UK International Development
Partnership | Progress | Prosperity

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

Data Driven Mentoring and Coaching for Multigrade Settings in Public School

Khyber Pakhtunkhwa

Bridging Technical Assistance for Governments (B-TAG)

Data-Driven Mentoring and Coaching for Multigrade Settings in Public Schools

Khyber Pakhtunkhwa (KP)

Bridging Technical Assistance for Governments (B-TAG)

Table of Contents

1. Executive Summary

Multigrade teaching is widespread across Khyber Pakhtunkhwa's public primary education system, where a large proportion of schools operate with limited teaching staff and a single teacher is frequently responsible for delivering instruction to two or more grades simultaneously. To strengthen multigrade teaching practices, the Bridging Technical Assistance for Governments (B-TAG) programme has been working closely with the Elementary and Secondary Education Department (E&SED) and the Directorate of Professional Development (DPD) in KP to support teachers through professional development, instructional resources, and data-driven academic support mechanisms. A key component of this support has been strengthening the capacity of School Leaders (SLs) and Assistant Sub-Divisional Education Officers (ASDEOs) to provide evidence-based mentoring and coaching using classroom observation data collected through a specialised Multigrade Classroom Observation Tool (MG-COT).

As part of this support, B-TAG designed and delivered a training programme on data-driven mentoring and coaching for multigrade settings in public schools of Khyber Pakhtunkhwa using the MG-COT. The programme was piloted with 70 ASDEOs (33 female, 37 male) and 126 School Leaders (59 female, 67 male) across six districts of the province. The ASDEO training was delivered at the DPD, Peshawar, and was followed by cascade training of School Leaders in February 2026. Both trainings were jointly facilitated by DPD and the B-TAG team and focused on applying MG-COT data for instructional supervision, structured feedback, and coaching in multigrade classroom settings.

An evaluation of the programme was conducted to assess whether the intervention contributed to improvements in multigrade teaching practices in public primary schools across selected districts of Khyber Pakhtunkhwa. The evaluation employed a quasi-experimental pre- and post-intervention design using treatment and control groups of schools. Baseline and endline classroom observations were conducted using the MG-COT framework, which assesses 13 core multigrade teaching practices related to classroom organisation, lesson planning, differentiated instruction, peer learning, formative assessment, and individualised feedback.

Overall, the evaluation covered 528 public primary schools (351 treatment, 177 control) across six districts of Khyber Pakhtunkhwa, including 251 girls' schools (48 percent) and 277 boys' schools (52 percent). A total of 904 classroom observations were conducted during the evaluation, distributed equally between female schools (452) and male schools (452).

The findings indicate substantial improvements in multigrade teaching practices between baseline and endline across both treatment and control schools; however, improvements were consistently and markedly greater in treatment schools linked with the trained SLs and ASDEOs.

The overall average MG-COT score in treatment schools increased from 1.72 at baseline to 3.70 (out of 5) at endline, representing an improvement of 1.98 points. In comparison, control schools improved from 1.61 to 2.09 (out of 5), representing an increase of 0.48 points.

Practice-wise analysis also showed stronger gains across treatment schools for all 13 MG-COT teaching practices. The largest improvements in treatment schools were observed in:

- use of multigrade lesson plans (2.08 points);

- effective classroom space use (2.02 points);
- content aligned with learning objectives (2.01 points);
- flexible student grouping (2.01 points); and
- use of multigrade academic calendars (1.99 points).
- Comparatively, improvements in control schools ranged between 0.41 and 0.52 points across the 13 practices.

The findings suggest that teachers in treatment schools adopted stronger student-centred and differentiated instructional practices over the intervention period relative to teachers in control schools. Gender-wise analysis similarly demonstrated stronger improvement in treatment schools for both female and male schools. Female treatment schools improved by 2.06 points (from 1.66 to 3.72) compared with 0.40 points in female control schools. Male treatment schools improved by 1.89 points (from 1.79 to 3.68) compared with 0.54 points in male control schools.

To estimate the causal impact of the intervention, the evaluation applied an Analysis of Covariance (ANCOVA) framework controlling for baseline MG-COT scores. The ANCOVA results confirm that the intervention had a positive and statistically significant impact on multigrade teaching practices.

The analysis shows that:

- treatment schools scored 1.626 points higher at endline compared with control schools after controlling for baseline scores ($p < 0.001$);
    - female treatment schools scored 1.691 points higher than female control schools ($p < 0.001$); and
    - male treatment schools scored 1.512 points higher than male control schools ($p < 0.001$).
- Out of the 13 indicators included in the MG-COT, there was an increase between +0.31 to 0.52 in overall practice gain.

The findings provide evidence that strengthening the mentoring and coaching capacity of SLs and ASDEOs contributed to measurable improvements in classroom instructional practices within multigrade public primary schools in Khyber Pakhtunkhwa.

The evaluation further highlights the importance of structured mentoring systems, classroom observation frameworks, and continuous professional support mechanisms for teachers working in multigrade classroom settings. The findings suggest that strengthening academic support systems through trained field-based education managers can improve instructional quality even within resource-constrained school environments.

The intervention was not designed or implemented in a way that would be expected to produce measurable changes in SBA outcomes within a single academic year.

Change in student learning outcomes was measured using routine School-Based Assessment (SBA) data. The intervention was implemented over a relatively short period from January 2026 to March 2026. Since the real-time MG-COT application was not fully operational during short tenure of implementation, the SBA analysis does not provide a definitive measure of programme impact on student learning. The classroom observation findings should therefore be regarded as the principal evidence of programme effectiveness, while the SBA results provide complementary contextual information.

2. Introduction

2.1 Background and Context

Improving the quality of teaching and learning in multigrade classrooms remains a significant challenge within Khyber Pakhtunkhwa's public primary education system, particularly in schools operating with limited teaching staff. In many four-teacher schools, a single teacher is responsible for managing students from multiple grade levels within the same classroom. Such settings require effective classroom organisation, differentiated instruction, lesson planning, time management, and continuous formative assessment to ensure meaningful student engagement and learning.

While the Government of Khyber Pakhtunkhwa continues to address teacher shortages and infrastructure gaps through teacher recruitment and classroom expansion, complementary efforts are also being undertaken to strengthen teaching practices and quality of instruction in multigrade settings. Strengthening academic mentoring and coaching systems for teachers has emerged as an important strategy for improving classroom practices and supporting teachers working in resource-constrained school environments.

Under the FCDO-funded B-TAG programme, technical assistance has been provided to the E&SED and the DPD to strengthen multigrade teaching practices across the province. In KP under the B-TAG support, more than 1,200 primary school teachers were trained through a face-to-face training delivered by the Directorate of Professional Development (DPD) with technical assistance from B-TAG. Building on the success of the pilot, the scale-up programme was designed as a blended model reaching out to more than 11,000 primary school teachers with face-to-face coaching and mentoring provided by 1,756 ASDEOs and SLs. As part of this work, a specialised MG-COT was developed to assess 13 core multigrade teaching practices and to support evidence-based classroom observation, feedback, and mentoring.

Building on these efforts, the current intervention focused on strengthening the mentoring and coaching capacity of SLs and ASDEOs using data-driven, evidence-informed practices in multigrade classroom settings.

2.2 Data-driven Coaching and Mentoring

In Khyber Pakhtunkhwa, ASDEOs and School Leaders serve as the critical link between education policy and classroom practice and play a central role in providing academic support to teachers through school visits, classroom observations, and professional feedback. To enhance this role, a training programme was designed to equip them with the skills necessary to use MG-COT data, give targeted feedback to teachers, and keep track of actions taken.

The intervention specifically focused on strengthening the understanding and application of the MG-COT framework among SLs and ASDEOs, as well as their ability to support teachers in key areas of multigrade pedagogy, including instructional planning, classroom management, differentiated instruction, peer learning, formative assessment, and student engagement. Through regular school visits, SLs and ASDEOs were expected to mentor teachers in adopting and sustaining effective multigrade teaching practices.

2.3 Pilot Implementation and Evaluation

The intervention was implemented in selected four-teacher public primary schools across six districts of Khyber Pakhtunkhwa using a treatment and control group design. The primary objective of the evaluation was to assess whether strengthening the mentoring and coaching capacity of SLs and ASDEOs contributes to improvements in multigrade teaching practices.

Specifically, the evaluation sought to:

- Assess the status of multigrade teaching practices in treatment schools at baseline and endline, to be compared with a control sample;
- Examine changes in MG-COT scores over the intervention period;
- Estimate the impact of the intervention on overall multigrade teaching practices;
- Assess intervention effects across individual MG-COT teaching practices and by school gender category.

With regular observation and feedback using the MG-COT framework, the intervention sought to establish a continuous cycle of professional support and improvement for teachers. Enhanced classroom practices were ultimately expected to contribute to improved teaching quality and more effective learning experiences for students in multigrade public primary schools.

This report presents the implementation details as well as the baseline and endline findings of the evaluation and examines the extent to which the intervention influenced multigrade teaching practices in treatment schools relative to control schools. The Analysis of Covariance (ANCOVA) framework was used to estimate intervention effects while controlling for baseline differences in classroom practices.

3. Methodology

3.1 Evaluation Design

The evaluation employed a quasi-experimental pre- and post-intervention design to assess the effectiveness of the training for SLs and ASDEOs on data-driven mentoring and coaching for multigrade settings in public schools of Khyber Pakhtunkhwa. The design included treatment and control groups of schools to enable comparison of multigrade teaching practices between schools exposed to the intervention and those not receiving the intervention during the study period.

Baseline data were collected prior to the implementation of the intervention in January 2026, while endline data were collected following the completion of the intervention period in March 2026. Classroom observations at both stages were conducted using the MG-COT to assess changes in teaching practices across multigrade classrooms.

The evaluation primarily focused on estimating the extent to which the intervention contributed to improvements in multigrade teaching practices in treatment schools relative to control schools.

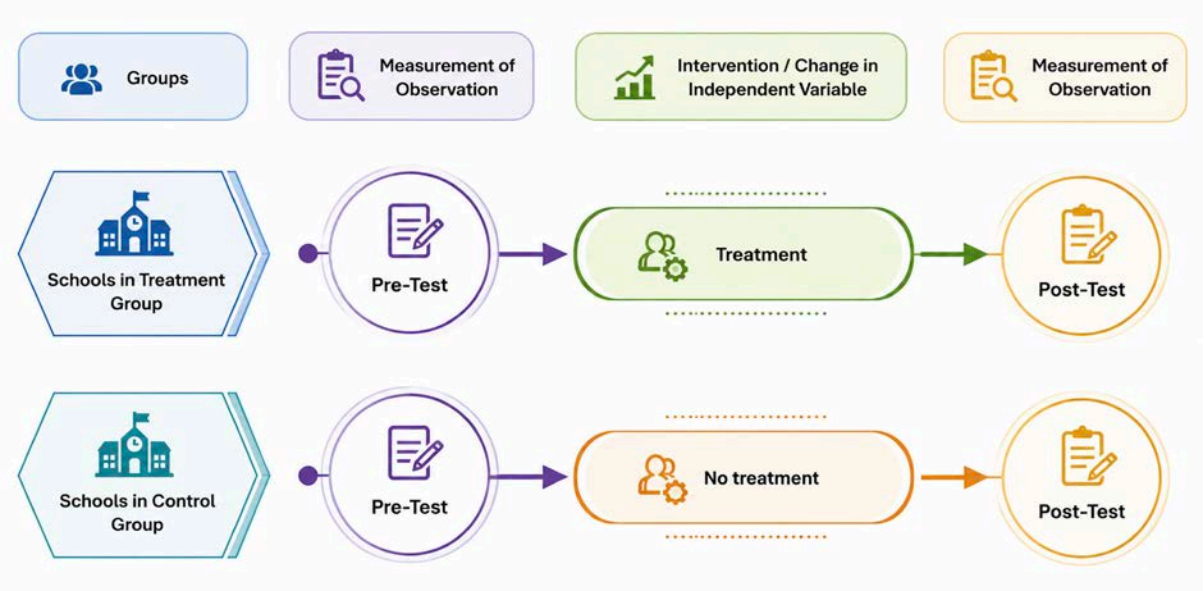


Figure 1: Evaluation Design of Baseline to Endline

3.2 Treatment and Control Groups

The evaluation sample consisted of 528 four-teacher public primary schools selected from six districts of Khyber Pakhtunkhwa. Schools included in the sample were randomly assigned into treatment and control groups, comprising 351 treatment schools and 177 control schools.

Schools assigned to the treatment group were linked with SLs and ASDEOs who received training on data-driven mentoring and coaching for multigrade classroom settings. These SLs and ASDEOs were trained on the use of the MG-COT framework and were expected to provide structured classroom observation, mentoring, coaching, and feedback support to teachers during routine school visits.

Schools assigned to the control group continued with regular academic monitoring mechanisms and did not receive the intervention during the evaluation period. The use of treatment and control groups allowed the evaluation to compare differences in classroom teaching practices while controlling for pre-existing baseline variations across schools.

To support implementation, 71 independent assessors were engaged in Khyber Pakhtunkhwa and trained to conduct classroom observations using the MG-COT at both baseline and endline.

Table 1: District-wise Distribution of Treatment and Control Schools

District	**Schools covered**	**Female schools**	**Male schools**
Charsadda	79	50	29
Haripur	71	35	36
Kohat	59	25	34
Nowshera	92	49	43
Peshawar	102	27	75
Swabi	125	65	60
TOTAL	**528**	**251**	**277**

3.3 Data Collection Approach

The evaluation relied primarily on classroom observations conducted using the MG-COT. Data collection was carried out at both baseline and endline stages across treatment and control schools.
Independent observers conducted classroom observations in sampled multigrade classrooms and assigned scores across 13 predefined teaching practices included within the MG-COT framework. Each practice was scored on a five-point scale, where a score of one represented the lowest level of implementation and a score of five represented the highest level of implementation.

The classroom observations focused on assessing practices related to classroom organisation, lesson planning, instructional delivery, differentiated instruction, time management, peer learning, formative assessment, and individualised feedback. Standardised observation protocols and data collection procedures were followed to ensure consistency and reliability across districts and school types.

3.4 Classroom Observation Framework

The MG-COT was specifically designed to assess the quality and implementation of teaching practices within multigrade classroom settings. The framework captures the following 13 core teaching practices considered essential for effective multigrade instruction:

i. *logical grade grouping;*

ii. *classroom space management;*

iii. *implementation of multigrade academic calendars;*

iv. *lesson planning practices;*

v. *articulation of learning objectives;*

vi. *alignment of instructional delivery with learning objectives;*

vii. *clarity and accuracy of content delivery;*

viii. *differentiated instruction;*

ix. *classroom time management;*

x. *flexible student grouping;*

xi. *promotion of peer learning;*

xii. *formative assessment practices; and*

xiii. *provision of individualised feedback.*

The use of a structured classroom observation framework enabled systematic assessment of teaching practices across schools and facilitated quantitative comparison between treatment and control groups.

3.5 Data Analysis Strategy

The analysis focused on assessing changes in multigrade teaching practices and estimating the impact of the intervention on MG-COT scores at endline.
Descriptive analysis was conducted to compare baseline and endline MG-COT scores across treatment and control groups, including overall scores, practice-wise scores, and school gender-wise comparisons.

To estimate the causal impact of the intervention, the evaluation employed the ANCOVA framework. The ANCOVA approach estimates differences in endline outcomes between treatment and control groups while controlling for baseline MG-COT scores. This approach improves the precision of treatment effect estimates and accounts for any pre-existing differences between groups.

3.5.1 ANCOVA Analytical Framework

The ANCOVA model estimated the impact of the intervention on endline MG-COT scores while controlling for baseline classroom practice scores. The primary outcome variable was the endline MG-COT score, measured either as an overall index or as individual practice-level scores.

The key explanatory variable was treatment assignment, indicating whether a school belonged to the intervention or control group. Baseline MG-COT scores were included in the model as covariates to adjust for pre-intervention differences across schools and classrooms.

