PAKISTAN INSTITUTE OF EDUCATION

NATIONAL ACHIEVEMENT TEST 2023

Pakistan Institute of Education
Ministry of Federal Education and Professional Training

Contents

Abbreviations

AEAC	Area Education Assessment Centre
AIR	American Institute for Research
AJK	Azad Jammu and Kashmir
CPEC	China Pakistan Economic Corridor
CRQ	Constructed Response Question
CTS	Critical Thinking Skills
ESR	Education sector reform
ECE	Early childhood education
FL	Foundational Literacy
GB	Gilgit-Baltistan
GPF	Global Proficiency Framework
ICT	Islamabad Capital Territory
ITA	Idara-e-Taleem-o-Aagahi
KP	Khyber Pakhtunkhwa
LMT	Lead Master Trainer
LSA	Large-scale assessment (the live NAT)
MCQ	Multiple choice question
MoFEPT	Ministry of Federal Education and Professional Training
NAT	National Achievement Test
NAW	National Assessment Wing
NC	National Curriculum
NEAS	National Education and Assessment System
NEMIS	National Education Management Information System
NMD	Newly Merged Districts
OMR	Optical Mark Recognition
PEAC	Provincial Education Assessment Centre
PIE	Pakistan Institute of Education
PPS	Probability Proportional to Size
PSLM	Pakistan Social and Living Standards Measurement
PTC	Primary Teaching Certificate
SD	Standard deviation
SDG	Sustainable Development Goal
SE	Standard error
SLO	Student Learning Outcome
SNC	Single National Curriculum
SOP	Standard Operating Procedure
TA	Test Administrator
TIMSS	Trends in International Maths and Science Study
TIPS	Teacher Induction Programmes
ToS	Tables of Specification

A message from the Federal Minister of Education

A country's human capital is relying on how well it prepares its future generations and how it is directly linked with the quality of education and education achievements of school going children. The Constitution of Pakistan mandates its duty bearers to lead and direct federal and provincial institutions through sound planning, policy reforms, and legislative framework that ensures that the development of country's human resource is closely associated with the socio-economic development of the country.

Government of Pakistan is fully aware of it responsibilities and the challenges that lie ahead to ensure that the provision of the Article 25A and similar provincial enactments. All of these provisions enable the federal and provincial governments to provide education to all its school-aged children without prejudice.

In light of this commitment, Pakistan Institute of Education, with the support of The World Bank and with technical partnership with Cambridge Partnership for Education, UK., has completed a comprehensive National Achievement Test across Pakistan, for grades 4 and 8. This assessment, I have been informed, provides a solid baseline on where learning achievement of our students stand today, notwithstanding the fact how the country's education system and its process were severely impacted during pandemic followed by unprecedented torrential floods.

The results are satisfactory, certainly not ideal but we know where we stand and what needs to be done next. This fact alone sets our agenda clearly for the next 5-10 years. I want to take this opportunity to thank all the provincial and regional education leadership and the departments who supported the Ministry in undertaking this very important exercise. This one-year exercise allowed the Ministry to test 23,000 students in different subjects and understand how students have performed in each province / region and against the learning outcomes. It provides analysis on how girls and boys from rural and urban areas have performed collectively and separately.

I have found the findings to be quite insightful, as they have shed light on some key fundamentals that directly impact the quality of teaching and learning in the classroom. I would urge the federal and provincial leadership and the institutions and bi-lateral donors to review these findings and recommendations, draw your own conclusions but work closely with the Ministry and Pakistan Institute of Education in developing policy recommendations based on this report and take this agenda forward in partnership.

I would like to take this opportunity to acknowledge the commitment of Director General PIE Dr. Muhammad Shahid Soroya, Muhammad Shakeel – Director National Assessment Wing PIE and their assessment team. We look forward to receiving more evidence from PIE which would help us make improved decisions for the children of Pakistan.

Mr. Madad Ali Sindhi,
Federal Minister
Ministry of Federal Education and Professional Training, Government of Pakistan

Message from Federal Secretary

National Achievement Test are conducted biennially in Pakistan, it provides an opportunity to understand and learn how students in grade 4 and 8 have performed in select set of subjects against well established, psychometrically tested items. This year 23,000 students participated in the learning achievement tests which were conducted across all provinces /regions for both girls and boys residing in urban and rural set-up. I want to congratulate Pakistan Institute of Education on achieving this major milestone.

On behalf of the Ministry, I want to extend my gratitude to The World Bank Team for their support to the Ministry and working closely with Pakistan Institute of Education (PIE) on this important milestone. In addition, engaging Cambridge Partnership for Education to work with PIE's Assessment Wing faculty, providing the technical lead throughout the process and building PIE's capacity in conducting high-stake assessment.

I want to take this opportunity to thank provincial /regional leadership and the education departments for their support to PIE throughout the year in undertaking this important exercise. These finding will also allow Pakistan to report on SGD 4 commitments.

Findings from the National Achievement Test 2023 are quite insightful and point to some conclusions that directly talk about the quality of teaching and learning in the classroom. I would urge the federal and provincial leadership and the institutions and bi-lateral donors to organize consultative dialogue at all levels, develop policy write-ups and recommend actions in the light of these findings.

I would like to take this opportunity to thank Director General PIE Dr. Muhammad Shahid Soroya, Muhammad Shakeel – Director National Assessment Wing PIE and their assessment team on this achievement and we look forward to PIE to continue to provide compelling evidence to make improved decisions for the children of Pakistan.

Mr. Waseem Ajmal Chaudhry,

Federal Secretary
Ministry of Federal Education and Professional Training
Government of Pakistan

Acknowledgement

Pakistan Institute of Education (PIE) on behalf of the Ministry of Federal Education and Professional Training (MoFE&PT) would like to offer our sincerely gratitude to The World Bank, FCDO and Cambridge Partnership for Education, UK., for their support in achieving this significant milestone.

This was a one of the best examples of collaborative work between the provinces /regions and PIE and I want to congratulate and extend my gratitude to the provincial /regions leadership and education departments in helping PIE achieve it amicably.

I want to take this opportunity to thank my assessment team ably lead by Director Muhammad Shakil, without my team's dedication and commitment, this exercise would not have been possible. Support provided by the administration and finance team at PIE is duly acknowledged in making this exercise a success.

Dr. Muhammad Shahid Soroya
Director General
Pakistan Institute of Education
Ministry of Federal Education & Professional Training

Executive summary

What is the National Achievement Test?

The National Achievement Test (NAT) is a sample-based assessment that takes place approximately every two years in a range of subjects at Grade 4 and Grade 8 in Pakistan. It is conducted by the National Assessment Wing (NAW) of the Pakistan Institute of Education (PIE), with support from provincial and area education departments. Before 2023, the most recent NAT was carried out in 2020, although it is known as NAT 2019.

What is the purpose of the NAT?

Pakistan has a high incidence of learning poverty (defined as being unable to read and understand a simple text by the age of 10) due to a range of structural factors. It is important to ensure continued improvement in the assessment system to help inform policy and deliver actions that reduce learning poverty.

If policy makers and other stakeholders do not know how successful (or unsuccessful) schools are in transforming resources into student learning, they risk maintaining suboptimal educational environments. When compared with total expenditure on education, a national assessment is a relatively inexpensive complement to reform efforts to improve learning.

The NAT is a low-stakes assessment for students and teachers that provides high-quality and data-driven information. Stakeholders can use this data to evaluate student attainment, measure the impact of teaching and learning, assess the efficacy of educational policies and reforms, and determine the cost benefits of investments in training and resources.

Policy makers and other stakeholders can use this empirical data from the NAT to make informed decisions regarding the allocation of resources. The core purposes of the NAT, as stated by NAW are to:

1. monitor student learning outcomes (SLO) and education quality levels over time
2. provide information for reforming standards and curriculum
3. generate data for supporting teacher training and materials development
4. inform policy reforms, reduce inequalities, and promote accountability
5. improve student learning outcomes.

What data does the NAT collect?

The NAT consists of two separate steps:

- Subject tests taken by students based on the curricula of English, Urdu, Sindhi (Sindhi was only taken in the Sindh Province and Urdu in other provinces), as well as Maths for Grade 4, and Maths and Science for Grade 8.
- Background questionnaires completed by headteachers, teachers, parents and students.

The subject tests cover content and cognitive domains in each curriculum. The background questionnaires collect data about, among others, demographics, learning and teaching, and school and home environments. Students' performance is analysed in connection with the demographic and learning factors in the questionnaires.

