).

```
-> SUBJECT = SCIENCE

      Source |       SS           df       MS      Number of obs   =     5,475
-------------+----------------------------------   F(3, 5471)      =     78.58
       Model |  26242.7042         3  8747.56807   Prob > F        =    0.0000
    Residual |  609064.422     5,471  111.325977   R-squared       =    0.0413
-------------+----------------------------------   Adj R-squared   =    0.0408
       Total |  635307.126     5,474  116.059029   Root MSE        =    10.551

------------------------------------------------------------------------------
       Score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1.treat |   1.715135   .3750312     4.57   0.000     .9799248    2.450345
      1.time |   1.403029   .4184916     3.35   0.001     .5826194     2.22344
             |
  treat#time |
        1 1  |   2.882513   .5783236     4.98   0.000     1.748769    4.016257
             |
       _cons |   18.91456   .2684318    70.46   0.000     18.38833     19.4408
------------------------------------------------------------------------------
```

Figure 23: DiD Regression Results – Science

Subject	Treatment effect (marks)	p-value	Statistical significance
English	1.949	0.001	Significant at 1 percent
Mathematics	1.322	0.011	Significant at 5 percent
Science	2.880	< 0.001	Significant at 1 percent

Table 4: Summary of Difference-in-Differences Regression Results

The regression results therefore corroborate the findings presented in the preceding sections. Across all three subjects, students exposed to the AI-enabled adaptive learning intervention achieved significantly greater improvements than students in control schools. The strongest evidence of programme impact is observed in science, followed by English and Mathematics, suggesting that the AI-enabled adaptive learning platform contributed positively to student learning outcomes during the implementation period.

7. Limitations

The findings presented in this report should be interpreted in light of a number of design, implementation and measurement constraints.

- **Limited geographic coverage.** The pilot was confined to two districts, Peshawar and Nowshera. Both are comparatively well-served in terms of digital infrastructure, and the results should not be generalized to the province as a whole, particularly to districts with weaker connectivity or fewer functional computer laboratories.
- **Short intervention duration.** Platform-based learning took place over a period of approximately six to eight weeks between teacher orientation and the commencement of endline assessment. This is a short window within which to expect and detect durable changes in learning.
- **Restriction to computer-equipped schools.** Only schools with functional computer laboratories were eligible for inclusion. The sampling frame therefore excludes a substantial share of public schools in the province, and the estimated effects apply to a relatively better-resourced segment of the system.
- **Single grade.** The intervention targeted Grade 8 only. Effects may differ at other grade levels where curricular demands and student digital fluency vary.
- **Sample composition skewed towards female students.** Female students accounted for approximately 79 percent of the assessed sample, a consequence of the census inclusion of eligible female schools in Peshawar. Aggregate results are therefore driven substantially by outcomes among girls.
- **Small and divergent male sub-sample.** Only 300 male students in 15 treatment schools were assessed, and this group recorded minimal change over the period. The reasons for this divergence cannot be established from the available data and require dedicated follow-up.
- **Baseline imbalance in two subjects.** Treatment schools recorded higher baseline averages in English (18.4 versus 17.5) and Science (20.6 versus 18.9). While the DiD framework accounts for differences in levels, it relies on the assumption that both groups would have followed parallel trajectories in the absence of the intervention. This assumption cannot be tested directly with two time points.
- **No long-term follow-up.** Learning was measured immediately at the end of the implementation period. The persistence of observed gains beyond that point is unknown.
- **Variation in implementation fidelity.** Monitoring visits recorded differences between schools in the availability of functional computers and the reliability of internet connectivity. Exposure to the platform is therefore likely to have varied across treatment schools, and platform usage data were not used to construct a formal measure of dosage.
- **Limited qualitative evidence.** Systematic qualitative data collection from students and teachers was not undertaken as part of the pilot, restricting the ability to explain the mechanisms underlying the observed quantitative patterns.

8. Conclusions

The AI-enabled adaptive learning pilot in Khyber Pakhtunkhwa demonstrates the potential of technology-enabled personalised learning to strengthen teaching and learning within the province's public education system. By combining curriculum-aligned diagnostic assessments, adaptive learning pathways and real-time progress monitoring, the intervention provided students with targeted learning support while equipping teachers with actionable insight into individual learning needs. The findings indicate that students who participated in the intervention achieved significantly greater improvements in English, Mathematics and Science than their peers in control schools, providing evidence that AI-enabled adaptive learning can contribute to improved learning outcomes when implemented within existing school structures.

The magnitude of the estimated effects is meaningful in the context of a short implementation period. An additional gain of 2.88 marks out of 50 in Science, achieved over roughly six to eight weeks of platform-based learning, is a substantial return relative to the incremental cost of deploying a web-based platform in schools that already possess the required hardware. The consistency of the direction of effect across all three subjects, and its statistical significance in each, strengthens confidence that the pattern is not the product of chance variation in a single subject.

At the same time, the results are not uniform. Gains were concentrated among female students, who made up the large majority of the sample, while male students in treatment schools showed little measurable change. This divergence is the most important open question arising from the pilot. It may reflect the small size of the male sub-sample, differences in baseline attainment, or variation in the conditions under which the platform was used in male and female schools. Resolving it should be a priority before decisions on wider rollout are taken, since a scaled programme that delivers benefits to only one part of the student population would fall short of its objective.