3.5.2 Regression Specification

The following regression specification was used to estimate intervention effects:

$$\text{Endline Score}_i = \beta_0 + \beta_1(\text{Treatment}_i) + \beta_2\,(\text{Baseline Score}_i) + \epsilon i$$

Where,

- Endline Score$_i$ is the MG-COT score recorded for school *i* after the intervention, serving as the outcome variable.
- β_0 is the intercept; the predicted endline score for a control school with a baseline score of zero.
- β_1 is the ANCOVA treatment effect; the estimated difference in endline scores between treatment and control schools, after adjusting for baseline performance.
- Treatment$_i$ is a binary indicator equal to 1 if school *i* was assigned to the treatment group, and 0 if control.
- β_2 is the coefficient on the baseline score; captures the relationship between prior performance and endline outcomes, controlling for regression to the mean.
- Baseline Score$_i$ is the pre-intervention MG-COT score for school *i*, included as a covariate to improve precision and reduce residual variance.
- ϵ_i is the error term; represents unexplained variation in endline scores not accounted for by treatment status or baseline performance.

3.5.3 Interpretation of Effect Sizes

The estimated treatment coefficient from the ANCOVA models represents the adjusted difference in endline MG-COT scores between treatment and control groups after accounting for baseline scores. Positive and statistically significant coefficients indicate that the intervention contributed to improvements in multigrade teaching practices in treatment schools relative to control schools. The magnitude of the coefficients reflects the extent of improvement associated with the intervention across overall and practice-level MG-COT scores.

3.6. Sampling Strategy and Endline Coverage

3.6.1 Sample Selection

The sampling strategy for the evaluation was designed to ensure representation across selected districts of Khyber Pakhtunkhwa while focusing specifically on four-teacher public primary schools operating in multigrade settings. Schools included in the evaluation were selected from districts representing different geographical and administrative contexts of the province.

The evaluation sample consisted of treatment and control schools identified through a random assignment process. The intervention targeted schools linked with SLs and ASDEOs who received training on data-driven mentoring and coaching for multigrade classroom settings.

Both baseline and endline classroom observations were conducted within the sampled schools using the MG-COT.

3.6.2 School Coverage at Endline

Overall, classroom observations were conducted across 528 public primary schools included in the evaluation sample. Of these, 251 schools (**48%**) were female schools, while 277 schools (**52%**) were male schools.

Figure 2 presents the overall gender-wise distribution of schools covered under the evaluation.

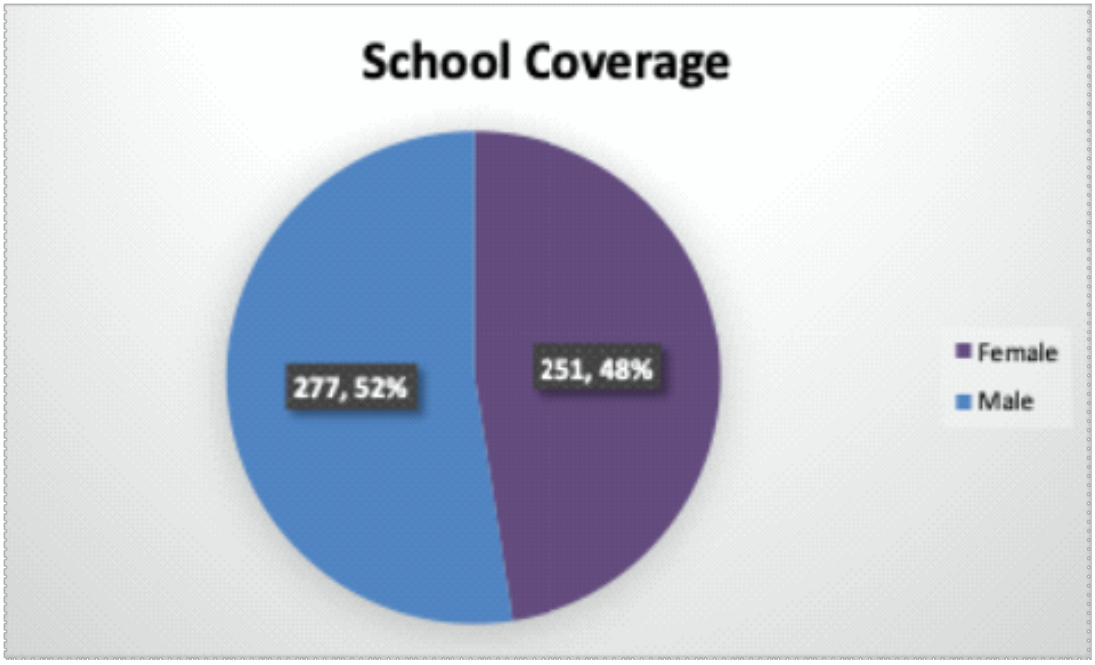


Figure 2: School Coverage by Gender

District-wise distribution of schools covered in the evaluation is presented in Figure 3. The highest number of schools were covered in Swabi (125), followed by Peshawar (102) and Nowshera (92). Comparatively fewer schools were covered in Kohat (59) and Haripur (71).

Female school coverage was comparatively higher in districts such as Charsadda, Nowshera, and Swabi, while male school coverage was relatively higher in Peshawar and Kohat.

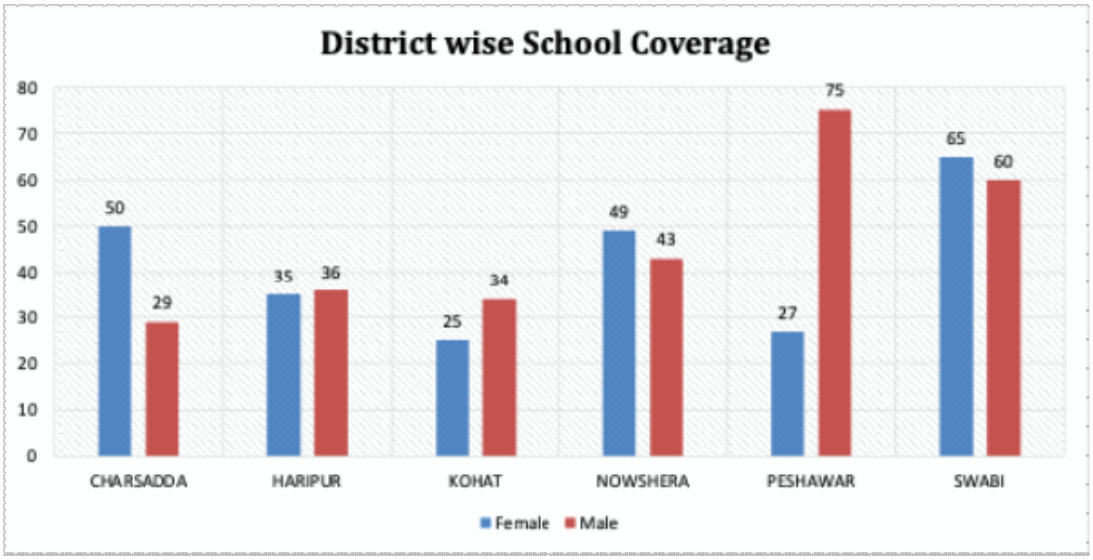


Figure 3: District-wise School Coverage by Gender

3.6.3 Observation Coverage at Endline

A total of 904 teacher classroom observations were conducted across sampled schools using the MG-COT framework. The observations were equally distributed across school gender categories, with 452 observations (**50%**) conducted in female schools and 452 observations (**50%**) conducted in male schools.

The teacher observations refer to classroom sessions observed during data collection and represent teachers observed within male and female public primary schools included in the evaluation.

Figure 4 presents the overall distribution of teacher observations by school gender.

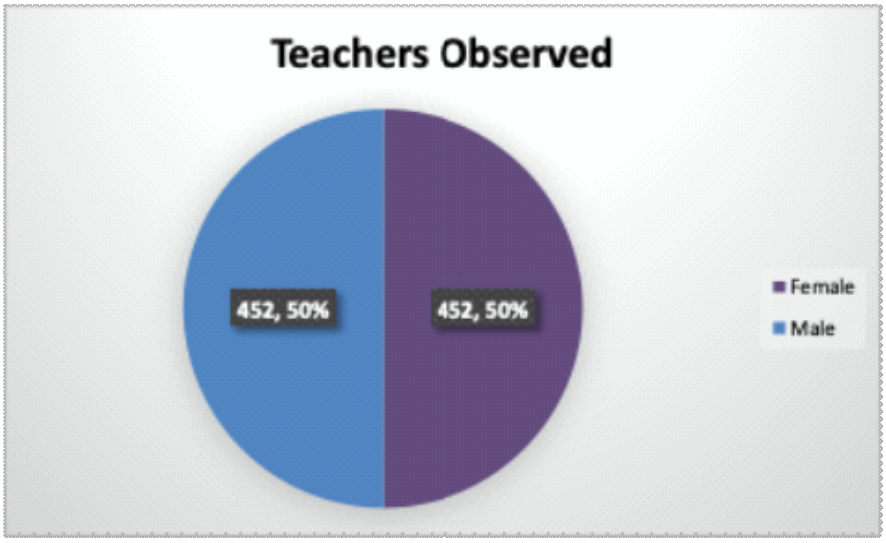


Figure 4: Teachers Observed by School Gender

District-wise distribution of teacher observations is presented in Figure 5. The highest number of teacher observations were conducted in Peshawar (234), followed by Swabi (191) and Nowshera (160). Comparatively fewer observations were conducted in Kohat (76) and Haripur (109).

Female school observations were comparatively higher in districts such as Swabi, Nowshera, and Charsadda, while observations in male schools were notably higher in Peshawar.

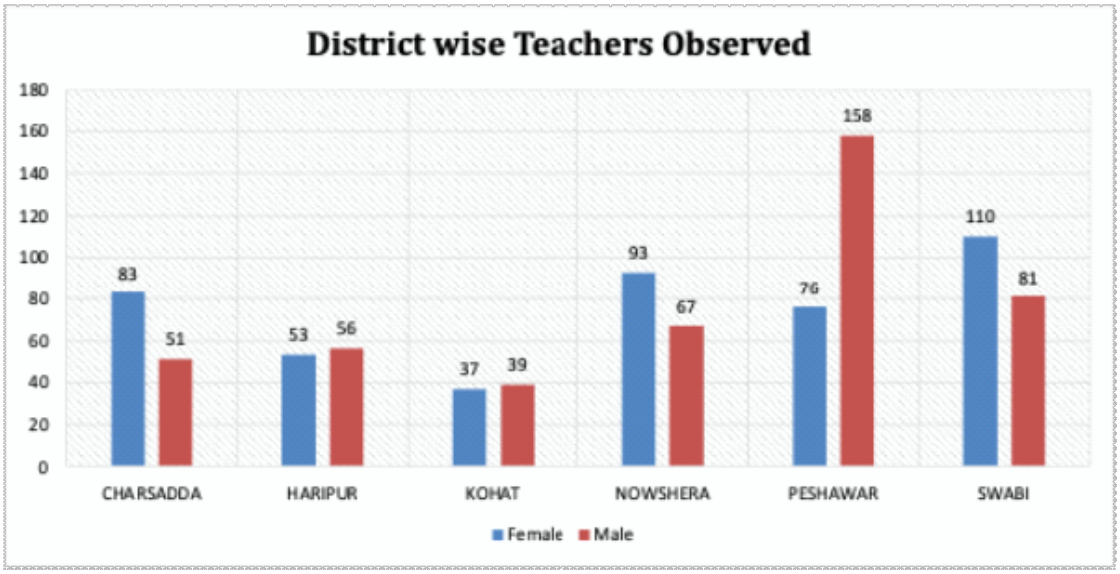


Figure 5: District-wise Teachers Observed by School Gender

4. Implementation Steps

4.1 Content Development and Review

A comprehensive set of training materials was developed to support data-driven coaching and mentoring using MG-COT. The material draws on the 13-indicator MG-COT framework and translates coaching principles into practical simulation-based routines suited to the supervisory and mentoring roles of SLs and ASDEOs. The MG-COT application demonstration formed part of the training module. A repository of instructional content mapped against all 13 MG-COT practices, including lesson plans, videos, and reading material, was also developed. The training package included a trainers' module, trainees' manuals, and a repository of resources. The training material and modules were reviewed by DPD from 22-24 December 2025.

4.2 Training of ASDEOs

Training on data-driven coaching and mentoring informed by MG-COT was delivered for 70 Assistant Sub-Divisional Education Officers (ASDEOs) including 33 females and 37 males at the DPD, Peshawar. The training was jointly facilitated by DPD and the B-TAG team and focused on applying MG-COT data for instructional supervision, structured feedback, and coaching in multigrade classroom settings. Participants were drawn from six districts (Charsadda, Haripur, Kohat, Nowshera, Peshawar, and Swabi), ensuring representation across diverse administrative and instructional contexts.

The training demonstrated strong learning outcomes, with average assessment scores increasing from 71.7 percent in the pre-assessment to 95.2 percent in the post-assessment, reflecting a net gain of 23.5 percentage points. Overall participant satisfaction was high, with an average rating of 4.51 out of 5.0, indicating strong perceived relevance and effectiveness of the training.

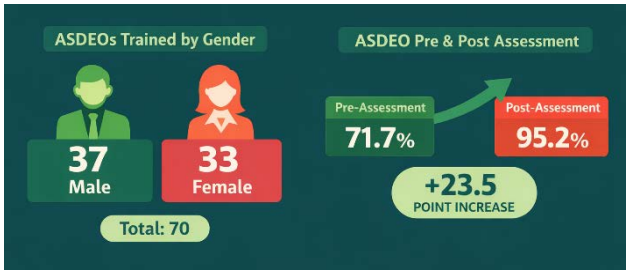


Figure 6: Pre- and Post-Assessment Scores - ASDEO Training

Table 2: ASDEOs Trained by District

District	Female	Male	Grand Total
Charsadda	9	4	13
Haripur	5	6	11
Kohat	2	4	6
Nowshera	6	3	9
Peshawar	4	10	14
Swabi	7	10	17
Grand Total	**33**	**37**	**70**

Figure 7: Training session of the ASDEOs in progress (DPD Peshawar)

4.3 Training of School Leaders

Building on the training of 70 ASDEOs at DPD in the previous quarter, 126 School Leaders (67 male, 59 female) were trained in February 2026. The training was jointly facilitated by DPD and B-TAG, focusing on applying MG-COT data for instructional supervision, structured feedback, and coaching in multigrade classroom settings.

Pre- and post-assessment results show an average increase of 14.25 percentage points in participant scores, improving from a baseline of 72.20 percent to 86.45 percent. Participant feedback was positive, with an overall satisfaction mean of 4.38 out of 5.0.

Table 3: School Leaders Trained by District

District	**SLs trained**	**Male**	**Female**
Charsadda	20	6	14
Haripur	6	0	6
Kohat	34	18	16
Nowshera	5	0	5
Peshawar	34	23	11
Swabi	27	20	7
Grand Total	**126**	**67**	**59**

Figure 8: Training session of the School Leaders in progress (DPD Peshawar)

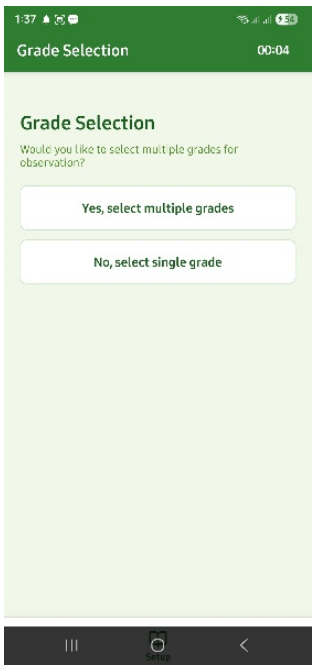
4.4 Ongoing Coaching and Mentoring

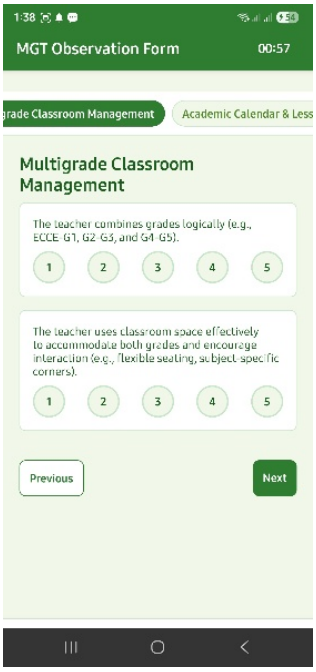
Following the adaptation of the MG-COT application and the training of ASDEOs and School Leaders, routine operational use of MG-COT data within the classroom observation process began. A comprehensive Monitoring and Evaluation (M&E) framework was developed to assess whether MG-COT-informed training translates into improved supervisory practice and classroom instruction.