Who took the 2023 NAT?

A representative sample of 1304 public sector schools (652 in each of Grades 4 and 8) was randomly selected and over 23000 students across the country took the NAT on 18 and 19 May 2023. The study used scientific sampling techniques to make sure the final sample of Grade 4 and 8 students was representative at the provincial and national level in terms of gender, rural or urban location and type of school. Of the sampled schools, 643 and 640 participated in the Grade 4 and Grade 8 studies respectively. A total of 11073 students participated in the Grade 4 study and 12383 in Grade 8.

How was the NAT developed?

Pilot papers were developed in December 2022 and were completed by a sample of students at 37 Grade 4 schools and 38 Grade 8 schools in January 2023. In March 2023, a panel of subject specialists developed the live NAT, using questions that demonstrated robust psychometric properties from the analysis of the pilot items.

This resulted in one test booklet being produced for Grade 4 Maths, English, Urdu and Sindhi reading as well as for Grade 8 Maths and Science. Sindhi and Urdu translations were produced for Maths and Science at both Grades. Grade 4 students also took a Foundational Literacy test made up of 15 questions that tested letter and word recognition at Grade 4.

How did students perform on the NAT?

As shown in Figure 1, on average, Grade 4 students answered 88% of items (13.2 out of 15) of the Foundational Literacy items correctly. Students also answered a high proportion of items correctly on average in Grade 4 Urdu and Sindhi (68% or 35.4 out of 52). On average, just over half of the items were answered correctly in Grade 4 English (56% or 26.9 out of 48) and Grade 8 Science (51% or 26.7 out of 52). However, the number of correct answers was slightly below half in Grade 4 Maths (49% or 23.7 out of 48) and Grade 8 Maths (42% or 21.6 out of 52).

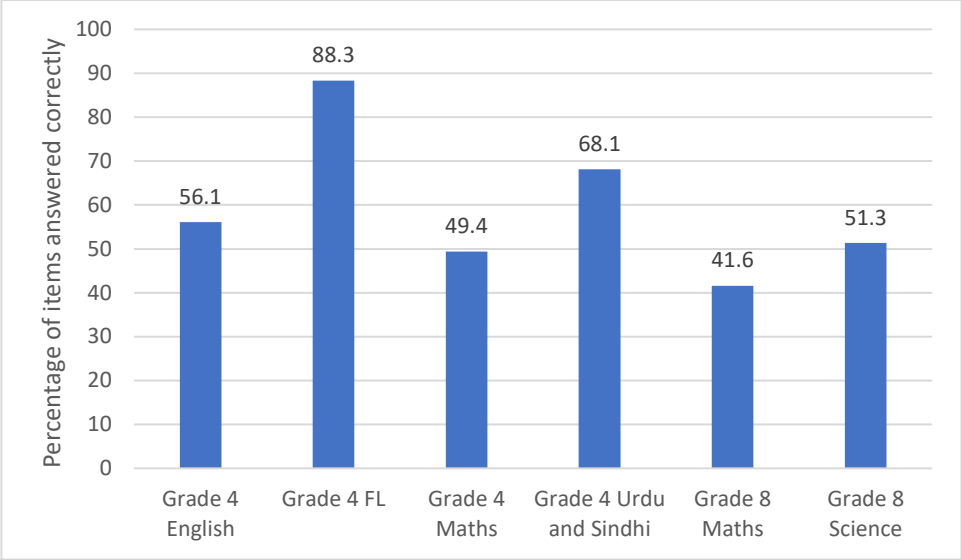


Figure 1: Average student scores as percentage of maximum for NAT 2023, Grade 4 and Grade 8, by subject

Each test consisted of four-option multiple choice questions. As such, on average, a student guessing at random would answer 25% of the items correctly. With this in mind, further inspection of the score distributions revealed that, in several subjects, many students struggled to achieve scores any better than might be achieved by guessing. In particular, approximately one in seven students in Grade 4 English, and one in six in both Grade 4 and Grade 8 Maths, answered no more than 25% of items correctly.

This is of concern because it indicates a noticeable proportion of students who either did not make any effort when completing the assessments or who genuinely lack all the skills needed to engage with them.

How does performance this year compare to the previous NAT?

Grade 4 English and Grade 8 Maths contained anchor items from NAT 2019[1]. However, these comprised less than the 20% of the test, as had been recommended by the American Institute for Research (AIR) after the 2019 NAT. As such, it is only possible to make tentative suggestions regarding overall changes in performance since NAT 2019.

In very broad terms, performance on the anchor items suggested the overall difficulty of tests in NAT 2023 was similar to NAT 2019. On this basis, we can get a rough sense of changes in performance by comparing the percentage of the total available marks achieved in each subject in NAT 2019 and NAT 2023. This comparison is shown in Figure 2[2].

[1] Note that, although we label the earlier test as "NAT 2019", for consistency with earlier reports, it was actually taken by students in autumn 2020.

[2] Values for NAT 2019 come from a draft version of the National Assessment Report 2019, which was provided to assist with preparation of this report.

Analysis suggests that Grade 4 English performance has improved since NAT 2019. Students answered 56% of items correctly in NAT 2023 compared to only achieving 51% of the available maximum in NAT 2019. This is equivalent to an improvement of 0.21 standard deviations. However, the underlying assumption in this comparison that the tests were of roughly equal difficulty is based on just five anchor items, all of which were in the same content domain. Ideally, anchor items would cover the same content and difficulty range as each of the full forms being equated [3].

Changes in Grade 8 Maths performance were inconsistent across the six anchor items that were available. Overall, there is no strong evidence of any change in performance, with students achieving 42% of the available marks in NAT 2023 compared to 43% in NAT 2019.

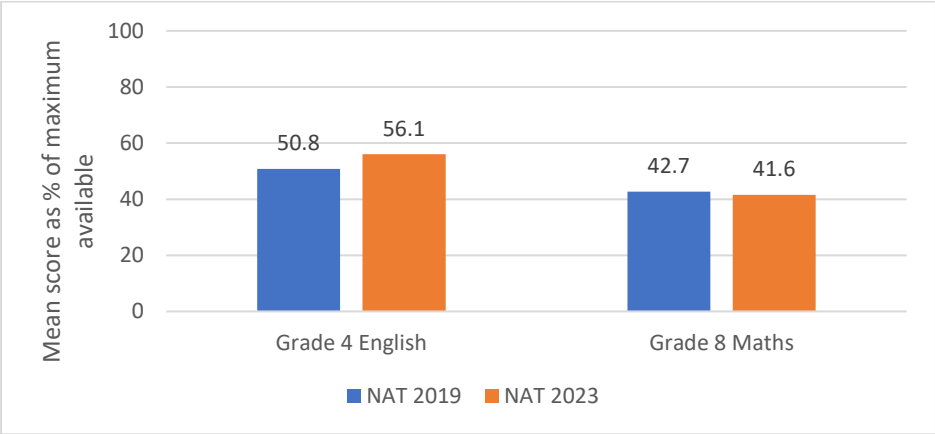


Figure 2: A comparison of mean scores (as a percentage of maximum available) between NAT 2019 and NAT 2023 for Grade 4 English and Grade 8 Maths

[3] For example, see page 98 of the 2014 Standards for Educational and Psychological Testing (https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf).

Key terminology

Anchor items: Questions that are included in multiple test versions to allow calibration of the difficulties of the various tests against each other and equating of test scores.

Equating: Equating is a statistical process by which scores on different tests are transformed to a common scale that accounts for the difficulty of the test items. When scores have been equated, we can easily identify which of two students has performed better even if they have done different tests. If scores have not been equated, we need to be more cautious in making comparisons.

Standard deviation: A number used in statistical analysis that shows the amount by which members of a group are different from the mean (otherwise known as the average) value for the group.

Statistical significance: A difference between groups is statistically significant if it is large enough that it is unlikely to have occurred purely due to random sampling. The probability of a difference occurring by chance in this way is represented by a **p-value**. A low p-value (e.g., below 0.05) indicates a statistically significant difference.

Variations in performance across provinces

Performance in NAT 2023 in each assessment is shown for each province in

Figure 3 (Grade 4 English and Foundational Literacy), Figure 4 (Grade 4 Maths and Urdu and Sindhi) and Figure 5 (Grade 8 Maths and Science). Each chart shows the percentage of items answered correctly in each assessment in each province. In each chart, provinces are sorted from lowest to highest performance.