Beyond its effect on student achievement, the pilot offers valuable lessons for the integration of digital learning technologies in government schools in Khyber Pakhtunkhwa. The experience highlights the importance of teacher capacity building, reliable digital infrastructure, curriculum alignment and continuous technical support for successful implementation. The relative ease with which 44 IT teachers were oriented and 50 schools brought onto the platform within a single academic term indicates that the operational model is workable within existing institutional arrangements, including the professional development architecture of the DPD and its regional centres.

Taken together, the implementation experience demonstrates that AI-enabled adaptive learning tools can be integrated into public sector schools within bounded and clearly specified conditions. The programme's delivery underscores the value of treating technology as a system-support mechanism, anchored in curriculum alignment, structured assessment and institutional coordination, rather than as a standalone solution. AI should be viewed as a tool to strengthen teacher effectiveness rather than to replace classroom teaching.

The pilot findings show that AI-enabled adaptive learning is technically feasible within Khyber Pakhtunkhwa's public education system, can be implemented using existing school infrastructure, and produces measurable improvements in student learning under the conditions tested. Building on these findings, future efforts should focus on refining the platform, investigating the differential effects observed between girls and boys, strengthening implementation mechanisms, and exploring opportunities for scaling the intervention to enable more personalised, data-driven learning across the province's education system.

9. Recommendations

Based on the implementation experience and evaluation findings, the following recommendations are proposed to guide the institutionalisation and scale-up of AI-enabled adaptive learning within Khyber Pakhtunkhwa's public education system.

9.1 Immediate Priorities (Next 12 Months)

1. **Investigate the differential effects observed between female and male students.** Before any expansion, E&SED should commission a focused diagnostic exercise in the male treatment schools to establish why gains were minimal. This should combine analysis of platform usage data, structured interviews with participating teachers and headteachers, and verification of computer laboratory functionality and connectivity. The findings should directly inform the design of any subsequent phase.
2. **Develop a provincial roadmap for scaling AI-enabled adaptive learning.** E&SED, working with the DPD, should develop a phased implementation strategy outlining priorities for expansion, infrastructure requirements, financing, governance arrangements and monitoring mechanisms. The roadmap should identify priority districts and grades for future implementation based on system readiness and learning needs.
3. **Institutionalise the AI-enabled learning platform within the department's digital ecosystem.** The platform should be hosted and managed within the department's own data infrastructure to ensure long-term sustainability, government ownership, data security and integration with existing education management and monitoring systems. Standard operating procedures should be established for platform maintenance, user management, technical support and data governance.
4. **Strengthen teacher capacity to use AI-enabled learning tools effectively.** The DPD should develop and deliver structured professional development that equips teachers to integrate AI-enabled adaptive learning into classroom instruction. Training should extend beyond platform operation to include interpretation of learner analytics, use of assessment data for instructional planning, and strategies for facilitating personalised learning. The existing RPDC network provides a ready delivery mechanism.
5. **Address infrastructure gaps identified during implementation.** A rapid audit of computer laboratory functionality and connectivity in schools identified for future rollout should be conducted, with minimum thresholds defined and remediation costed before schools are onboarded.

9.2 Medium-Term Priorities (1–3 Years)

1. **Implement a phased expansion of the intervention.** Following successful institutionalisation, E&SED should gradually expand the AI-enabled adaptive learning model to additional districts and grade levels, prioritising schools with adequate digital infrastructure and demonstrated implementation readiness. Expansion should be accompanied by continuous monitoring, technical support and iterative refinement based on implementation evidence.

2. **Extend the intervention period and measure dosage.** Future implementation should allow a longer window of platform-based learning and should record and analyse individual student exposure, so that effects can be related to actual usage rather than to assignment alone.

3. **Develop integrated AI-supported teaching and learning resources.** Curriculum-aligned digital teaching and learning resources should be developed that integrate AI-enabled adaptive learning with classroom instruction. These resources should include lesson plans, teacher guides, assessment activities and differentiated learning materials linked to SLOs.

4. **Embed AI-enabled learning within teacher professional development systems.** The DPD should integrate AI literacy, digital pedagogy and the effective use of adaptive learning technologies into both pre-service teacher education and continuous professional development programmes, ensuring that teachers are equipped to use emerging technologies as part of routine classroom practice.

9.3 Long-Term System Strengthening

1. **Institutionalise AI-enabled adaptive learning within the provincial education system.** E&SED should position AI-enabled learning as a complementary component of the province's teaching and learning strategy by integrating adaptive learning platforms with curriculum implementation, classroom assessment and student support systems.

2. **Establish an integrated learning analytics framework.** Mechanisms should be developed to combine data from AI-enabled learning platforms, school-based assessments and other education management systems to generate actionable evidence for teachers, school leaders and policymakers. Such integration would enable more effective monitoring of student progress and inform evidence-based planning and resource allocation.

3. **Conduct long-term evaluations of effectiveness, scalability and cost-efficiency.** Future research should assess the sustainability of learning gains, implementation fidelity at scale, and the cost-effectiveness of AI-enabled adaptive learning. Evaluations should be designed with sufficient statistical power to detect effects separately for girls and boys. Findings should inform policy decisions regarding province-wide adoption and future investments in educational technology.

4. **Strengthen partnerships for AI innovation in education.** E&SED should foster collaboration between government, academia, technology partners and development organisations to continuously improve AI-enabled learning solutions, evaluate emerging technologies, and ensure that innovations remain aligned with curriculum standards, ethical considerations and evolving educational needs.