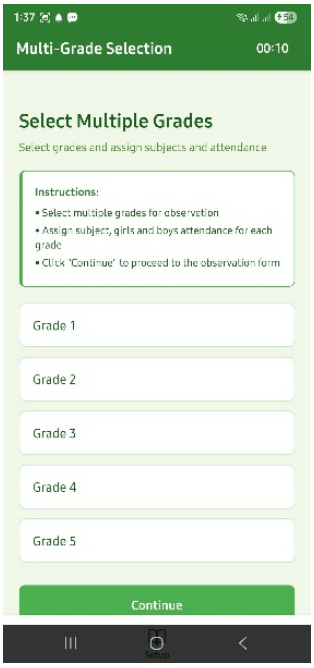
5. Digital Classroom Observation Tool

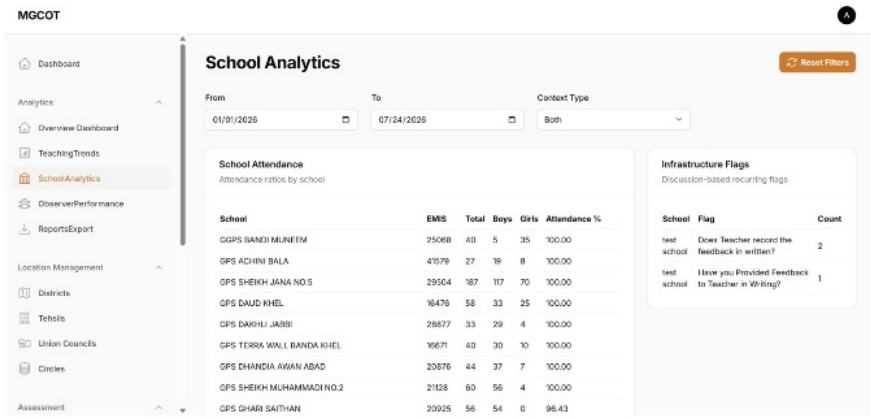
5.1 MG-COT Application

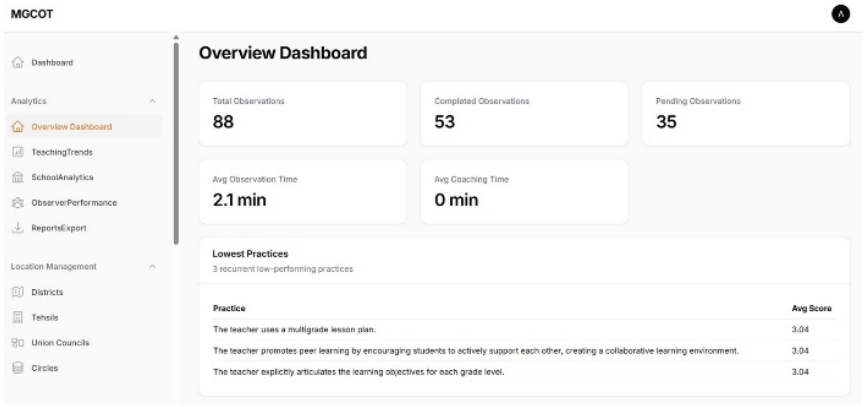
In KP Multigrade application was developed for classroom observation with support from B-TAG. The existing COT application was updated to include MG-COT, allowing users to select between monograde and multigrade observations, with 13 instructional practices for multigrade settings. The system generates real-time analysis, tracks teacher performance trends across practices, identifies targeted areas for improvement, integrates aligned instructional content, and records feedback sessions and agreed action points between coaches and teachers. The application was hosted on KP Data center. User IDs were created for ASDEOs and SL to observe classroom practices. The application will be adopted in selected district and will be later scaled up to all districts. The Directorate of Curriculum & Teacher Education (DCTE) conducted a review workshop from 22/12/2025 to 25/12/2025 to review the MG-COT tool and approved on 7-01-2025.








MGCOT
Dashboard
Analytics
Overview Dashboard
TeachingTrends
SchoolAnalytics
ObserverPerformance
ReportsExport
Location Management
Districts
Tehsils
Union Councils
Circles
Assessment
School Analytics
Reset Filters
From
01/01/2026
To
07/24/2026
Context Type
Both
School Attendance
Attendance ratios by school
School EMIS Total Boys Girls Attendance %
GGPS BANDI MUNEEM 25068 40 5 35 100.00
GPS ACHINI BALA 41579 27 19 8 100.00
GPS SHEIKH JANA NO.5 29504 187 117 70 100.00
GPS DAUD KHEL 16476 58 33 25 100.00
GPS DAKHLI JABBI 28877 33 29 4 100.00
GPS TERRA WALL BANDA KHEL 16671 40 30 10 100.00
GPS DHANDIA AWAN ABAD 20876 44 37 7 100.00
GPS SHEIKH MUHAMMADI NO.2 21128 60 56 4 100.00
GPS GHARI SAITHAN 20925 56 54 0 98.43
Infrastructure Flags
Discussion-based recurring flags
School Flag Count
test school Does Teacher record the feedback in written? 2
test school I have you Provided Feedback to Teacher in Writing? 1


MGCOT
Dashboard
Analytics
Overview Dashboard
TeachingTrends
SchoolAnalytics
ObserverPerformance
ReportsExport
Location Management
Districts
Tehsils
Union Councils
Circles
Overview Dashboard
Total Observations
88
Completed Observations
53
Pending Observations
35
Avg Observation Time
2.1 min
Avg Coaching Time
0 min
Lowest Practices
3 recurrent low-performing practices
Practice Avg Score
The teacher uses a multigrade lesson plan. 3.04
The teacher promotes peer learning by encouraging students to actively support each other, creating a collaborative learning environment. 3.04
The teacher explicitly articulates the learning objectives for each grade level. 3.04

6. Baseline and Endline Status of MG-COT Teaching Practices

6.1 Overall MG-COT Index: Baseline vs Endline

The classroom observations conducted using the MG-COT indicate substantial improvement in multigrade teaching practices between baseline and endline across both treatment and control schools. However, the magnitude of improvement was considerably higher in treatment schools compared to control schools.

At baseline, the overall average MG-COT score in treatment schools was 1.72, compared to 1.61 in control schools, indicating that both groups started from broadly comparable levels of classroom practice. At endline, the average MG-COT score increased sharply to 3.70 in treatment schools and 2.09 in control schools.

The findings show that:

- treatment schools improved by 1.98 points between baseline and endline; while
- control schools improved by 0.48 points during the same period.

The findings suggest that schools supported through trained School Leaders and ASDEOs experienced substantially greater strengthening in multigrade classroom teaching practices relative to schools operating under regular academic supervision mechanisms.

Figure 9 presents the comparison of overall baseline and endline MG-COT scores across treatment and control schools.

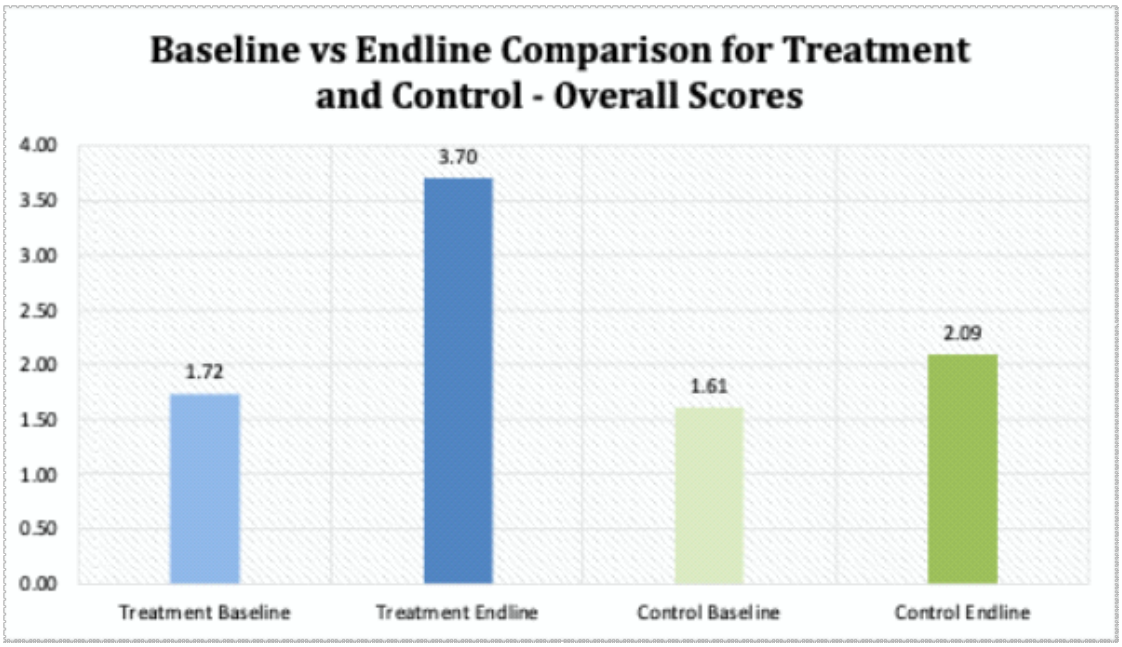


Figure 9: Baseline vs Endline Comparison for Treatment and Control — Overall Scores

Figure 10 presents the overall change in MG-COT scores between baseline and endline across treatment and control groups.

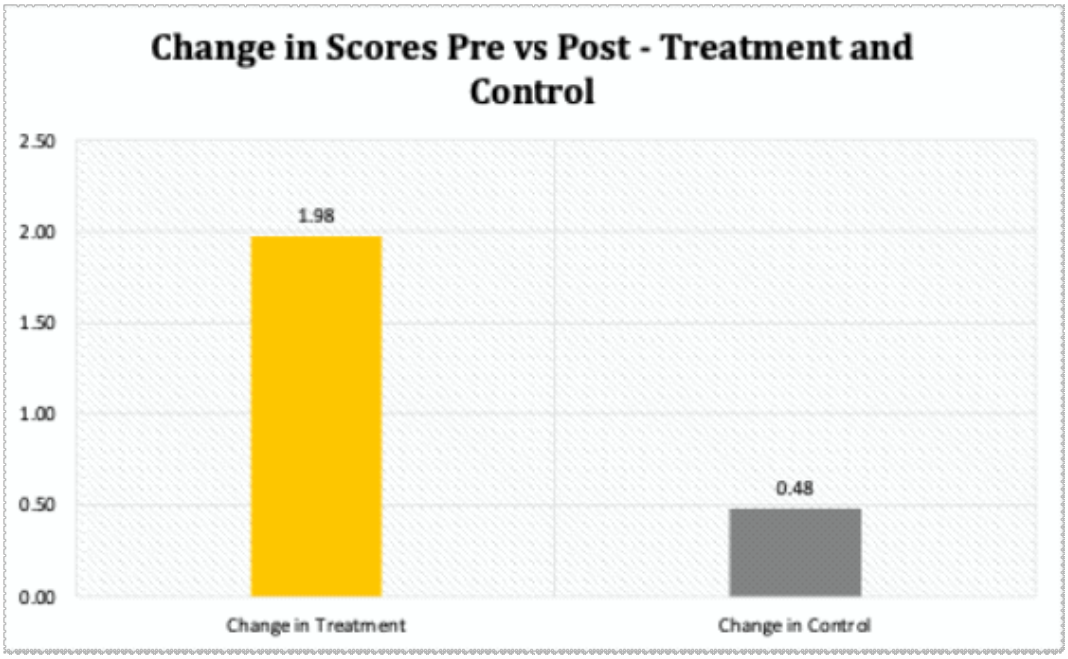


Figure 10: Change in Scores Pre vs Post — Treatment and Control

It is important to note that the intervention was implemented within the regular academic cycle; therefore, it is expected that teachers in both groups would improve their classroom practices through regular teaching experience, school-based support, supervision and mentorship, and other professional development opportunities, such as trainings.

Consequently, the improvements observed in the control group do not imply the absence of programme impact. Rather, they reflect the changes in teaching practices that would be expected even without the intervention as teachers continue to gain experience and benefit from routine support mechanisms available within the education system.

The purpose of the impact evaluation is therefore not to determine whether teaching practices improved over time, but rather the intervention produced **additional improvements beyond those that would have occurred in its absence**. The analysis is therefore designed to isolate the incremental effect of the intervention after accounting for the improvements that would be expected under normal conditions.

Figure 10 represents unadjusted changes over time and does not account for any differences in baseline teaching practice scores between the two groups. To estimate the intervention's independent contribution, an Analysis of Covariance (ANCOVA) was conducted using endline teaching practice scores as the outcome variable while controlling for baseline scores and clustering standard errors at the school level.

The ANCOVA results indicate that, after adjusting for baseline performance, the intervention had a positive and statistically significant effect on endline teaching practice scores ($\beta = 1.626$, $p < 0.001$). This indicates that teachers in treatment schools achieved endline teaching practice scores that were, on average, 1.626 points higher than those in comparable control schools after accounting for their initial performance.

The observed improvement in the control group is consistent with the natural strengthening of teaching practices that can occur during an academic year through routine teaching experience, ongoing school support, supervisory feedback, and other contextual influences.

The combined evidence from the ANCOVA and the descriptive comparison demonstrates that the intervention produced additional improvements beyond those that would have occurred in its absence. While both groups exhibited improvements in teaching practices over the study period, participation in the intervention resulted in a large and statistically significant additional improvement in teaching practices.

6.2 Practice-wise MG-COT Scores: Baseline vs Endline

Practice-wise analysis further shows improvement across all 13 MG-COT teaching practices for both treatment and control schools between baseline and endline. However, the magnitude of improvement was consistently and substantially higher in treatment schools across every observed classroom practice.

Among treatment schools, the largest improvements were observed in:

- use of multigrade lesson plans (2.08 points);
- effective classroom space use (2.02 points);
- content aligned with learning objectives (2.01 points);
- flexible student grouping (2.01 points); and
- use of multigrade academic calendars (1.99 points).

Similarly, strong improvements were also observed in:

- clear learning objectives;
- balanced instructional time across grades;
- differentiated instruction by grade; and
- logical grade grouping.

Comparatively, improvements in control schools remained modest, ranging between 0.41 and 0.52 points across the 13 practices.

The analysis indicates that teachers in treatment schools adopted substantially stronger multigrade instructional practices over the intervention period, particularly in areas related to classroom organisation, lesson planning, instructional alignment, student grouping, and learner-centred teaching approaches.

Figure 11 presents the baseline and endline comparison of practice-wise MG-COT scores across treatment and control schools.