In every subject except Grade 4 Foundational Literacy, Punjab achieved the highest level of performance and Sindh Province saw the second highest level of performance. In Foundational Literacy, most provinces displayed a very high level of performance, with the highest average score displayed by Islamabad Capital Territory (ICT). Performance in Foundational Literacy was noticeably lower in Sindh than in other provinces. This may potentially indicate that, for this assessment, the translation of the assessment from Urdu to Sindhi increased the difficulty of the items.

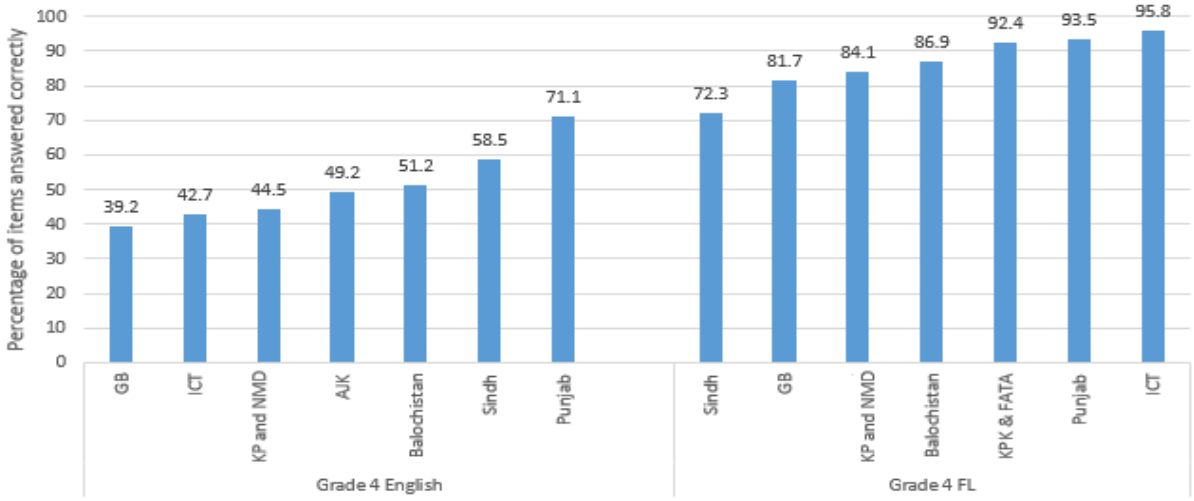


Figure 3: Performance in Grade 4 in NAT 2023 in English and Foundational Literacy

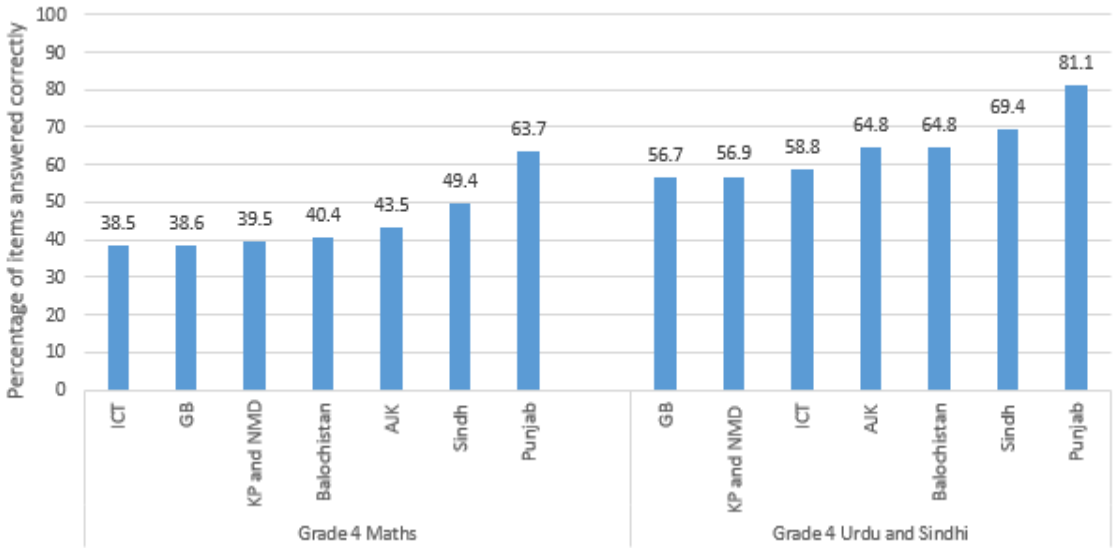


Figure 4: Performance in Grade 4 in NAT 2023 in Maths, and Urdu and Sindhi

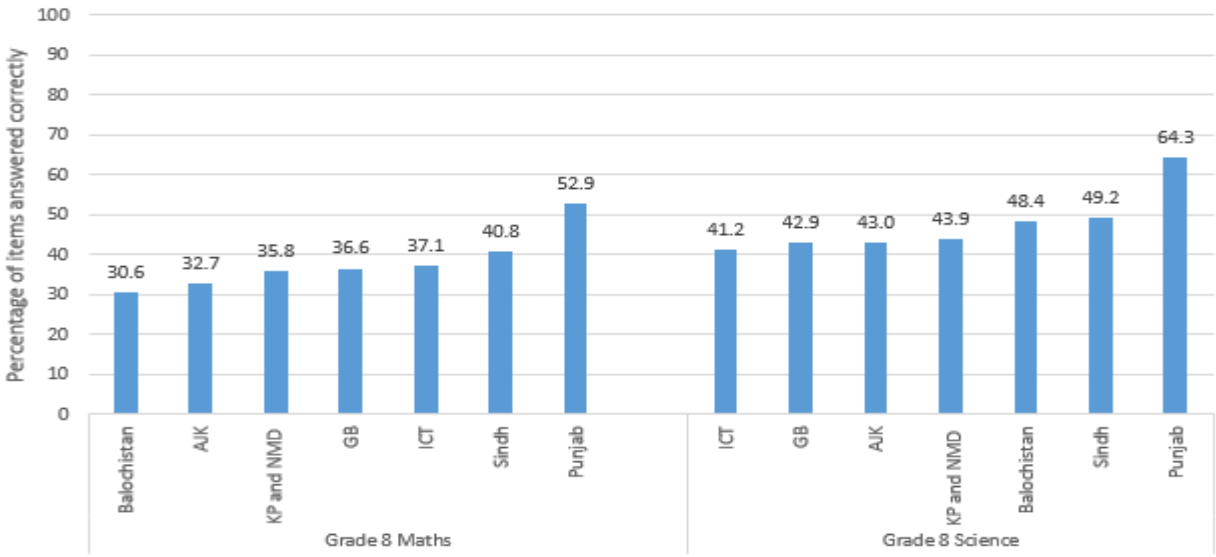


Figure 5: Performance in Grade 8 in NAT 2023 in Maths and Science

Figure 6 shows changes in performance in each province in Grade 4 English and Grade 8 Maths since NAT 2019. Across provinces, the largest improvement in Grade 4 English scores was seen in Punjab. Although Punjab was not the highest performing province in this subject in NAT 2019, it is now outperforming all other provinces. Further analysis linked performance in Grade 4 to teachers stating that the course 'always' ends in time (see below). The fact that teachers in Punjab were the most likely to report this may partially explain the high performance in this province.

In Grade 8 Maths, the pattern of results across provinces in NAT 2023 is consistent with that displayed in NAT 2019.