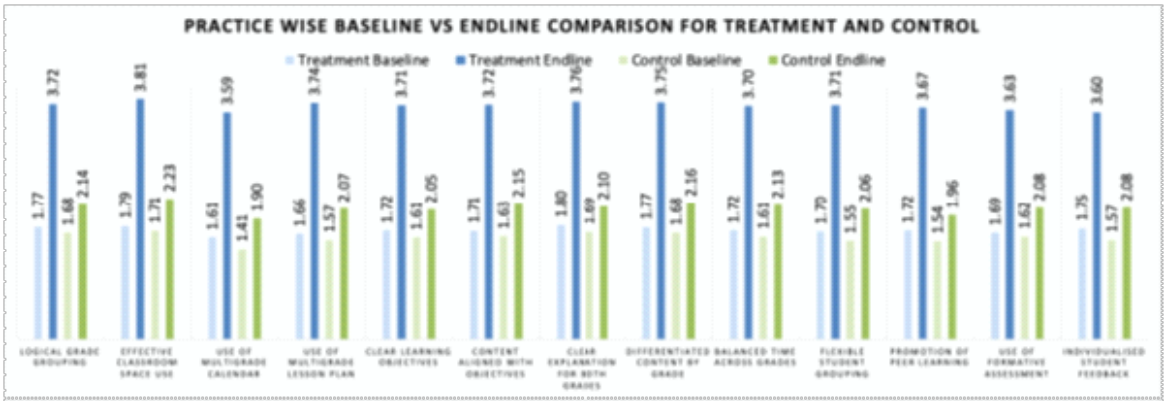


Figure 11: Practice-wise Baseline vs Endline Comparison for Treatment and Control

Figure 12 presents the change in practice-wise scores between baseline and endline for treatment and control schools.

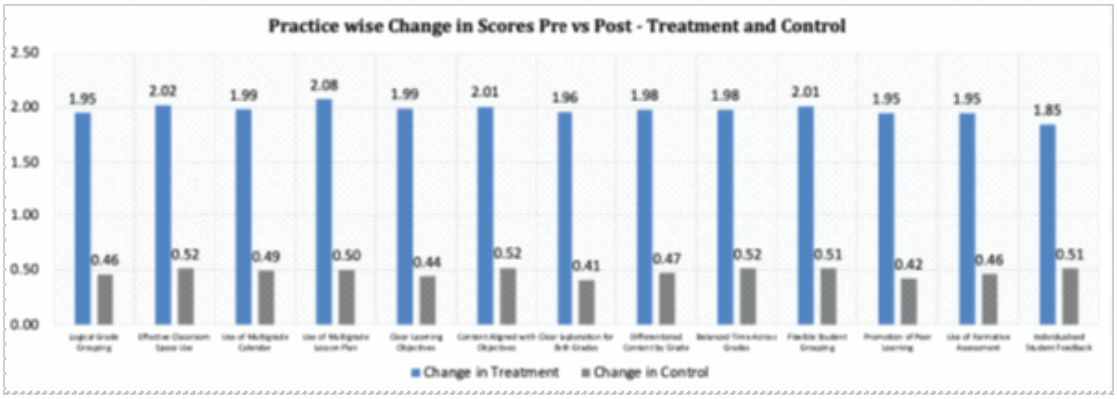


Figure 12: Practice-wise Change in Scores Pre vs Post — Treatment and Control

6.3 School Gender-wise MG-COT Scores: Baseline vs Endline

6.3.1 Female Schools

The comparison of MG-COT scores for female schools shows substantial improvement in classroom teaching practices between baseline and endline for both treatment and control schools. However, improvements in female treatment schools were considerably larger.

At baseline, the average MG-COT score for female treatment schools was 1.66, compared with 1.64 for female control schools. At endline, the average score increased to 3.72 in treatment schools and 2.04 in control schools.

The overall improvement observed was:

- 2.06 points in female treatment schools; and
- 0.40 points in female control schools.

The findings suggest that female schools linked with trained School Leaders and ASDEOs experienced markedly stronger strengthening of classroom instructional practices over the intervention period.

Figure 13 presents the comparison of baseline and endline MG-COT scores across female treatment and control schools.

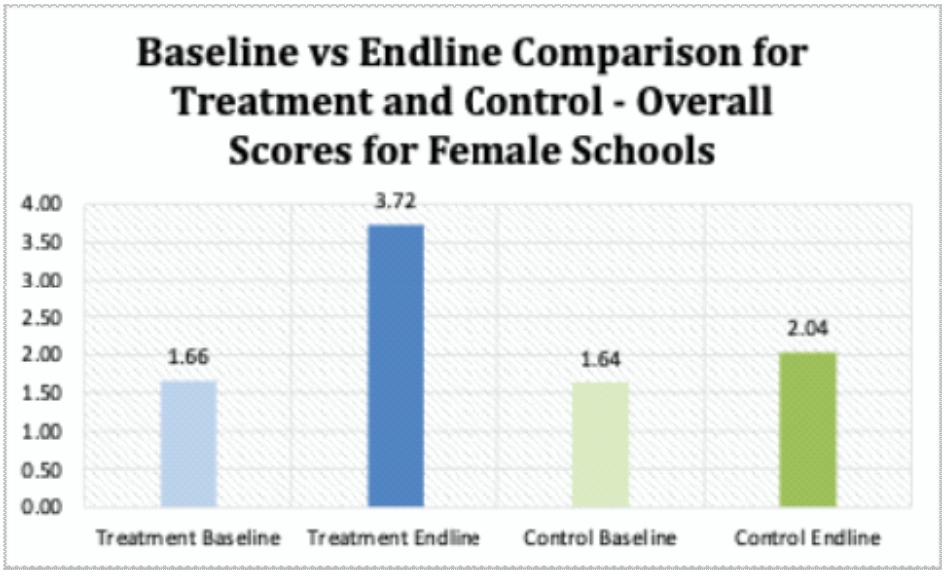


Figure 13: Baseline vs Endline Comparison for Treatment and Control — Female Schools

Figure 14 presents the change in scores between baseline and endline for female treatment and control schools.

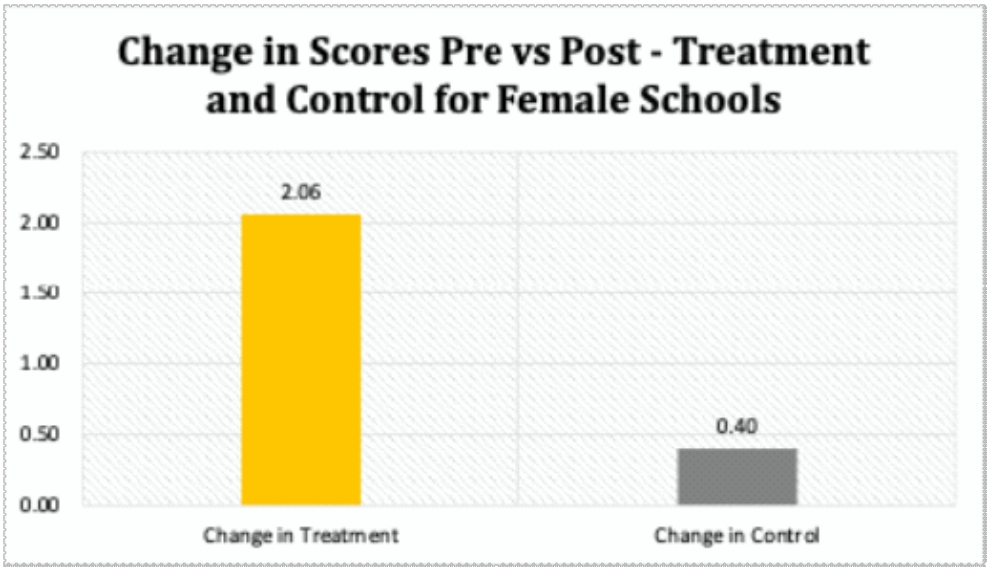


Figure 14: Change in Scores Pre vs Post — Treatment and Control for Female Schools

The stronger improvements observed among female treatment schools indicate notable gains in classroom organisation, instructional delivery, differentiated teaching, and student engagement practices relative to female schools in the control group.

6.3.2 Male Schools

A similar pattern was observed across male schools, where both treatment and control schools demonstrated improvement in MG-COT scores between baseline and endline, although treatment schools again recorded substantially larger gains.

At baseline, the average MG-COT score among male treatment schools was 1.79, compared with 1.58 for male control schools. By endline, the average score increased to 3.68 for male treatment schools and 2.12 for male control schools.

The overall improvement observed was:

- 1.89 points in male treatment schools; and
- 0.54 points in male control schools.

These findings indicate that the intervention also generated substantial strengthening of multigrade teaching practices within male schools associated with trained School Leaders and ASDEOs.

Figure 15 presents the comparison of baseline and endline MG-COT scores across male treatment and control schools.

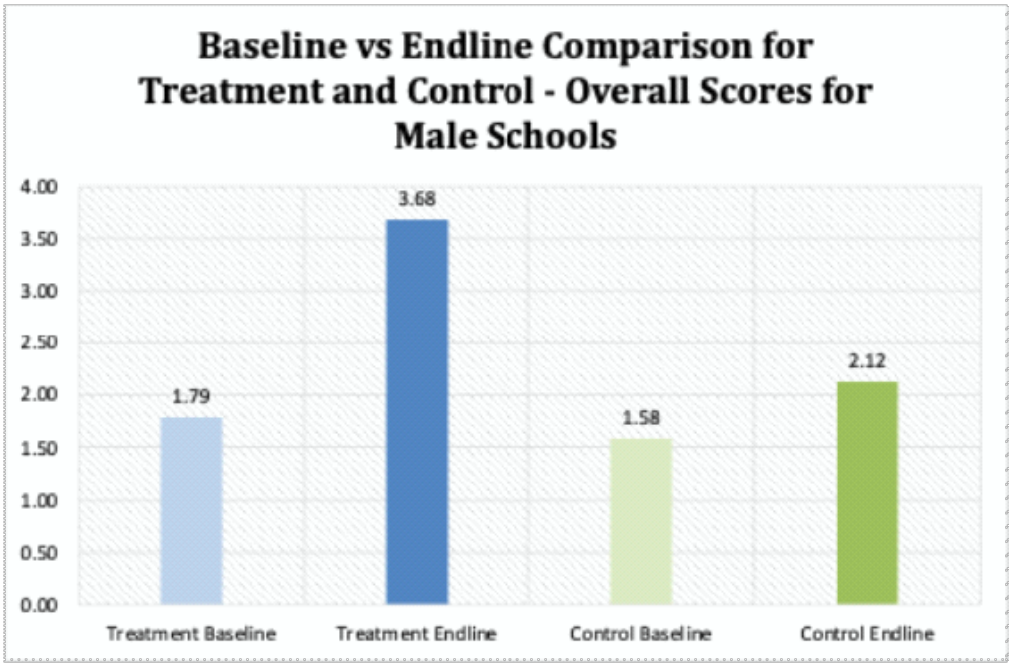


Figure 15: Baseline vs Endline Comparison for Treatment and Control — Male Schools

Figure 16 presents the change in scores between baseline and endline for male treatment and control schools.

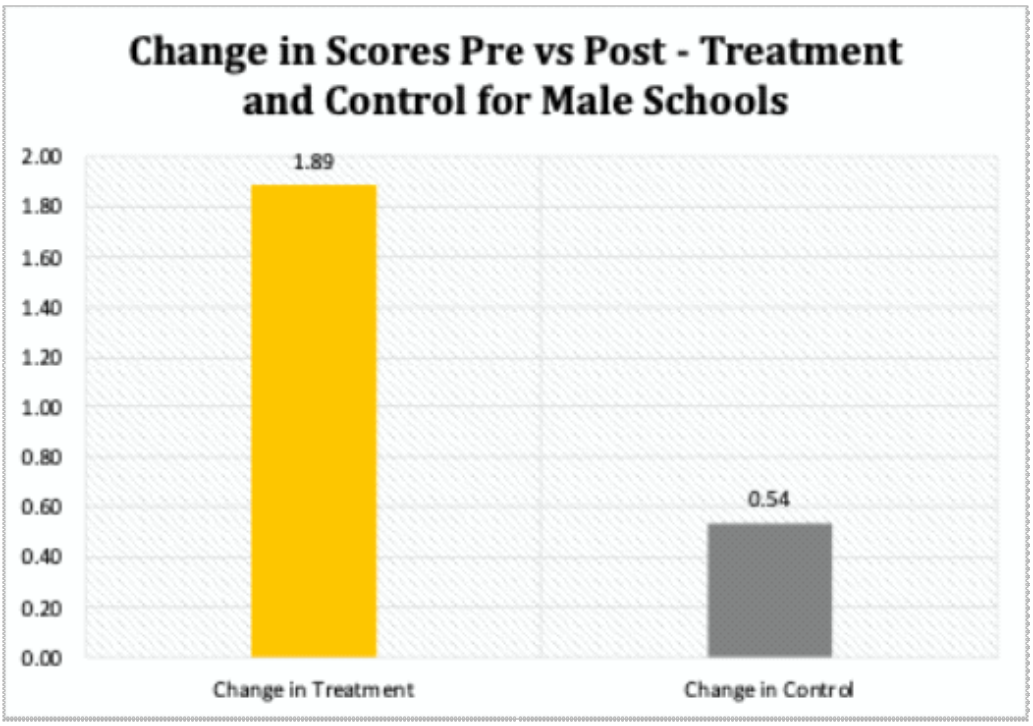


Figure 16: Change in Scores Pre vs Post — Treatment and Control for Male Schools

Overall, the findings across both female and male schools consistently demonstrate stronger improvements in multigrade classroom practices among treatment schools relative to control schools. The descriptive analysis suggests that the intervention contributed positively towards improving instructional quality, classroom organisation, lesson planning, and learner engagement practices within multigrade public primary schools across Khyber Pakhtunkhwa.

7. Impact of the Intervention: ANCOVA Results

7.1 Analytical Approach

To estimate the impact of the intervention on multigrade teaching practices, the evaluation employed the ANCOVA framework. The ANCOVA model estimated differences in endline MG-COT scores between treatment and control schools while controlling for baseline classroom practice scores.

The analysis focused on:

- the overall MG-COT index (average score); and
- subgroup analysis by school gender.

Standard errors were clustered at the school level using EMIS codes to account for correlation among classroom observations conducted within the same schools.

7.2 Impact on Overall MG-COT Index

The ANCOVA results show that the intervention had a large and statistically significant positive impact on overall MG-COT scores in treatment schools.

After controlling for baseline scores, treatment schools scored, on average, 1.626 points higher at endline compared with control schools. The estimated treatment effect was statistically significant at the 1 percent level ($p < 0.001$).

The results indicate that schools supported through School Leaders and ASDEOs who received training on data-driven mentoring and coaching demonstrated substantially stronger multigrade teaching practices at endline relative to schools operating under regular academic supervision arrangements.

Unlike the Punjab analysis, the baseline MG-COT score was not found to be a statistically significant predictor of endline classroom practice scores in the overall KP model, indicating that the intervention effect remained strong even after adjusting for baseline differences.

Table 4 presents the ANCOVA results for the overall MG-COT index.

```
. reg SCORE treat school_baseline if time==1 & PROVINCE=="KP", vce(cluster EMIS)

Linear regression                               Number of obs     =        618
                                                F(2, 512)         =     312.07
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6025
                                                Root MSE          =      .6249

                                    (Std. Err. adjusted for 513 clusters in EMIS)
```

SCORE	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
treat	1.62636	.0667036	24.38	0.000	1.495314	1.757407
school_baseline	-.0020433	.0648551	-0.03	0.975	-.1294582	.1253717
_cons	2.09214	.11036	18.96	0.000	1.875326	2.308954

Table 4: ANCOVA Results — Overall MG-COT Index

The magnitude of the treatment coefficient indicates substantial improvement in multigrade classroom instructional practices among treatment schools associated with trained School Leaders and ASDEOs.

7.3 Gender-wise Impact Analysis

7.3.1 Female Schools

The ANCOVA analysis for female schools shows that the intervention had a strong and statistically significant effect on MG-COT scores among female schools.
After adjusting for baseline scores, female treatment schools scored 1.691 points higher at endline compared with female control schools. The estimated treatment effect was statistically significant at the 1 percent level ($p < 0.001$).

The findings indicate that mentoring and coaching support provided through trained School Leaders and ASDEOs contributed to substantial improvements in multigrade classroom practices within female schools.

In the female school model, the baseline MG-COT score was statistically significant and negatively associated with endline scores, suggesting that female schools with relatively lower baseline scores experienced comparatively larger gains during the intervention period.

Table 5 presents the ANCOVA results for female schools.

Linear regression

Number of obs	= 305
F(2, 245)	= 172.31
Prob > F	= 0.0000
R-squared	= 0.6587
Root MSE	= .56235

(Std. Err. adjusted for 246 clusters in EMIS)

SCORE	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
treat	1.691134	.092817	18.22	0.000	1.508313	1.873955
school_baseline	-.2212491	.0714962	-3.09	0.002	-.3620746	-.0804235
_cons	2.41411	.1407936	17.15	0.000	2.13679	2.691431

Table 5: ANCOVA Results — Female Schools

The results suggest that the intervention was highly effective in improving classroom organisation, instructional delivery, differentiated teaching, and student engagement practices within female multigrade schools.

7.3.2 Male Schools

The ANCOVA findings for male schools also demonstrate a strong and statistically significant positive impact of the intervention on MG-COT scores.

After controlling for baseline scores, male treatment schools scored 1.512 points higher at endline compared with male control schools. The treatment effect was statistically significant at the 1 percent level ($p < 0.001$).

The magnitude of improvement observed among male schools further confirms that the intervention substantially strengthened multigrade classroom teaching practices across different school categories.

Unlike female schools, the baseline MG-COT score for male schools was positively associated with endline scores, indicating that schools with relatively stronger baseline instructional practices also tended to maintain stronger classroom performance at endline.

Table 6 presents the ANCOVA results for male schools.

Linear regression

Number of obs	**313**
F(2, 266)	**181.64**
Prob > F	**0.0000**
R-squared	**0.5864**
Root MSE	**.65475**

(Std. Err. adjusted for **267** clusters in EMIS)

SCORE	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
treat	**1.511552**	**.0916429**	**16.49**	**0.000**	**1.331114**	**1.69199**
school_baseline	**.3057844**	**.0899284**	**3.40**	**0.001**	**.1287223**	**.4828466**
_cons	**1.632143**	**.1541286**	**10.59**	**0.000**	**1.328675**	**1.93561**

Table 6: ANCOVA Results — Male Schools

Overall, the ANCOVA findings across both female and male schools consistently demonstrate that the intervention contributed significantly towards improving multigrade classroom teaching practices in treatment schools relative to control schools. The large and statistically significant treatment effects provide strong evidence that strengthening the mentoring and coaching capacity of School Leaders and ASDEOs can substantially improve instructional quality within multigrade public primary school settings across Khyber Pakhtunkhwa.

8. Impact on Student Learning

8.1 Results from School Based Assessment

The SBA analysis provides useful insights into changes in student performance between the two assessment periods across treatment and control schools. However, the findings should be interpreted in light of several contextual and implementation factors.

First, the intervention was designed primarily to strengthen instructional coaching and mentoring practices rather than directly improve student learning outcomes within a single assessment cycle. Its principal objective was to enhance the quality of classroom teaching by enabling School Leaders and ASDEOs to provide more structured, evidence-informed feedback to teachers using MG-COT data.

Second, School Leaders and ASDEOs were unable to use the planned real-time classroom observation application to identify teaching trends and provide data-driven mentoring during the intervention period. The intervention therefore assessed the effects of refresher training alone rather than the complete data-driven coaching model originally envisaged.

Third, the intervention period was relatively short. The complete data-driven mentoring model could not be implemented because the application was delayed. Both treatment and control schools continued to receive routine government support and had previously benefited from B-TAG's multigrade work. Student learning was measured using routine SBA data rather than a purpose-designed impact assessment.

These factors mean that differences in SBA performance should be interpreted cautiously and should not be viewed as a direct measure of the effectiveness of the refresher training.

The analysis presented in this section is limited to a descriptive pre- and post-assessment of students in treatment schools and is intended to provide contextual information on student performance across the two assessment periods. It should not be interpreted as an estimate of the causal impact of the intervention on student learning.

Furthermore, the learning outcomes were drawn from routine SBA data rather than from a purpose-designed assessment aligned to the evaluation's impact objectives. As a result, the observed student performance provides useful complementary information on learning outcomes during the evaluation period, but it does not present a definitive attribution of changes in student performance to the intervention. The classroom observation results therefore remain the principal evidence for assessing programme effectiveness, while the SBA findings should be considered supporting contextual evidence.

8.2 Semester 1

The overall subject-wise results for Semester-1 indicate that the average score was highest in **Mathematics (58.43%)**, followed by **General Knowledge (56.86%)**, **English (53.93%)**, and **Science (52.44%)**. Mathematics therefore recorded the highest average performance among the four subjects, while Science recorded the lowest. The subject-wise distribution is presented in Figure 17.

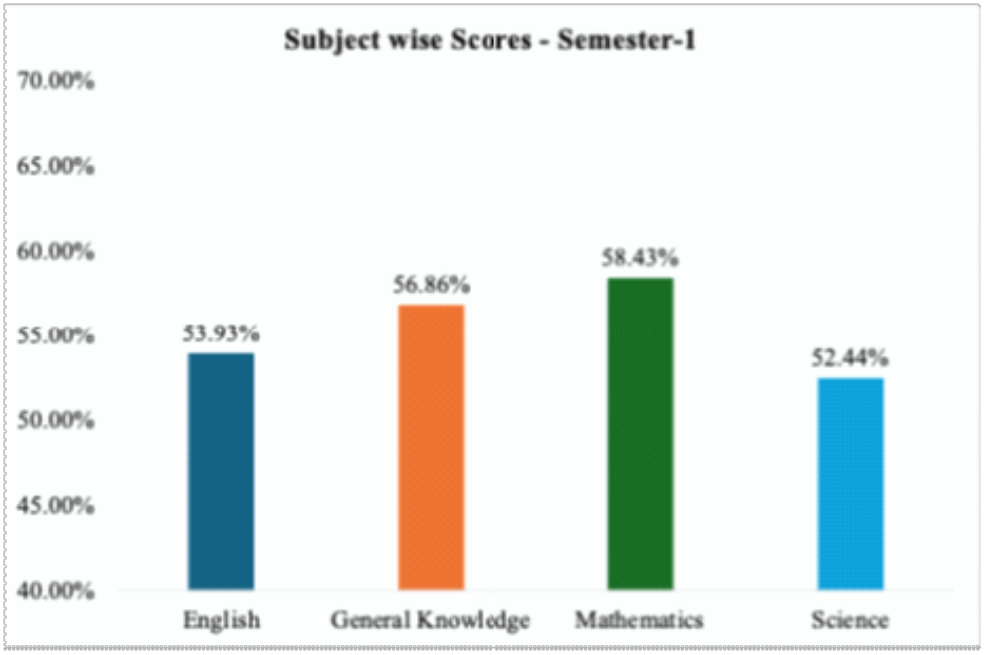


Figure 17: Subject-wise Scores - Semester-1

At district level, the results show variation in performance across subjects. In Charsadda, average scores were 60.3% in English, 63.3% in General Knowledge, 64.6% in Mathematics, and 59.9% in Science. In Haripur, the corresponding scores were 62.6%, 60.9%, 63.2%, and 56.1%, respectively. Kohat recorded average scores of 50.5% in English, 55.1% in General Knowledge, 55.3% in Mathematics, and 50.0% in Science.

In Nowshera, the average scores were 60.0% in English, 62.6% in General Knowledge, 65.3% in Mathematics, and 60.5% in Science. Peshawar recorded 53.1% in English, 57.2% in General Knowledge, 55.8% in Mathematics, and 50.0% in Science, while Swabi recorded 44.2% in English, 48.5% in General Knowledge, 51.7% in Mathematics, and 44.2% in Science.

Across the district-level results, the highest observed score was recorded for Mathematics in Nowshera (65.3%), while the lowest scores were 44.2% for English and Science in Swabi. The district-wise results are presented in Figure 18.

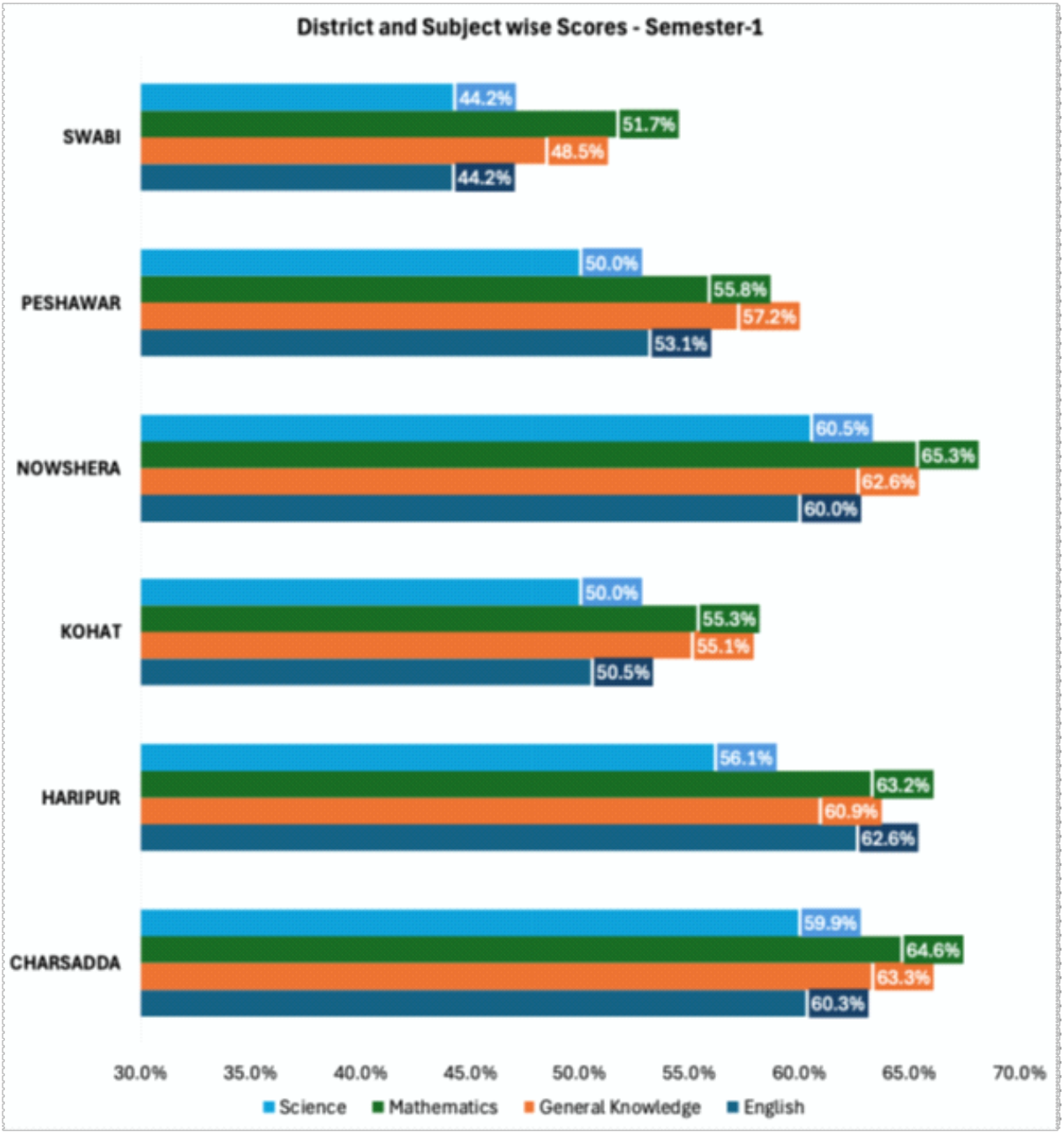


Figure 18: District and Subject-wise Scores - Semester-1

8.3 Semester 2

For Semester-2, the overall subject-wise results show average scores of 64.72% in English, 63.25% in General Knowledge, 63.83% in Mathematics, and 60.10% in Science. English recorded the highest average score among the four subjects, while Science recorded the lowest. The overall subject-wise results are presented in Figure 19.

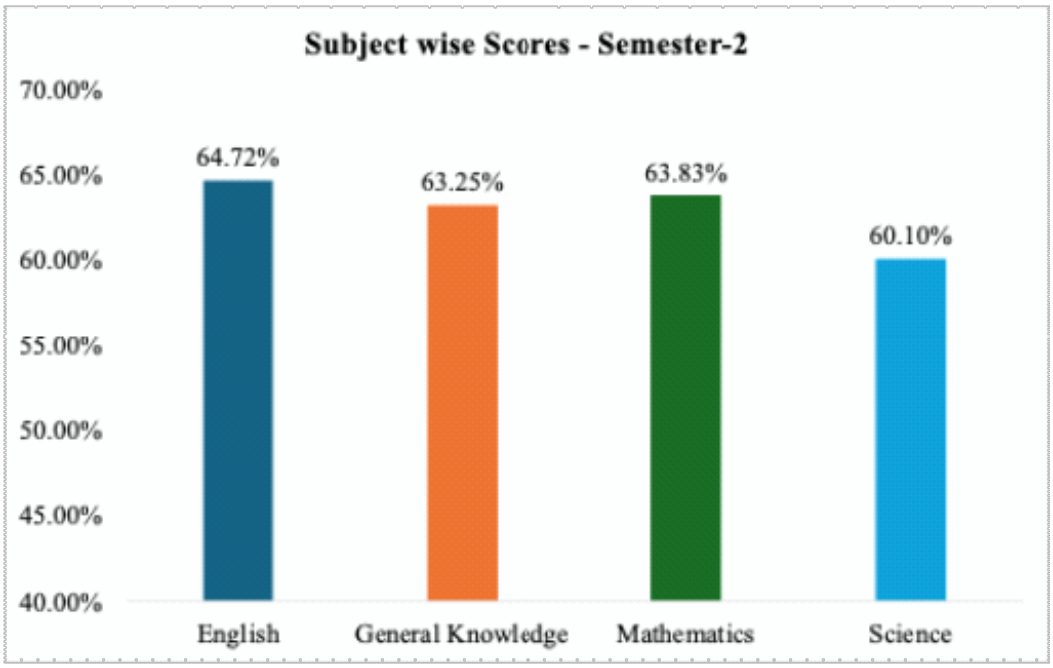


Figure 19: Subject-wise Scores - Semester-2

The district-level results also show variation across subjects. In Charsadda, average scores were 71.7% in English, 65.6% in General Knowledge, 71.6% in Mathematics, and 68.7% in Science. In Haripur, the corresponding scores were 68.3%, 67.5%, 66.5%, and 63.1%, respectively. Kohat recorded average scores of 61.5% in English, 58.1% in General Knowledge, 59.6% in Mathematics, and 59.1% in Science.

In Nowshera, the average scores were 68.1% in English, 65.8% in General Knowledge, 66.0% in Mathematics, and 61.8% in Science. Peshawar recorded 63.3% in English, 67.4% in General Knowledge, 63.3% in Mathematics, and 63.4% in Science, while Swabi recorded 59.5% in English, 58.8% in General Knowledge, 59.4% in Mathematics, and 51.6% in Science.

At district level, the highest observed scores were 71.7% in English and 71.6% in Mathematics in Charsadda, while the highest General Knowledge score was 67.4% in Peshawar and the highest Science score was 68.7% in Charsadda. The lowest observed score across the district and subject combinations was 51.6% in Science in Swabi. The district-wise results are presented in Figure 20.