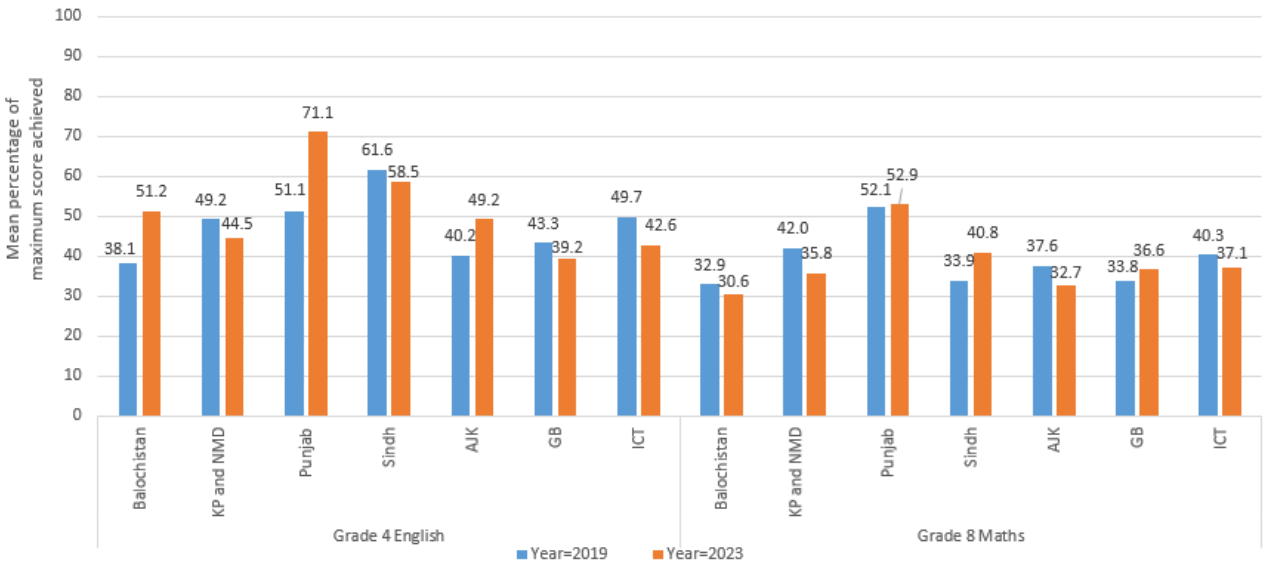


Figure 6: Comparison of performance in Grade 8 Maths and Grade 4 English between NAT 2019 and NAT 2023

Variations in performance across subgroups

A comparison of the average performance of boys and girls is shown in Figure 7. Across all subjects, girls achieved statistically significantly higher scores than boys, except in Maths, where girls and boys achieved similar scores on average. There was no statistically significant variation in the gender gaps across different provinces.

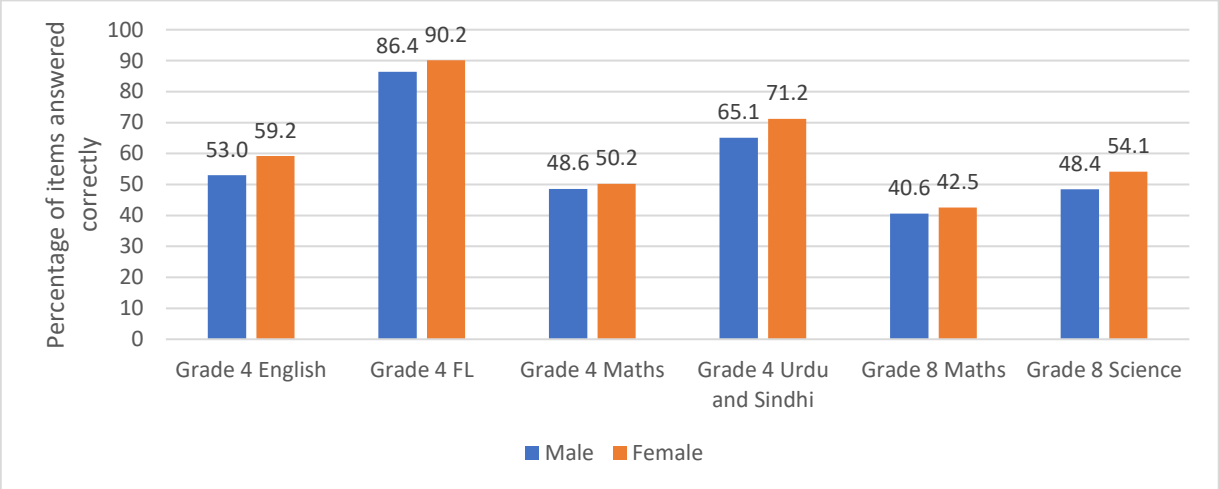


Figure 7: A comparison of performance across boys and girls in NAT 2023

A comparison of average performance of urban and rural areas is shown in Figure 8. These comparisons mostly revealed no significant differences in performance. An exception was in Grade 8 Maths, where rural areas achieved slightly higher scores on average. Further analysis also revealed few significant differences between urban and rural areas within individual provinces. However, the proportion of students performing no better than would be achieved on average through random guessing was notably high in Grade 4 Maths in rural Balochistan (40% of students).

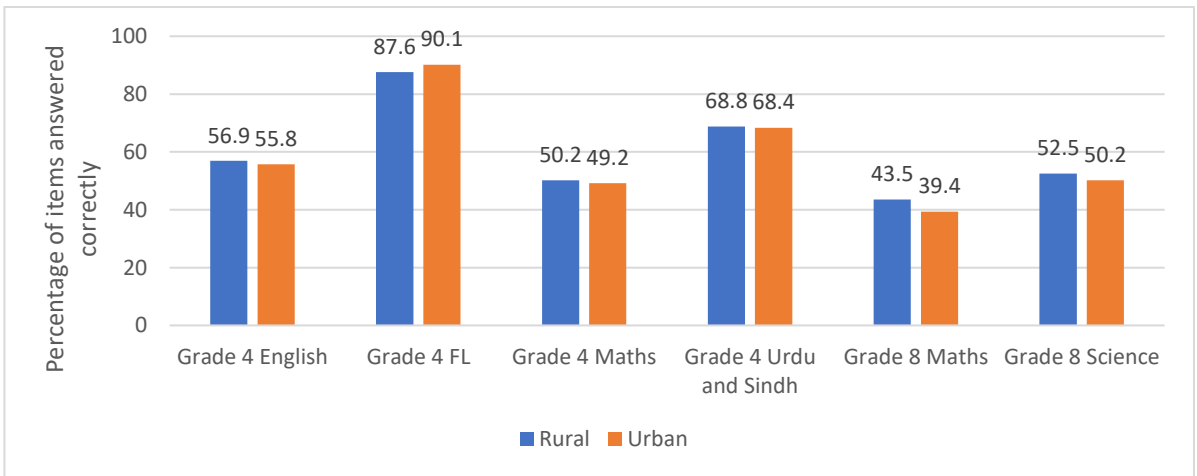


Figure 8: A comparison of performance in urban and rural areas in NAT 2023

Comparison of student and teacher performances

In every school where students took assessments, their teacher was also invited to participate. Figure 9 compares the performances of students and teachers. As expected, on average, teachers achieved much higher scores than students. More interestingly, further analysis revealed a statistically significant link between teacher and student performance: within individual schools it was very rare for the average performance of students in a school to noticeably exceed the performance of their teacher.

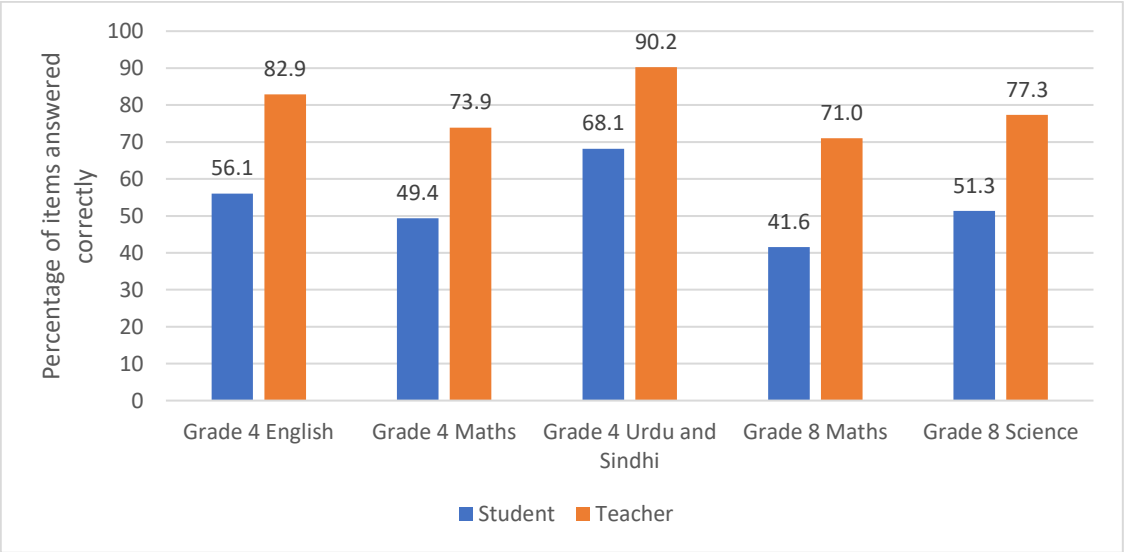


Figure 9: A comparison of student and teacher performance in NAT 2023

Variation in performance across content domains

Within any given assessment, there was little obvious variation in performance across cognitive and content domains. Figure 10 (focused on Grade 4) and Figure 11 (focused on Grade 8) display the average performance of students and teachers in each content domain. Note that judging the statistical significance of differences is tricky and needs to account for the number of items in each domain as well as variation in performance across them (details of how this was addressed are in the main report).