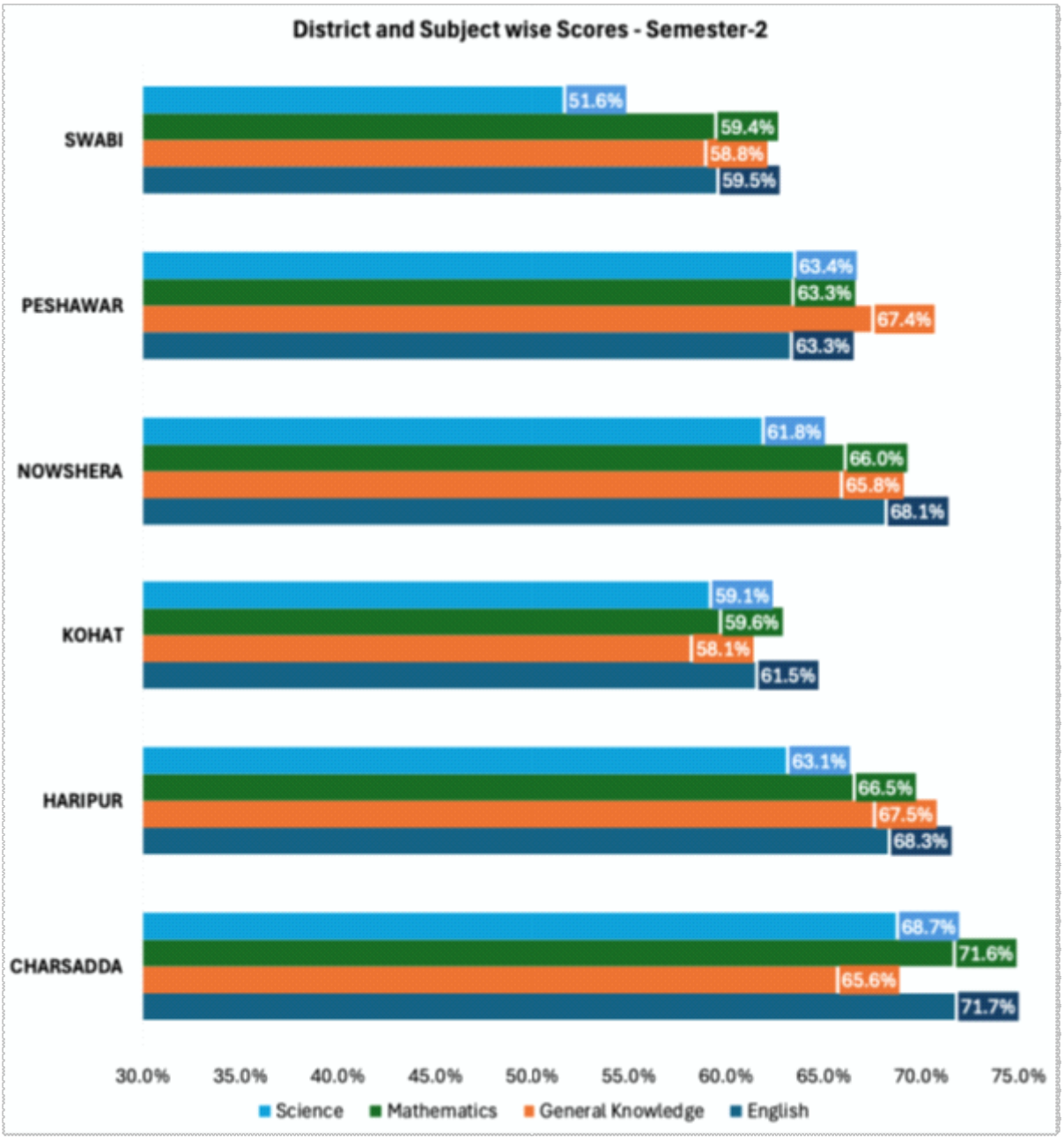


Figure 20: District and Subject-wise Scores - Semester-2

8.4 Overall Pre and Post Assessment Results by Treatment Status

To provide a broader descriptive view of student performance, the analysis also examines changes between Semester-1 and Semester-2 assessment results separately for treatment and control schools. These results are presented at the overall level. The analysis remains descriptive and should not be interpreted as a causal estimate of the intervention's effect on student learning.

For **English**,

the average score among treatment schools increased from **53.93% in Semester-1 to 64.72% in Semester-2**, representing a reported change of **10.79 percentage points**.

Among control schools, the corresponding average score increased from **52.05% to 63.25%**, representing a change of **11.20 percentage points**.

The results are presented in Figure 21.

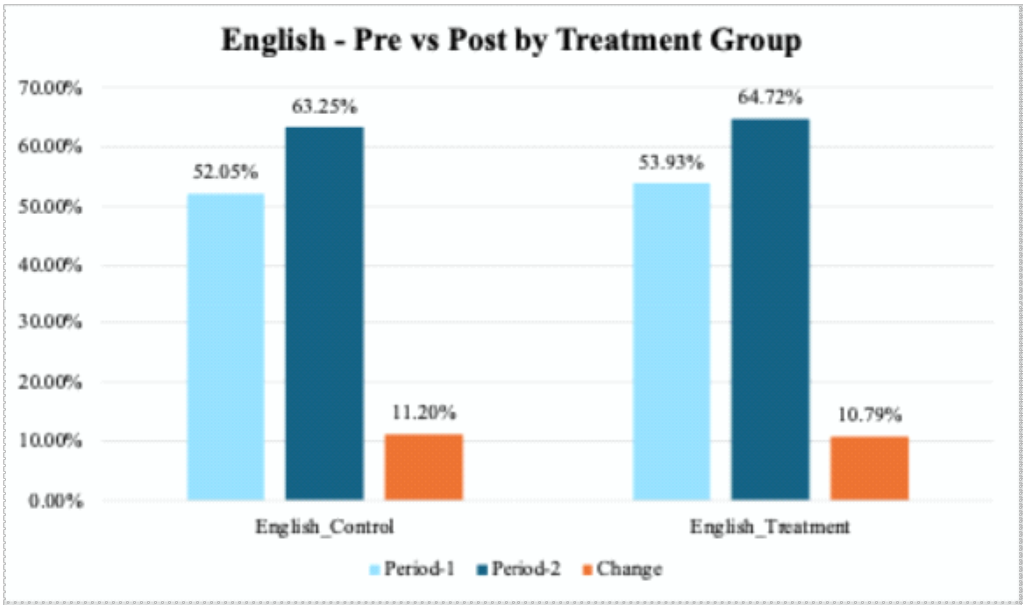


Figure 21: English - Pre vs Post by Treatment Group

For **Mathematics**,

treatment schools recorded an average score of **58.43% in Semester-1** and **63.83% in Semester-2**, with a reported change of **5.40 percentage points**.

In control schools, the average score increased from **57.04% to 61.71%**, representing a change of **4.67 percentage points**.

The results are presented in Figure 22.

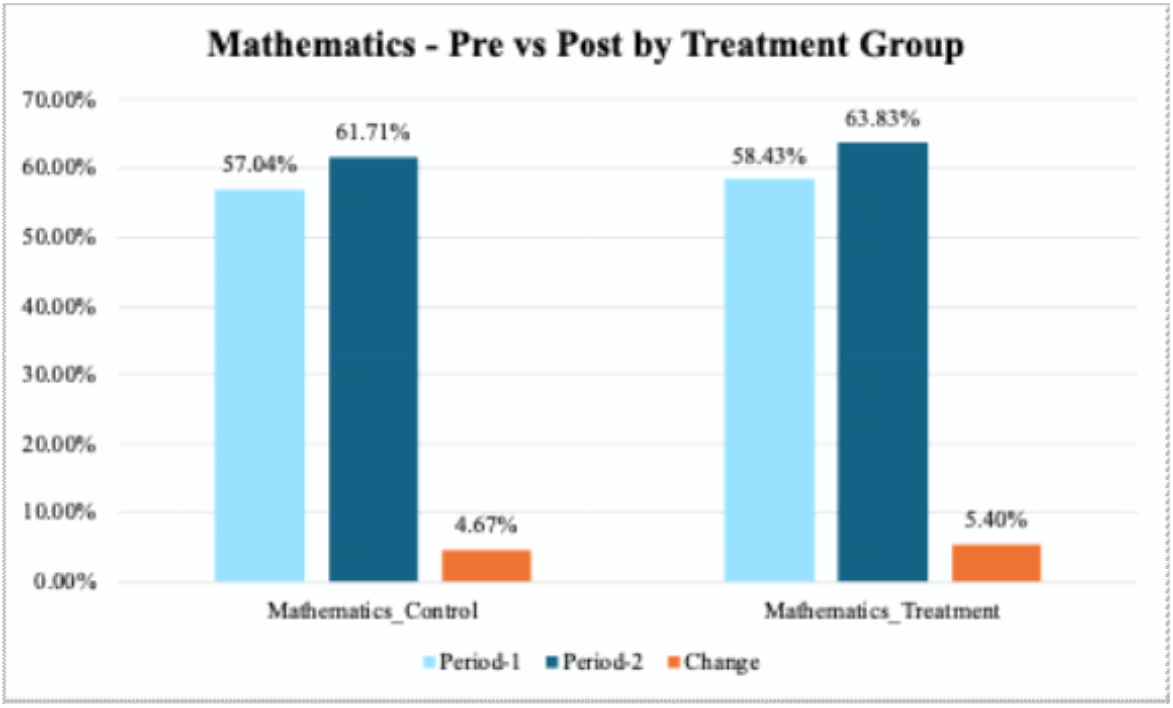


Figure 22: Mathematics - Pre vs Post by Treatment Group

For **General Knowledge**,

the average score in treatment schools was **56.86% in Semester-1** and **63.25% in Semester-2**, corresponding to a reported change of **6.39 percentage points**.

Among control schools, the average score increased from **54.24% to 64.12%**, representing a change of **9.88 percentage points**.

The results are presented in Figure 23.

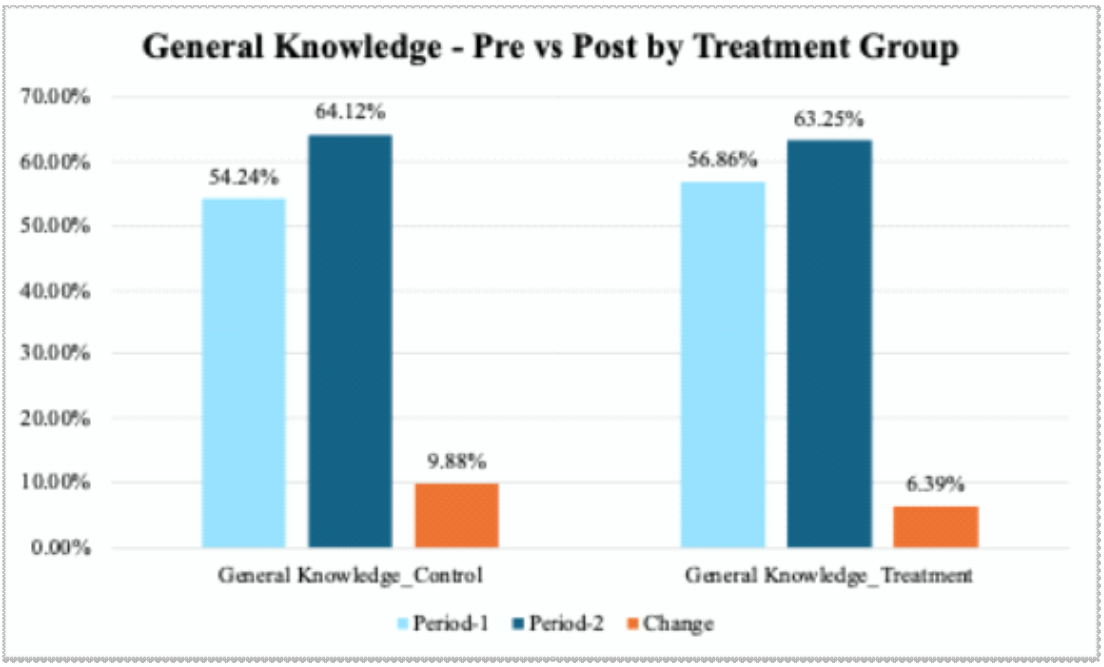


Figure 23: General Knowledge - Pre vs Post by Treatment Group

For **Science**,

treatment schools recorded an average score of **52.44% in Semester-1** and **60.10% in Semester-2**, with a reported change of **7.66 percentage points**.

In control schools, the corresponding scores were **50.86% in Semester-1** and **59.06% in Semester-2**, representing a reported change of **8.20 percentage points**.

The results are presented in Figure 24.

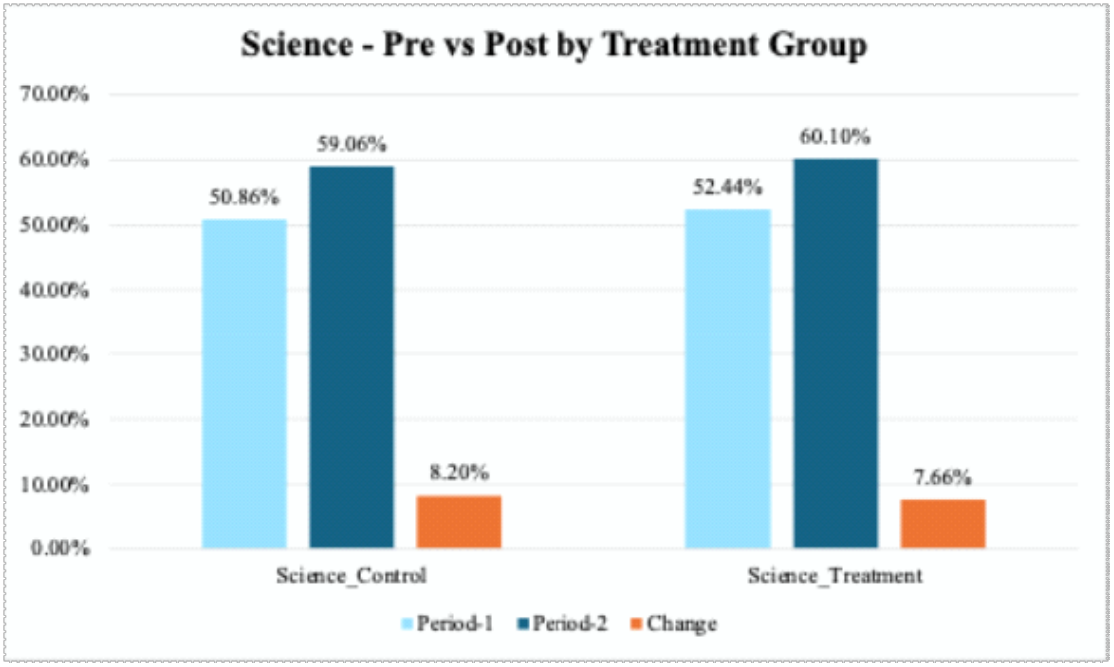


Figure 24: Science - Pre vs Post by Treatment Group

Taken together, the descriptive results show changes in student assessment scores between the two assessment periods in both treatment and control schools across all four subjects. However, these changes should not be interpreted in isolation as evidence of an intervention effect on student learning. In particular, the analysis presented here is a pre- and post-assessment comparison and does not account for other factors that may have influenced student performance during the assessment period.

9. Limitations of the Evaluation

Although the evaluation provides robust evidence on the effectiveness of data-driven mentoring and coaching for strengthening multigrade teaching practices, several limitations should be considered when interpreting the findings.

First, the evaluation measured changes in classroom instructional practices over a relatively short implementation period, with baseline observations in January 2026 and endline observations in March 2026. The findings therefore demonstrate improvements in teaching practice following the intervention but do not indicate whether these improvements will be sustained over longer periods. Continued classroom observations would be required to assess the durability of the observed changes.

Second, the evaluation focused primarily on classroom teaching practices as measured through the MG-COT. Although improvements in instructional quality are widely recognised as an important precursor to improved student learning, the evaluation could not directly examine the impact of the intervention on student learning outcomes. Consequently, no conclusions can be drawn regarding the effect of the intervention on student achievement.

Third, the magnitude of the treatment effects observed in KP is notably large. While the consistency of the findings across descriptive, practice-level, gender-disaggregated, and ANCOVA analyses supports their robustness, the low baseline scores recorded across both groups mean that a substantial part of the measured change reflects movement from a very low starting point. Results should therefore be interpreted alongside the baseline distribution rather than in isolation.

Fourth, implementation fidelity may have varied across districts because of differences in local contexts, school conditions, and the frequency and quality of mentoring visits conducted by individual School Leaders and ASDEOs. While common training materials, standardised observation tools, and structured coaching processes were used across all participating districts, some variation in implementation is expected under field conditions. In addition, the delayed training of School Leaders, caused by the non-renewal of SL contracts, meant that the full mentoring model was operational for a shorter period in some schools than in others.

Finally, the evaluation was conducted in four-teacher public primary schools, which represent a substantial proportion of Khyber Pakhtunkhwa's multigrade schools. Although many of the findings are likely to be relevant to other multigrade settings, caution should be exercised when generalising the results to schools with markedly different staffing patterns or organisational contexts.

Despite these limitations, the consistency of the findings across descriptive analyses, practice-level comparisons, gender-disaggregated analyses, and ANCOVA estimates provides strong evidence that strengthening the mentoring and coaching capacity of

School Leaders and ASDEOs contributed to measurable improvements in multigrade teaching practices.

10. Discussion

10.1 Implications for Multigrade Teaching

The evaluation provides evidence that strengthening the mentoring and coaching capacity of School Leaders and ASDEOs can improve the quality of instructional practice in multigrade classrooms. Both the descriptive analysis and the ANCOVA estimates consistently demonstrate considerably stronger improvement among treatment schools relative to control schools across overall MG-COT scores and individual classroom teaching practices.

The scale of improvement observed in Khyber Pakhtunkhwa is particularly notable given the operational complexity associated with multigrade classrooms in four-teacher schools. Teachers working in such settings are required to simultaneously manage students from multiple grade levels while ensuring effective lesson delivery, classroom management, differentiated instruction, and learner engagement. The findings suggest that targeted mentoring and coaching support can significantly strengthen teachers' ability to manage these instructional demands more effectively.

Substantial improvements were observed in practices related to lesson planning, instructional alignment, classroom organisation, student grouping, peer learning, and formative assessment. These findings indicate that teachers in treatment schools increasingly adopted structured and learner-centred teaching approaches during the intervention period.

The strong gains observed across almost all MG-COT practices further suggest that structured mentoring systems can contribute not only to incremental improvements in isolated classroom behaviours, but also to broader transformation in classroom instructional processes within multigrade settings.

The findings additionally reinforce the usefulness of structured classroom observation frameworks such as MG-COT in supporting instructional improvement. By providing clear performance benchmarks and observation criteria, the framework appears to have facilitated more focused classroom support and constructive feedback processes.

10.2 Implications for Mentoring and Coaching Systems

The evaluation findings highlight the importance of strengthening field-level academic support systems through capacity building of School Leaders and Assistant Sub-Divisional Education Officers.

The large and statistically significant treatment effects observed across treatment schools suggest that the training enabled School Leaders and ASDEOs to provide more effective mentoring and coaching support to teachers working in multigrade classrooms. The intervention appears to have strengthened the quality of classroom observation processes, feedback mechanisms, and instructional guidance provided during school visits.

The findings support a shift away from purely compliance-focused monitoring systems towards more developmental and instructional mentoring approaches. Structured observation tools such as MG-COT appear to have enabled School Leaders and ASDEOs to engage more systematically with classroom teaching practices and provide evidence-based academic support to teachers.

The results also demonstrate the importance of continuous professional support mechanisms for teachers operating in complex multigrade environments. In schools where teachers face significant instructional and classroom management challenges, sustained mentoring and coaching support may play a critical role in improving classroom quality over time.

Additionally, the stronger improvements observed among treatment schools suggest that school-level instructional change can be accelerated when academic supervision systems are linked with structured coaching, classroom observation, and ongoing professional guidance.

10.3 Lessons for Scale-up

The evaluation offers several important lessons for future scale-up of multigrade mentoring and coaching interventions within Khyber Pakhtunkhwa's public education system.

First, the findings demonstrate that strengthening the mentoring and coaching capacity of School Leaders and ASDEOs can generate substantial improvements in classroom instructional practices. This suggests that investing in field-based academic support systems can be an effective strategy for improving teaching quality in multigrade schools.

Second, the intervention highlights the importance of structured classroom observation frameworks in supporting instructional improvement. Integrating MG-COT or similar tools into routine school support mechanisms may strengthen consistency, quality, and focus within classroom observation processes.

Third, the findings indicate that substantial classroom improvement is achievable even within schools operating under teacher shortages and multigrade constraints. This has important implications for remote, underserved, and resource-constrained schools where multigrade teaching remains common.

Fourth, the intervention demonstrates the importance of combining technical training with sustained follow-up support. Future scale-up efforts may therefore benefit from integrating periodic refresher trainings, continuous mentoring support, and peer learning mechanisms for teachers and field-level education managers.

Fifth, the delay in training School Leaders arising from contract renewal issues underlines the importance of aligning the timing of capacity-building activities with human resource and contracting cycles, so that the full mentoring model becomes operational as planned.

Finally, the results suggest that strengthening instructional leadership and mentoring systems at the local level can contribute meaningfully towards broader improvements in teaching quality within public primary schools across Khyber Pakhtunkhwa.

11. Conclusion

This evaluation demonstrates that strengthening the mentoring and coaching capacity of School Leaders and ASDEOs can contribute to measurable improvements in multigrade teaching practices in Khyber Pakhtunkhwa's public primary schools. Across all analyses, treatment schools consistently demonstrated markedly greater improvements than control schools in overall MG-COT scores as well as across the core instructional practices assessed by the classroom observation framework. The ANCOVA analysis further confirms that these improvements remained large and statistically significant after controlling for baseline differences.

The intervention was particularly effective in strengthening practices associated with learner-centred pedagogy, including lesson planning, classroom space use, instructional alignment, flexible student grouping, differentiated instruction, and formative assessment. These improvements are especially relevant in multigrade classrooms, where effective classroom organisation and instructional adaptation are essential for ensuring meaningful learning opportunities for students across multiple grade levels.

Beyond improving classroom practice, the evaluation demonstrates the potential of a system-based approach to instructional improvement. By equipping School Leaders and ASDEOs with structured observation tools, practical coaching skills, and digital monitoring systems, the intervention strengthened the quality of academic support available to teachers while building institutional capacity within existing government structures.

Although student learning outcomes could not be assessed within the scope of this evaluation, the evidence provides a strong basis for expanding data-driven mentoring and coaching across multigrade schools in Khyber Pakhtunkhwa. Continued institutionalisation of MG-COT within routine supervision systems, together with strengthened assessment data and ongoing professional support, offers a practical pathway for improving instructional quality at scale.

12. Recommendations

Based on the evaluation findings, the following recommendations are proposed, covering policy as well as implementation level:

12.1 Immediate Priorities (Next 12 Months)

1. Institutionalise the MG-COT within Khyber Pakhtunkhwa's routine academic supervision system by fully integrating it into the existing classroom observation application and the reporting processes of School Leaders and ASDEOs.

2. Ensure that all School Leaders and ASDEOs responsible for multigrade schools receive standardised training on evidence-based mentoring and coaching, supported by annual refresher programmes and practical coaching simulations.

3. Align the scheduling of capacity-building activities with School Leader contracting cycles so that trained mentors remain in post for the full duration of the mentoring cycle.

4. Operationalise the MG-COT dashboard at provincial, district, and sub-division levels to enable regular monitoring of classroom practices, coaching coverage, and follow-up actions, with routine use of dashboard evidence in academic review meetings.

5. Develop and disseminate a comprehensive package of multigrade teaching resources, including lesson plans, academic calendars, classroom organisation strategies, differentiated learning activities, and formative assessment tools aligned with the MG-COT framework.

12.2 Medium-Term Priorities (1–3 Years)

1. Implement a phased expansion of the mentoring and coaching model to additional districts, prioritising schools with high levels of multigrade teaching and teacher shortages.

2. Strengthen the role of School Leaders and ASDEOs as instructional coaches by ensuring regular school visits, structured coaching conversations, documented action plans, and follow-up support rather than compliance-focused monitoring.

3. Integrate multigrade pedagogy, coaching methodologies, and classroom observation practices into both pre-service teacher education and continuous professional development programmes delivered through the DPD.

4. Use MG-COT data routinely to identify provincial and district training priorities, enabling teacher professional development programmes to respond directly to observed classroom needs.

12.3 Long-Term System Strengthening

1. Strengthen Khyber Pakhtunkhwa's student assessment systems by improving data quality, completeness, and reporting at student and learning-outcome levels so that

future evaluations can examine the relationship between improved teaching practices and student learning.

2. Conduct a follow-up evaluation to assess the sustainability of classroom practice improvements and examine whether regular mentoring produces lasting changes in instructional quality.

3. Undertake a cost-effectiveness and scalability analysis of the mentoring and coaching model to inform province-wide implementation and long-term resource allocation.

4. Establish a continuous monitoring and learning framework that combines classroom observation, coaching records, teacher professional development, and student assessment data to support evidence-based decision-making across the education system.

Annexure I: Training of ASDEO on Coaching and Mentoring

Directorate of Professional Development, Khyber Pakhtunkhwa
Charsadda Road Peshawar
Phone & Fax No. 091-9224783 **Email:** directordpdkp@gmail.com

No. 1598 /Mutligrade Teaching /2025 **Dated Peshawar, the December 23rd, 2025**

The District Education Officers (M&F)
Charsadda, Haripur, Kohat, Nowshera, Peshawar, and Swabi.

Subject: **NOMINATION OF ASDEOs FROM THE SELECTED CIRCLES TO ATTEND TRAINING ON COACHING AND MENTORING SKILLS FOR EFFECTIVE MULTIGRADE INSTRUCTION**

Memo:

I am directed to refer to the subject cited above and to state that in continuation of the training of ASDEOs held in December 2024 on Multigrade Teaching, conducted as part of the B-TAG/I-SAPS approved work plan of the Elementary & Secondary Education Department, Khyber Pakhtunkhwa, a follow-up training for ASDEOs from selected circles in your district has been scheduled on **"Coaching and Mentoring Skills for Effective Multigrade Instruction"** as per the given schedule.

Venue: Directorate of Professional Development (DPD), Larama Landy Sarak, Charsadda Road, Peshawar
Date & Time: December 30–31, 2025, from 09:00 AM to 04:00 PM

You are requested to nominate ASDEOs from the concerned circles (list attached) and direct them to attend this training at DPD Peshawar.

NOTE:

B-TAG/I-SAPS will provide lunch, stationery, and an honorarium of PKR 5,000 per day for the training to the participants

Enclosure: District-wise list of circles.

SENIOR INSTRUCTOR

Endst: Even No. & Date:
Copy for information to the:

1. Secretary to the Government of Khyber Pakhtunkhwa, E&SED Peshawar.
2. Director Elementary & Secondary Education Khyber Pakhtunkhwa Peshawar.
3. Additional Directors (Acad and Admn) Local Directorate
4. P.S to Director Local Directorate.
5. Project Director ISAPS-B-TAG Peshawar with the reference to his email dated Dec 23, 2025

Annexure II: MG-COT Approval Notification

DIRECTORATE OF CURRICULUM AND TEACHER EDUCATION
KHYBER PAKHTUNKHWA ABBOTTABAD
Phone # 0992-382634 Fax # 0992-381527 Email: director.dcte@kpese.gov.pk

No. 342-44/C&R/B-TAG/AI, dated: 07/01/2025

Subject: **REVIEW AND APPROVAL OF MULTIGRADE CLASSROOM OBSERVATION TOOL (MG-COT)**

The Directorate of Curriculum & Teacher Education (DCTE) conducted a review workshop from 22/12/2025 to 25/12/2025. Following the review, the committee is pleased to approve the Multigrade Classroom Observation Tool (MG-COT) for use in multigrade classroom observations.

DIRECTOR

Even No & Date.
Copy for information to the:

1. PS to Secretary to Khyber Pakhtunkhwa E&SE Department, Peshawar
2. Provincial Team Lead B-TAG Peshawar.
3. P.S to Director, Local Directorate

(DR. IRUM NAZ)
SS (Material Development)

Annexure III: Multigrade Classroom Observation Tool (COT)

Part A: General Information				
School name		MIS code		
Teacher name		CNIC		
Name of observer		Observation Date		
Observer designation (circle one)	School leaders	Headteacher	Teacher	
About the lesson you observe				
Grades Combined		Number of children in class today		
Subject				
Topic				
Main SLO				

Part B: MGT CLASSROOM OBSERVATION TOOL					
No.	**Indicator**	**Score**			**Observation**
		Emerging (1)	**Developing (2)**	**Proficient (3)**	
Multigrade Classroom Management					
1.	The teacher combines grades logically (e.g., ECCE-G1, G2-G3, and G4-G5).				
2.	The teacher uses classroom space effectively to accommodate both grades and encourage interaction (e.g., flexible seating, subject-specific corners).				
Academic Calendar & Lesson Planning					
3.	The teacher implements multigrade academic calendar.				
4.	The teacher uses a multigrade lesson plan.				
Instructional Strategies					
5.	The teacher explicitly articulates the learning objectives for each grade level.				
6.	The teacher's explanation of the content aligns with the stated learning objectives for each grade level.				
7.	The teacher's explanation of the content is clear and accurate catering to students of both grades.				
8.	The content presented in the lesson is appropriate,				

	relevant, and differentiated for each grade level.				
9.	The teacher manages time effectively to balance instructional attention between grades.				
Student Engagement					
10.	The teacher uses flexible grouping strategies to enable students from both grades to collaborate based on their needs and abilities.				
11.	The teacher promotes peer learning by encouraging students to actively support each other, creating a collaborative learning environment.				
Assessment & Feedback					
12.	The teacher uses a variety of formative assessment techniques, including peer and self-assessment, to monitor and support student progress across both grades.				
13.	The teacher provides individualised feedback to students, tailored to their grade level and specific needs.				

Scoring Rubric for the Classroom Observation Tool with Examples

The rubric describes the practice at three levels, low, medium and high. Note that the rating scale in the classroom observation tool allows observers to rate the practice from 1(low), 2 (medium), and 3 (high).

No.	Indicator	Score		
		Low (1)	Medium (2)	High (3)
Multigrade Classroom Management				
1.	**The teacher combines grades logically (e.g., ECCE-G1, G2-G3, and G4-G5).**	The teacher does not combine grades logically, leading to confusion and frustration. ***Example**: The teacher combines ECCE and Grade 2 students for the same literacy activity. ECCE students struggle with basic letter recognition, while Grade 2 students find the material too easy and unchallenging. This misalignment results in both groups feeling disengaged and confused.*	The teacher combines grades logically but does not fully consider the different learning needs, leading to partial effectiveness. ***Example**: The teacher combines Grade 2 and Grade 3 students for a literacy lesson. While both groups read the same passage, Grade 3 students require more advanced comprehension tasks, while Grade 2 students find the content too challenging. This leads to uneven progress and engagement.*	The teacher consistently combines grades logically, addressing the learning needs of each group and ensuring differentiated instruction. ***Example**: The teacher combines ECCE and Grade 1 students for a literacy lesson. ECCE students work on letter recognition, while Grade 1 students focus on forming simple sentences. The teacher provides tasks appropriate to each grade level, ensuring an organised and engaging learning environment for all students.*
2.	**The teacher uses classroom space effectively to accommodate both**	The teacher does not use classroom space effectively to	The teacher makes some effort to use the classroom space, but it does not fully accommodate	The teacher organises classroom space effectively,

	grades and encourage interaction (e.g., flexible seating, subject-specific corners).	accommodate both grades and encourage interaction. ***Example**: The teacher arranges desks in rows without creating flexible seating or subject-specific corners for Grade 2 and Grade 3 students. There is no space for collaborative work, and students from both grades struggle to engage with the lesson, leading to confusion and lack of interaction.*	both grades and encourage effective interaction. ***Example**: The classroom includes a reading corner and some flexible seating, but it is not well-suited to both Grade 2 and Grade 3 needs. While Grade 3 students use the space for independent work, Grade 2 students are less engaged, and interaction between the grades is limited.*	accommodating both grades and encouraging smooth interaction. ***Example**: The teacher designs subject-specific corners, such as a phonics area for Grade 2 and a maths station for Grade 3, while shared spaces like a collaborative project table allow for interaction. Flexible seating options encourage movement between independent work and group activities, fostering an organised and interactive environment for both grades.*
Academic Calendar & Lesson Planning				
3.	**The teacher implements multigrade academic calendar.**	The teacher implements a monograde academic calendar, resulting in confusion and misalignment for the combined grades. ***Example**: The teacher follows a Grade 3 academic calendar without adjustments for Grade 2 students. Both grades are given the same "Measurement" tasks, which are too advanced for Grade 2 students, leaving them*	The teacher partially implements the multigrade academic calendar. Some activities do not fully integrate overlapping content or align with the term-wise plan. ***Example**: The academic calendar schedules "Measurement" for both Grade 2 and Grade 3 in the first term. The teacher plans lessons where Grade 2 students measure objects using rulers, while Grade*	The teacher consistently implements the multigrade academic calendar, effectively integrating overlapping content and following the term-wise plan. ***Example**: The academic calendar schedules "Measurement" for both Grade 2 and Grade 3 during the first term. The teacher plans lessons where Grade 2 students measure objects with non-standard units (e.g., paperclips), while Grade 3*