Figure 11 shows that students and teachers both answered a smaller proportion of items correctly in the Grade 8 Maths content domain of statistics and probability. This may relate to this content domain being more recently introduced to the curriculum. The results also showed that the performance of students and teachers in Grade 8 tended to be better in Life Sciences than in either Earth and Space Sciences or Physical Sciences.

Figure 10 also shows that, compared to other content domains, students answered fewer Grammar items correctly in the Grade 4 assessment of Urdu and Sindhi.

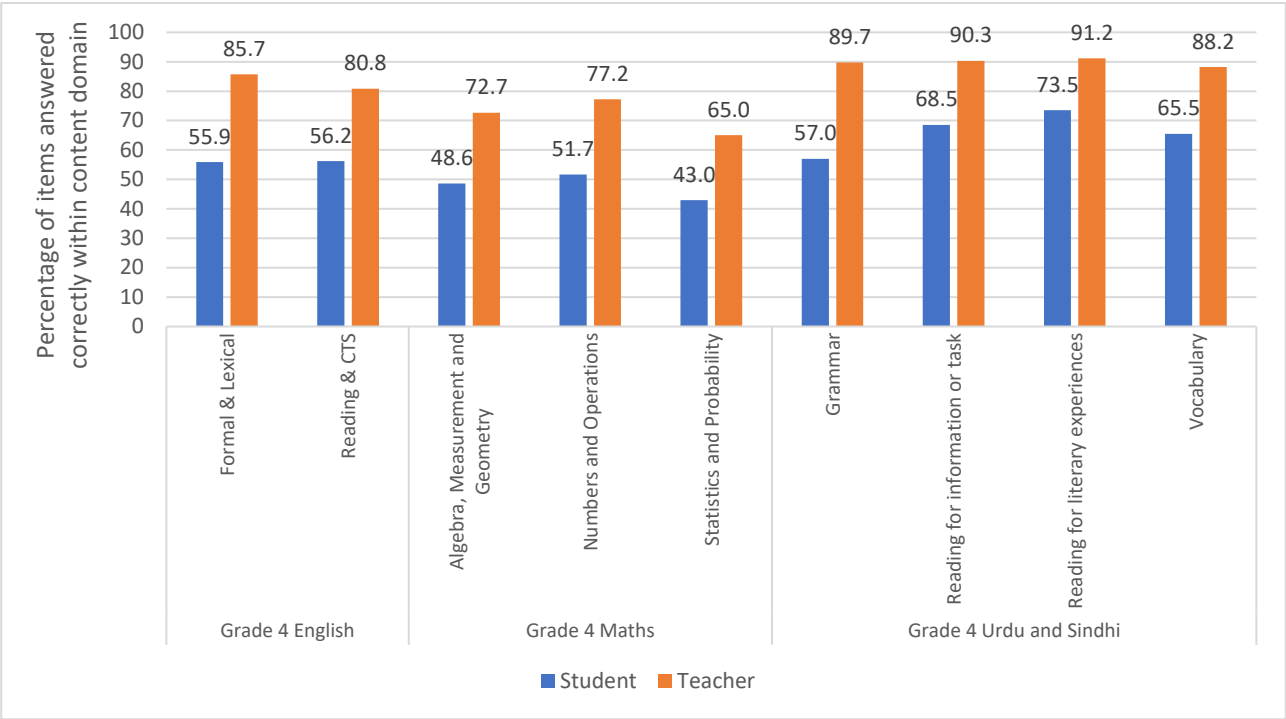


Figure 10: Performance of students and teachers across different content domains in Grade 4 assessments

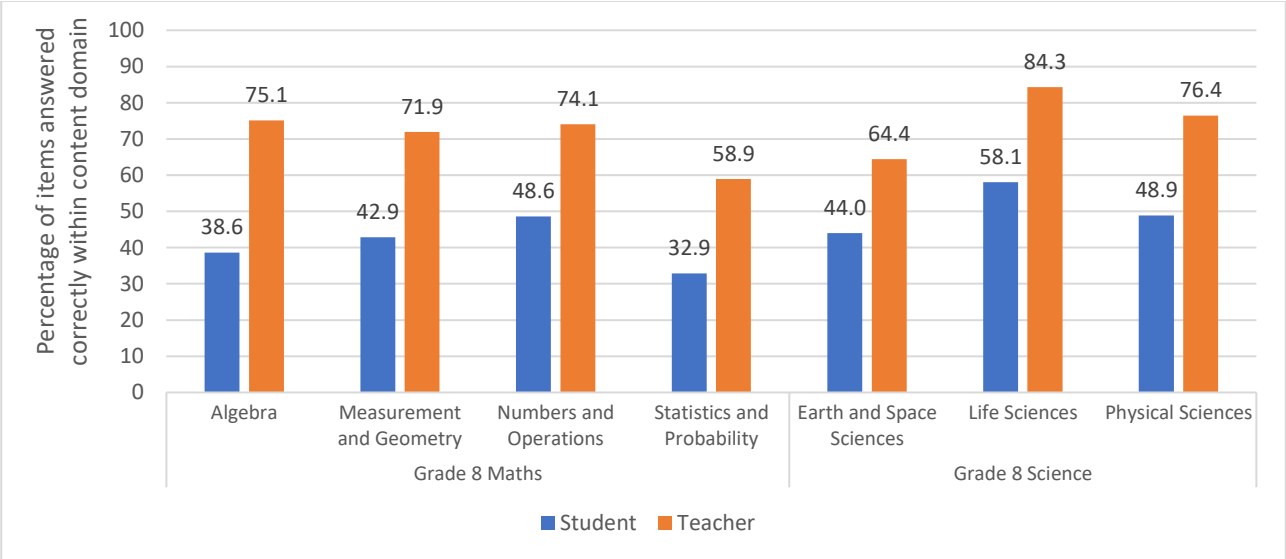


Figure 11: Performance of students and teachers across different content domains in Grade 8 assessments

Reasons for strong student performance

We reviewed the associations between what teachers, students and parents said in questionnaires and students' performances on the assessments. Several factors emerged that were significantly associated with student performance. The strongest associations are described below. Although caution is advised

in assuming these relationships are causal, it is worth pointing out that these factors align with international research findings for high-quality teaching and reflect teaching ideals that were not always achieved.

- **Completion of the course**: In schools where teachers reported that the course was always completed in time, students achieved higher test scores on average (see Figure 12) and were far less likely to have scores at or below a level that would be expected by guessing. Note that the importance of course completion was also highlighted after NAT 2016[4].
- **Homework**: The assignment, completion and checking of homework by teachers consistently emerged as being significantly associated with attainment. Figure 13 shows this relationship in relation to Grade 4 Maths. Note that the importance of homework being assigned, checked, and corrected was also noted in reporting on NAT 2016.
- **Language of instruction**: Students tended to achieve higher scores in English when lessons were taught in their local or mother tongue for at least some of the time (see Figure 14).
- **Self-expression and confidence**: Students who felt they could express their ideas in class and students whose parents stated they had self-confidence tended to achieve higher scores (see Figure 15).

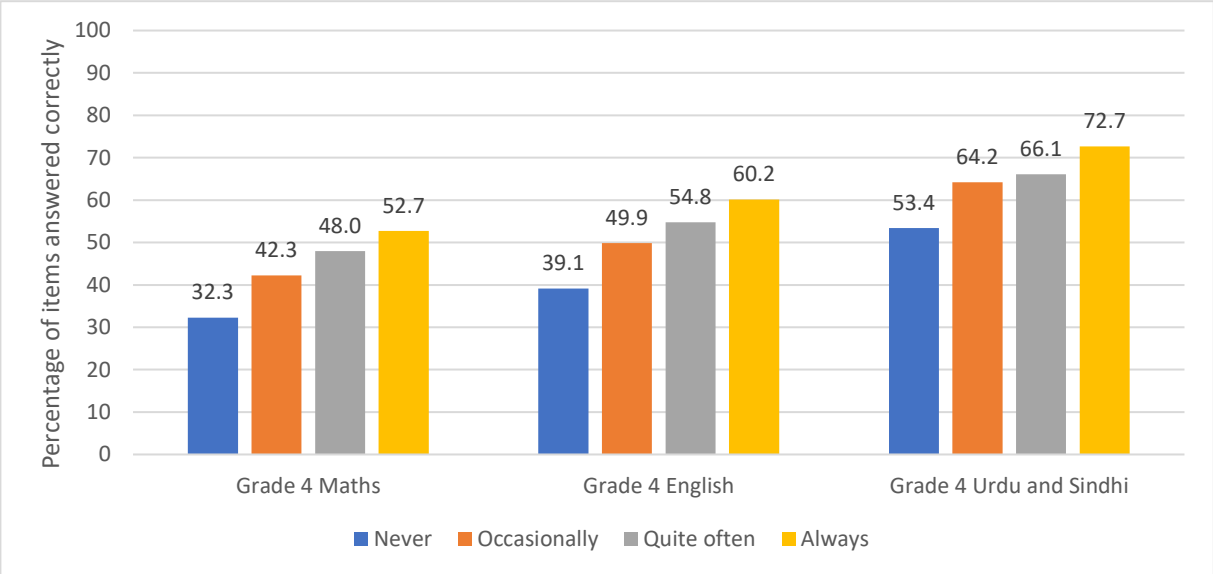


Figure 12: Performance of Grade 4 students in each subject split by the extent to which their teacher agrees that "the course ends in time"

[4] See Dissemination of National Achievement Test Findings 2016–2017, Fifth Stakeholders Conference, National Education Assessment System Ministry of Federal Education and Professional Training Islamabad: https://allchildrenlearning.org/wp-content/uploads/2020/01/Presentation-disseminating-findings-from-Nat-2016-17.pdf.

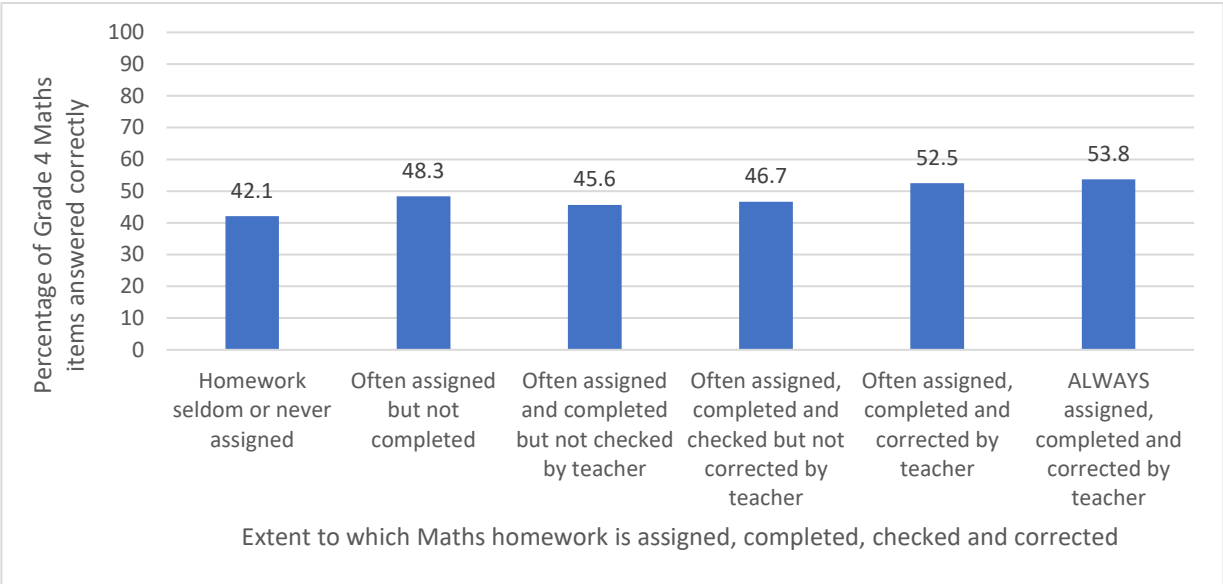


Figure 13: Performance of students in Grade 4 Maths by the extent to which they say homework is assigned, completed, checked, and corrected

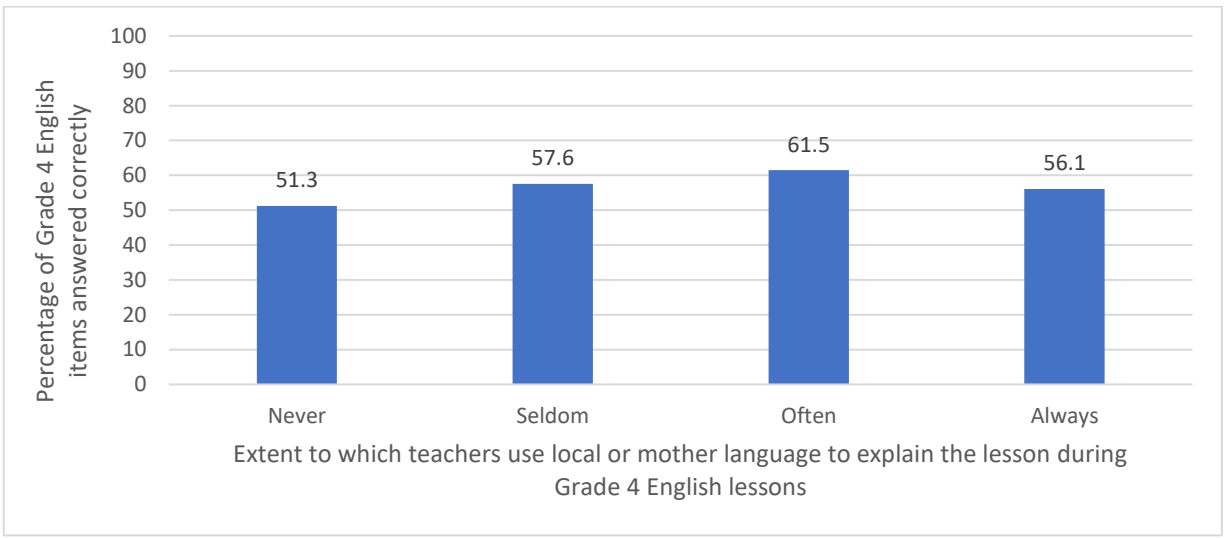


Figure 14: Performance of students in Grade 4 English by the extent to which teachers use local or mother language in English lessons

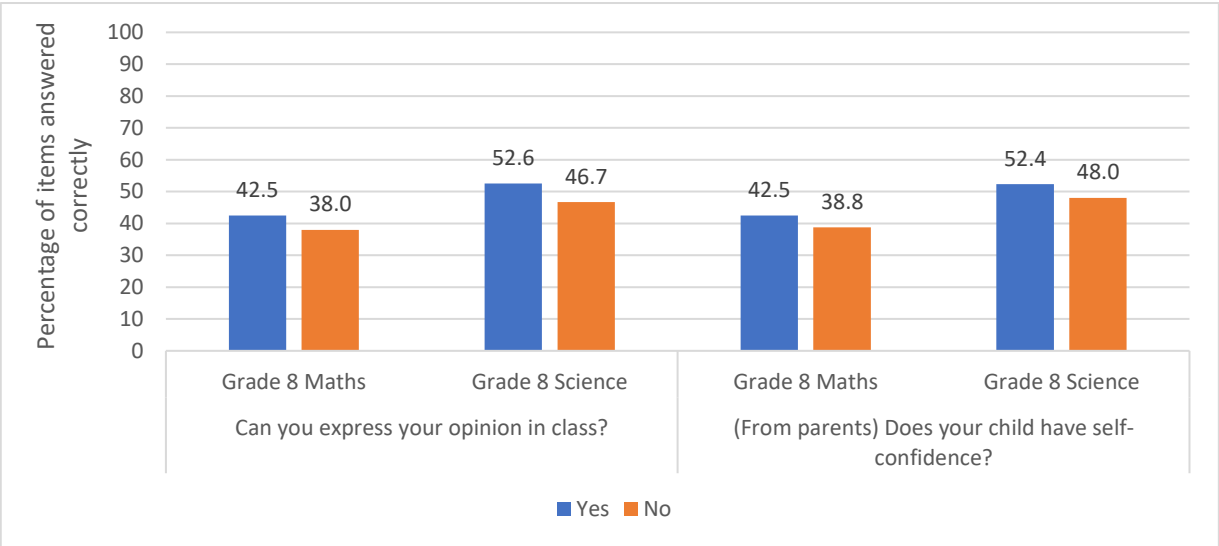


Figure 15: Performance of students in Grade 8 Maths and Science by whether students report that they can express their opinion in class and whether parents report that their child has self-confidence

Teaching ideals

Know it: Students rarely performed better than their own teacher. This emphasises the importance of every teacher having a full understanding of the content that they are teaching.