		confused and unable to keep up, while Grade 3 students proceed without difficulty. This leads to disorganised lessons and gaps in learning for the younger students.	*3 students calculate area and perimeter. However, the teacher sometimes introduces unrelated topics like "Time," which disrupts the flow and causes confusion for the students.*	*students work on calculating area and perimeter. Both grades collaborate, with Grade 3 students helping Grade 2 students, ensuring every lesson follows the academic calendar and creating a structured and engaging learning experience.*
4.	**The teacher uses a multigrade lesson plan.**	The teacher uses two separate monograde lesson plans, resulting in a lack of integration and alignment across the combined grades. ***Example:*** *The teacher prepares one lesson plan for Grade 2 focused on "Addition" with basic two-digit problems (e.g., 25 + 13) and another for Grade 3 with more advanced three-digit subtraction problems (e.g., 456 - 123). Because these plans are not aligned, there is no opportunity for collaboration or shared learning between the grades. As a result, both grades work on different topics simultaneously, creating confusion and*	The teacher uses a multigrade lesson plan aligned with the academic calendar but with partial differentiation between grades, resulting in inconsistencies. ***Example****: The teacher plans a lesson on "Addition and Subtraction" for both Grade 2 and Grade 3. Grade 2 students work on two-digit problems without regrouping (e.g., 34 + 12), while Grade 3 students solve three-digit problems with regrouping (e.g., 235 + 178). However, when they work together on mixed problems, the younger students get confused by the more advanced material. The lesson somewhat follows the academic calendar, but the lack*	The teacher prepares and implements a well-structured multigrade lesson plan that is fully differentiated for each grade and aligned with the academic calendar. ***Example****: The teacher plans a lesson on "Addition and Subtraction" where Grade 2 students practise two-digit problems without regrouping (e.g., 48 + 21), while Grade 3 students tackle multi-digit problems with regrouping (e.g., 235 + 178) and real-world applications. During group work, Grade 3 students assist Grade 2 students in understanding regrouping. The lesson is perfectly aligned with the academic calendar, ensuring*

		making it difficult for the teacher to manage the class effectively.	*of clear differentiation causes confusion during group work.*	*both grades are appropriately challenged and engaged.*
Instructional Strategies				
5.	**The teacher explicitly articulates the learning objectives for each grade level.**	The teacher does not state the learning objectives for both grades, and the activities do not clarify the objectives. ***Example:*** *The teacher begins the lesson by having both Grade 4 and Grade 5 students read a story but fails to explain the learning objectives. Grade 4 students focus on identifying key details, and Grade 5 students are tasked with writing a summary. Since the teacher does not clarify the purpose behind these tasks or connect them to the learning objectives, students are left confused about the lesson's focus.*	The teacher states a general objective for both grades but does not clearly differentiate or explain the specific objectives for each grade. ***Example****: The teacher says, "Today we are working on reading comprehension," but does not specify distinct objectives for Grade 4 and Grade 5. While Grade 4 students work on identifying details and Grade 5 students write a summary, the teacher does not clearly explain the specific goals for each grade, leading to some confusion about the lesson's expectations.*	The teacher clearly articulates the differentiated learning objectives for each grade, and the activities align with these objectives. ***Example****: The teacher begins the lesson by saying, "Grade 4 will focus on identifying key details from the text, while Grade 5 will work on summarising the story." The teacher explains how the activities are designed to meet these objectives, ensuring that Grade 4 students concentrate on details and Grade 5 students focus on writing summaries, with both tasks clearly aligned to their respective learning objectives.*
6.	**The teacher's explanation of the content aligns with the stated learning**	The teacher's explanation does not align with the stated lesson objectives, and the content does not support the	The teacher's explanation partially aligns with the stated learning objectives, but there are significant distractions or gaps.	The teacher's explanation fully aligns with the stated learning objectives, and the content

	objectives for each grade level.	learning goals for either grade. ***Example**: The stated objective is for ECCE students to listen to letter sounds and for Grade 1 students to practise saying the letters. However, the teacher spends the lesson talking about animal names and showing pictures of animals. This has nothing to do with the alphabet sounds, and students are left confused, without practising their listening or speaking skills.*	***Example**: The teacher's objective is for ECCE students to listen to the alphabet sounds and for Grade 1 students to practise saying the letters. The teacher starts by introducing some letter sounds, but then spends too much time playing unrelated phonics games. While students are exposed to some relevant content, they don't spend enough time focusing on the actual objectives, reducing the lesson's effectiveness.*	clearly supports the goals for each grade. ***Example**: The teacher clearly explains that ECCE students will listen to the alphabet sounds, and Grade 1 students will practise saying the letters. The teacher uses a song to help ECCE students hear the letter sounds, then asks Grade 1 students to say each sound aloud. The activities are well-matched to the objectives, and both groups are focused on learning their specific skills.*
7.	**The teacher's explanation of the content is clear and accurate catering to students of both grades.**	The teacher's explanation is unclear or incorrect, leaving students confused or with misconceptions. ***Example**: The teacher attempts to explain multiplication to both Grade 4 and Grade 5 but inaccurately describes it as "adding numbers together." This confuses both grades, and when the teacher incorrectly*	The teacher's explanation is mostly correct but unclear or too simple for one or both grades. ***Example**: The teacher gives a correct explanation of multiplication but only addresses basic two-digit multiplication. This explanation proves too challenging for some Grade 4 students and too simplistic for udents, who are ready for more advanced concepts. Although*	The teacher provides a clear and correct explanation for both grades, adjusting the explanation to each grade's level. ***Example**: The teacher explains multiplication clearly for both Grade 4 and Grade 5. For Grade 4, the teacher uses simple examples like "3 groups of 4" and draws arrays to show basic multiplication. For Grade 5, the teacher explains more complex*

		states that multiplying by "zero" is the same as multiplying by "one," it leads to further misconceptions. As a result, neither grade understands multiplication properly, causing frustration among students.	*the explanation is accurate, it does not adequately cater to the varying levels of both grades, leading to only partial understanding.*	*multiplication, like three-digit by two-digit problems, and introduces regrouping. Each grade receives a tailored explanation that helps them understand the concept fully.*
8.	**The teacher's explanation of the content aligns with the stated learning objectives for each grade level.**	The teacher's explanation does not align with the stated lesson objectives, and the content does not support the learning goals for either grade. ***Example**: The stated objective is for ECCE students to listen to letter sounds and for Grade 1 students to practise saying the letters. However, the teacher spends the lesson talking about animal names and showing pictures of animals. This has nothing to do with the alphabet sounds, and students are left confused, without practising their listening or speaking skills.*	The teacher's explanation partially aligns with the stated learning objectives, but there are significant distractions or gaps. ***Example**: The teacher's objective is for ECCE students to listen to the alphabet sounds and for Grade 1 students to practise saying the letters. The teacher starts by introducing some letter sounds, but then spends too much time playing unrelated phonics games. While students are exposed to some relevant content, they don't spend enough time focusing on the actual objectives, reducing the lesson's effectiveness.*	The teacher's explanation fully aligns with the stated learning objectives, and the content clearly supports the goals for each grade. ***Example**: The teacher clearly explains that ECCE students will listen to the alphabet sounds, and Grade 1 students will practise saying the letters. The teacher uses a song to help ECCE students hear the letter sounds, then asks Grade 1 students to say each sound aloud. The activities are well-matched to the objectives, and both groups are focused on learning their specific skills.*

9.	**The content presented in the lesson is appropriate, relevant, and differentiated for each grade level.**	The content is not appropriate for the grade level(s), either too advanced or too simple, and does not support students' learning needs. ***Example**: The teacher presents both Grade 2 and Grade 3 students with difficult words like "unbelievable" and "complicated" without first teaching simpler words. This is too hard for Grade 2 students, who are still learning basic words, and does not challenge Grade 3 students appropriately. The lack of differentiation confuses and disengages both groups.*	The content is somewhat appropriate but lacks sufficient differentiation for each grade, limiting its effectiveness. ***Example**: The teacher introduces "cat" for Grade 2 and "elephant" for Grade 3, but then asks both groups to practise with "butterfly." While "butterfly" is a good word for Grade 3, it is a bit difficult for Grade 2 students, who need simpler words. The content is partially relevant, but the lack of differentiation causes some students to struggle and others to be under-challenged.*	The content is fully appropriate and well-differentiated for each grade, supporting the learning objectives effectively. ***Example**: The teacher presents Grade 2 students with words like "rabbit" and "sunset" (two-syllable words) and gives Grade 3 students more complex words like "elephant" and "butterfly" (three-syllable words). Each grade works on their level, with Grade 2 mastering basic syllables and Grade 3 tackling more challenging words. The content is well-matched to each grade, engaging all students appropriately.*
10.	**The teacher manages time effectively to balance instructional attention between grades.**	The teacher does not manage time effectively, causing disorganisation and loss of instructional flow. ***Example:** During a math lesson, the teacher spends too long on direct instruction for Grade 2, leaving little time for Grade 3 students to complete their tasks. As a*	The teacher manages time somewhat effectively but struggles to maintain a smooth instructional pace. ***Example:** The teacher alternates between helping Grade 2 students with basic subtraction and Grade 3 students with more advanced problems. However, the teacher*	The teacher manages time effectively, ensuring smooth transitions and lesson flow for both grades. ***Example:** The teacher starts with a focused explanation for Grade 2 students and sets them up with an independent task while providing more complex instruction to Grade 3. The*

		result, the lesson runs overtime, and neither grade finishes the planned activities.	*spends too much time with Grade 2, causing Grade 3 students to wait and run out of time to complete their tasks, resulting in a rushed and uneven lesson.*	*teacher tracks the time carefully, ensuring that both grades have sufficient time to work on their assignments and complete the activities before transitioning to the next part of the lesson.*
Student Engagement				
11.	**The teacher uses flexible grouping strategies that enables students to collaborate with peers from both their own or other grades, based on needs and abilities.**	The teacher does not use flexible grouping strategies, and students are grouped strictly by grade or randomly, without considering their needs or abilities. ***Example**: The teacher puts all Grade 2 students in one group and all Grade 3 students in another for a reading comprehension activity. There is no attention to different reading levels or specific needs. As a result, stronger students are underchallenged, and struggling students are not supported. There is no collaboration between the grades, and the groups don't reflect students' abilities,*	The teacher uses some flexible grouping strategies, considering either grade level or abilities, but not both. Some collaboration occurs across grades, but grouping is not fully based on individual needs. ***Example**: The teacher allows a few strong Grade 2 readers to work with Grade 3 students who need support during a reading activity. However, most students remain in grade-based groups. While some cross-grade collaboration occurs, the teacher doesn't consistently use abilities to form groups, leading to mixed results, where only some students benefit from the collaboration.*	The teacher effectively uses flexible grouping strategies that allow collaboration across grades, based on students' needs and abilities. Groups are dynamic and change during the lesson to maximise learning. ***Example**: The teacher groups students for reading based on their skills, not grade. A strong Grade 2 reader is paired with a Grade 3 student to read a more advanced text, while a struggling Grade 3 student joins Grade 2 students to work on simpler texts. As the lesson progresses, the teacher adjusts the groups to meet changing needs, ensuring all students are challenged and supported through peer*

		leading to limited engagement.		*collaboration, enhancing learning for everyone.*
12.	**The teacher promotes peer learning by encouraging students to actively support each other, creating a collaborative learning environment.**	The teacher does not encourage peer learning or collaboration between older and younger students. Each grade works independently. ***Example**: The teacher separates Grade 2 and Grade 3 students into different groups and teaches them separately. There is no interaction or collaboration between the two grades, and older Grade 3 students do not help the younger Grade 2 students. This isolation limits opportunities for peer learning and collaboration.*	The teacher occasionally encourages peer learning, but the collaboration between older and younger students is limited. Older students may help younger students, but it is informal and lacks structure. ***Example**: The teacher allows Grade 3 students to help Grade 2 students if they finish their work early, but this happens without structure. While some Grade 3 students assist their younger peers, it is not planned, and the support is inconsistent. Peer learning happens randomly, and younger students don't always get the help they need.*	The teacher actively promotes and structures peer learning, with older students consistently supporting younger students in a collaborative environment. ***Example**: The teacher pairs Grade 3 students with Grade 2 students during a reading activity. Grade 3 students are given clear roles to help their younger peers understand the reading and answer questions. The teacher structures the tasks so that Grade 3 students explain the material, reinforcing their own learning while supporting Grade 2 students. This approach is consistently used to create a collaborative learning environment.*
Assessment & Feedback				
13.	**The teacher uses a variety of formative assessment techniques, including peer and self-assessment, to**	The teacher does not use a variety of formative assessment techniques or incorporate peer or self-	The teacher uses some formative assessment techniques but does not fully incorporate peer or self-	The teacher consistently uses a variety of formative assessment techniques, including peer and

	monitor and support student progress across both grades.	assessment, resulting in missed opportunities for monitoring progress. ***Example**: The teacher gives both Grade 2 and Grade 3 students the same worksheet on subtraction without adapting it to their level. There are no opportunities for peer or self-assessment, and students complete the task independently. The teacher does not know how well students from either grade are progressing.*	assessment, leading to uneven monitoring of student progress. ***Example**: During a subtraction lesson, the teacher uses questioning to assess Grade 3 students but does not provide sufficient assessment for Grade 2 students. The teacher briefly mentions peer-assessment but does not structure it properly. As a result, Grade 2 students do not have a clear understanding of their progress, and peer-assessment is inconsistent.*	self-assessment, to monitor progress across both grades. ***Example**: During the subtraction lesson, the teacher uses tailored worksheets: Grade 2 students practice basic subtraction while Grade 3 students solve more complex problems. The teacher integrates peer-assessment by having Grade 3 students review the work of Grade 2 students, and Grade 2 students use a self-assessment checklist to reflect on their performance. These varied assessment techniques allow the teacher to effectively monitor the progress of both grades.*
14.	**The teacher provides individualised feedback to students, tailored to their grade level and specific needs.**	The teacher provides generic or minimal feedback that does not address individual grade levels or student needs. ***Example**: After a writing assignment, the teacher tells both Grade 4 and Grade 5 students general comments like "Good effort" or "Try harder next time," without pointing out specific strengths*	The teacher provides some individualized feedback, but it is not fully tailored to both grade level and individual student needs. ***Example**: After a reading comprehension activity, the teacher gives feedback focused on main ideas for both Grade 4 and Grade 5 students. While this feedback is relevant for Grade 4*	The teacher provides individualized feedback consistently, tailored to both grade level and specific student needs. ***Example**: After a creative writing activity, the teacher gives Grade 4 students feedback on sentence structure and descriptive language, while Grade 5 students receive feedback on*

		or areas for improvement. The feedback does not help students understand what they did well or what they need to work on, leaving them without clear guidance.	*students, Grade 5 students, who need more advanced feedback on analysing themes, only receive general comments. As a result, some students benefit, but others do not get the specific feedback they need for their grade level.*	*organizing paragraphs and developing stronger arguments. Additionally, individual needs are considered: a Grade 4 student struggling with grammar gets tips on improving punctuation, while a high-achieving Grade 5 student is encouraged to use more advanced vocabulary. This feedback helps each student improve according to their grade level and abilities.*