Teach it: There is a very strong association between the extent to which teachers stated that "The course ends in time" and student performance. Although the importance of completing teaching is obvious, many teachers report that this is not achieved.

Check it: Homework consistently displayed a highly significant association with performance. Using homework to check that students have understood what is being taught and that misunderstandings are dealt with is an important part of teaching. Also, there is a significant association between periodic assessment for children and their achievement in the NAT.

Summary of recommendations

Preparing for NAT 2025:

1. Increase awareness of the importance of the NAT for national monitoring and policy decisions. It is important to emphasise that, whilst the tests are low stakes for schools and students, they are nevertheless very important for monitoring purposes. It is critical that student motivation and effort is increased to help reduce the proportion of students scoring no better than guessing and give a truer impression of the educational landscape.
2. Collect a wider sample of data in NAT 2025, to include more schools and extend to those in the private sector.
3. Reflect on lessons learned from the limitations of the 2023 NAT and implement changes to the 2025 NAT questionnaires.

Curriculum:

4. Research barriers that teachers and schools face in completing courses of learning.
5. Provide additional support for teachers and schools to administer homework through, for example, providing teacher training and support materials such as the development of a homework policy and homework workbooks.
6. Provide additional support for teachers and schools to use formative and summative assessment effectively, such as with the development of a rigorous school assessment policy.
7. Encourage the use of the students' mother tongue in English lessons to clarify concepts and help understanding.
8. Support children's reading skills through developing library provision, increasing the availability of books, and implementing or increasing library lessons.
9. Research the extent to which the 2022 National Curriculum has been implemented in schools.

Teacher preparation, recruitment, and professional development:

10. Provide school leaders with guidance on interpreting the results of the NAT and how recommendations can be applied to schools.
11. Provide further support for teachers on the new aspects of the National Curriculum.
12. Understand teachers' needs and identify any gaps in their subject knowledge by carrying out skills audits that can be used to plan teacher training and support.

Parental involvement and student engagement recommendations:

13. Increase parents' involvement in their child's schooling – for example, by encouraging and facilitating regular parent-teacher meetings, curriculum newsletters, guidance on supporting their child, and workshops on literacy and numeracy for parents.

1. Introduction

Background information

Access to quality education is the right of every child in Pakistan. The Government of Pakistan realises the importance of investment in human resource development as a key element for national development. As per Article 25-A of the Constitution of Pakistan, "[The] State shall provide free and compulsory education to all children of the age of five to sixteen years".

In line with the United Nations Sustainable Development Goal (SDG) 4, which sets 10 international targets for ensuring inclusive and equitable quality education by 2030, Pakistan has declared that improving the quality of education is a priority at the national level. The SDG targets include eliminating all discrimination in education, as well as providing universal literacy and numeracy, increasing the supply of qualified teachers, and building and upgrading safe schools.

The Government is working on various initiatives to provide quality education to its citizens. One of the interventions that is designed to support the improvement of quality education in Pakistan is the National Achievement Test (NAT), which is a sample-based national assessment that aims to provide stakeholders with a reliable means of evaluating student performance. From the results of the NAT, evidence-based recommendations can be made which can then be employed by stakeholders to help to provide the quality of education desired. Cambridge University Press & Assessment were contracted by the World Bank to provide the National Assessment Wing (NAW) of the Pakistan Institute of Education with technical assistance to develop and deliver the 2023 NAT and then analyse the data, report on the findings, and provide recommendations.

Purpose of the report

This report was produced to document the 2023 NAT for the benefit of, and use by, current and future stakeholders. This documentation covers the development and administration of the NAT, key findings from the tests and the questionnaires, associations with student performance and recommendations for policy and practice to try to address issues identified or gather more data. It is hoped that this report will provide useful information at both National and Provincial level to guide decision making and prioritisation.

Report organisation

Due to the volume of information to be shared, it is necessary for this report to be lengthy. To assist with navigating through it, the report is split into sections. A brief summary of these sections follows.

- **The National Achievement Test (NAT) –** this section focuses on the purpose, aims and objectives of the NAT. The need for a robust large-scale assessment such as the NAT is also discussed.
- **Development of the NAT** – this section provides an overview of how the tests for the 2023 NAT were produced. The role of NAW, provincial representatives and Cambridge are discussed. The sampling design is explained, along with a description of how the final papers were produced via a piloting phase.

- **Student achievement** – after an overview of the main findings, this section reports on the performance of students by province location and gender, before finally comparing the performance of students versus their teachers.
- **Student achievement in different content and cognitive domains** – after an overview of the main findings, this section reports on the performance of students in the different content and cognitive demands in each of the subjects.
- **Analysis of factors associated with student performance** – after an overview of the main findings, this section discusses the initial exploration of relationships between student performance and factors commonly of interest (teacher qualification, parent education, possessions in the home and student attendance). Focus then shifts to those factors which show the strongest relationships with student performance (teacher vs student performance and aspects of the student and parental questionnaires with the most significant associations). The main factors associated with student learning are finally summarised in section 6.5.
- **Policy recommendations** – this section focuses on the findings of the 2023 NAT and the implications of this for policy and practice in Pakistan. A series of recommendations is presented under key areas of focus (curriculum and instruction; teacher preparation, recruitment and professional development; parental involvement and student engagement).
- **Limitations of the 2023 NAT and recommendations for future iterations** – whilst acknowledging the great success that was the 2023 NAT and reinforcing that this should be celebrated, there were compromises and limitations. This section summarises the key limitations of the current study (relating to sampling, Foundational Literacy items, background questionnaires and engagement with the NAT), with suggestions made as to how to remedy these for the 2025 iteration of the NAT.
- **Appendices** – the detailed appendices cover the NAT methodology and subject frameworks, the data collection procedures, specifics of the psychometric analysis (methodology and results), responses to the background questionnaires and finally a summary of results for each province.

2. The National Achievement Test (NAT)

This section focuses on the purpose, aims and objectives of the NAT. The need for a robust large-scale assessment like the NAT is also discussed.

2.1. The purpose of the NAT

The National Achievement Test (NAT) is a sample-based assessment that takes place approximately every two years in a range of subjects (English, Urdu, Sindhi, Maths and Science) at Grade 4 and Grade 8 in Pakistan. NAT 2023 is the eighth assessment cycle of this national large-scale cross-sectional assessment survey. A total of 1304 schools (652 schools for Grade 4 and 652 schools for Grade 8) in four provinces, two areas and the Islamabad Capital Territory (ICT) were selected for inclusion in the study.

The purpose of the NAT is to provide stakeholders with a reliable means of evaluating student attainment. The data are used to inform decision makers, to measure the impact of teaching and learning, to evaluate the efficacy of educational policies and reforms, and to determine the value for money of investments in training and resources.

The NAT is a low-stakes assessment for students and teachers, providing high-quality, data-driven information that is designed to improve the quality of education in Pakistan.

Greaney and Kellaghan (2008[5]) state that all large-scale assessments seek answers to one or more of the following questions:

- How well are students learning in this education system? Are they meeting specific learning standards?
- Are there particular strengths and weaknesses in student knowledge and skills?
- Do some subgroups perform worse than others? Are there disparities, for example, between the performance of boys and girls, students in urban and rural locations, or students from different language groups?
- What factors are associated with student achievement? To what extent does achievement vary according to the characteristics of the learning environment (for example, school resources or teacher preparation) or according to students' home circumstances?
- Does student achievement change over time? What factors are linked to changes in student achievement over time?

2.2. The need for robust large-scale assessments

The role of Pakistan's National Assessment Wing (NAW) in developing and executing robust assessment studies is critical given the severe learning challenges in Pakistan, which has a high incidence of learning poverty due to a range of structural factors. The total out-of-school population in Pakistan is the second highest in the world, with an estimated 22.8 million children aged 5 to 16 not attending school, according

[5] Greaney and Kellaghan (2008). Quoted in Clarke, M. and Luna-Bazaldua, D. (2021) *Primer on Large-Scale Assessments of Educational Achievement*, World Bank.

to the United Nations Children's Fund (UNICEF)[6]. They represent 44% of the total population in this age group (Pakistan Social and Living Measurement (PSLM), 2018–19).

The Government's 2021–22 Pakistan Economic Survey estimates that, from 2020–21, there were 14.4 million students enrolled in pre-primary education, 25.7 million in primary education (Grades 1 to 5) and approximately 8.3 million in middle education (Grades 6 to 8). In this context, it is important to ensure continued improvement in the assessment system so that it informs remedial policies and actions that are beneficial for the teaching and learning landscape in the country.

Rigorous and periodic assessments of student learning levels are necessary to provide policy makers and other stakeholders with information about the impact of the resources allocated to education on student learning. A national assessment provides information about the quality of student learning with reference to the national curriculum, the implementation of the curriculum, public perceptions about what students should be able to do, and whether or not students are properly prepared for future life. When compared with total expenditure on education, a national assessment is a relatively inexpensive complement to reform efforts to improve learning.

Assessment data can be used to monitor change in achievement over time. Reliable and valid data is necessary to answer the question, "Is the quality of the education system, in terms of learning outcomes, improving?".

If policy makers and other stakeholders do not know how successful (or unsuccessful) schools are in transforming resources into student learning, they risk maintaining suboptimal educational environments.

Policy makers and other stakeholders can draw on the empirical data from the assessments to make informed decisions about the allocation of resources. For example, a national assessment can identify areas of the curriculum where a considerable proportion of students are underachieving. This underachievement may also be associated with specific factors, such as location, type of school and medium of instruction. Subsequent action may involve the provision of in-service courses for teachers or additional resources and materials allocated to schools in specific categories.

2.3. Aims and objectives of the NAT

The 2023 NAT aims to:

- provide a snapshot of learning achievement in Grade 4 and Grade 8 for a range of content domains in five subjects (English reading, Urdu reading, Sindhi reading, Maths and Science)
- establish a systematic way of developing, implementing and using assessments for strengthening the quality of teaching and learning
- enable the bridging of information gaps by providing a platform to all stakeholders for discussion and use of assessment results for improved practices
- guide policy making in areas such as teaching and learning, resource allocation and curriculum development and design

[6] According to https://www.unicef.org/pakistan/education (accessed on 22 November 2023).

- enable the comparison of the quality and effectiveness of education systems at national, regional and international levels, such as the Global Proficiency Framework (GPF)
- enable the comparison of student performance over time through the regular administration of the NAT with a nationally representative stratified random sample of schools, students and teachers
- look at trends in student performance over time which may help to identify any potential learning loss due to COVID 19 or the floods in 2022
- model best practices of assessments appropriate to the context of Pakistan and outline the process for the collection of information on student performance aligned to international standards. This will enable reporting against Sustainable Development Goal (SDG) 4.1 (**Target 4.1:** By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes. **Indicator 4.1.1:** Proportion of children and young people (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) Maths, by sex).

3. Development of the NAT

Section 3 provides an overview of how the tests for the 2023 NAT were produced. The role of NAW, provincial representatives and Cambridge are discussed. The sampling design is explained, along with a description of how the final papers were produced via a piloting phase.

3.1. The role of NAW in the design and delivery of the NAT

The NAT is the largest assessment activity that Pakistan's National Assessment Wing (NAW) carries out, approximately every two years, with support from provincial and area education departments. The NAT assesses Grade 4 and 8 student performances in the core subjects of Maths, General Science and languages across the country. The study uses scientific sampling techniques to ensure that the results are representative at the provincial and national level, and for specified strata of the total population.

Subject and assessment experts thoroughly review the assessment tools, including test booklets, background questionnaires and test administrators' manuals. NAW assigns assessment items based on student learning outcomes defined in the curriculum as opposed to the books developed by different textbook boards.

NAW then uses a process of technical validation to finalise the assessment items. Tables of specification (ToS) that outline the weight or percentage allocated to each piece of content and cognitive domain are then used to design and finalise the test booklets.

NAW planned to deliver a NAT in 2021 but this was delayed due to COVID-19. The items from the 2021 NAT were revised and repiloted to test the 2023 NATs for Grade 4 English, Urdu, Sindhi and Maths and Grade 8 Science. The items for Grade 8 Maths were newly commissioned and reviewed.

Table 1 shows the subjects tested in each grade for each round of the NAT so far.

Table 1: Subjects tested in NATs from 2014–2023

Subject	NAT 2014		NAT 2016		NAT 2019		NAT 2023	
	Grade 4	**Grade 8**	**Grade 4**	**Grade 8**	**Grade 4**	**Grade 8**	**Grade 4**	**Grade 8**
English	✓			✓	✓		✓	
Urdu		✓	✓			✓	✓	
Sindhi						✓	✓	
Maths		✓	✓			✓	✓	✓
Science	✓			✓	✓			✓
(Social studies)					✓			

In 2023, 15 items were added for the first time to assess Foundational Literacy and learning poverty in Urdu and Sindhi. These were administered as a separate test to the Grade 4 students following the completion of the live NAT. The purpose of the inclusion of these Foundational Literacy items was to assess basic literacy skills such as letter and word recognition.

3.2. NAT sample design

The sample of schools and students for live tests was drawn by the civil society organisation Idara-e-Taleem-o-Aagahi (ITA). The approach to sampling was based on the methodology used within the Trends in International Maths and Science Study (TIMSS) and was designed to achieve a sample that was representative of students in public schools in Pakistan. This should enable accurate estimates of achievement at a national level.

The sample was drawn in two stages. To begin with, a sample of schools was chosen from the National Education Management Information System (NEMIS) school directory, based on academic year 2016–17. The NEMIS database is updated every year and contains consolidated data from annual provincial and federal education censuses. It contains information about public sector schools across the country. The school directory from 2016–17 was used as the basis for sampling as it was the most up-to-date and complete data set available to ITA at the time when sampling was required.

The number of schools sampled from each province was chosen to reflect the proportion of students nationally enrolled within each one. The number of schools sampled in 2023 in each grade and each province is shown in the final two columns of Table 2. For comparison, sample sizes from previous NAT studies are also included. Note that all schools sampled in 2023 were public schools. The same number of schools were sampled in both Grade 4 and Grade 8. Note that schools were included in the sampling process regardless of the size of their enrolment.

As can be seen from Table 2, the sample size for NAT 2023 was somewhat larger than that for NAT 2019 and similar to that used in NAT 2016. The larger sample is useful as it allows a more accurate provincial-level representation of results.

Within each province, the sample was explicitly stratified by school gender and whether the school was in an urban or a rural location. Specifically, within each province, an equal number of schools were sampled within each of the following four categories: rural female, rural male, urban female, urban male. For example, in AJK, in Grade 4, 12 rural female, 12 rural male, 12 urban female and 12 urban male schools were sampled making 48 schools in total. For every sampled school, two replacement schools with the same characteristics were also identified in case the initially sampled schools did not participate.

The sampling of schools was implicitly stratified by school type in the following four categories: Primary, Elementary, High, and Higher Secondary. Implicit stratification works alongside the explicit stratification described above by sorting the sampling frame by variables of interest prior to sampling. When used alongside systematic random sampling, it can help achieve a sample where students are in each category of school type in the same proportions as those found in the population as a whole (see Olson, Martin and Mullis, 2008, page 84)[7]. School sampling was also completed using probability proportional to size

[7] Olson, J. F., Martin, M. O., and Mullis, I. V. (Eds.). (2008). *TIMSS 2007 technical report*. TIMSS and PIRLS International Study Center.

(PPS) in each school to ensure that the final sample was representative in terms of the sizes of schools that student attend.

Table 2: Number of schools sampled in each NAT

Sample size (schools)	NAT 2014		NAT 2016		NAT 2019		NAT 2023	
Province	**Grade 4**	**Grade 8**	**Grade 4**	**Grade 8**	**Grade 4**	**Grade 8**	**Grade 4**	**Grade 8**
AJK	9	9	44	48	14	14	48	48
Balochistan	21	23	59	52	19	20	60	60
Gilgit-Baltistan	11	9	23	26	8	8	32	32
ICT	11	11	28	34	8	8	28	28
KP	32	37	126	128	42	42	156	156
Punjab	88	82	194	214	60	60	188	188
Sindh	30	39	132	117	44	46	140	140
Total	**202**	**210**	**606**	**619**	**195**	**198**	**652**	**652**

Within each grade, the second stage of sampling was to select a single class randomly within each school. Finally, within each of these classes, 20 students were selected from the attendance register using systematic random sampling. In this context, systematic random sampling involves choosing a starting student at random for the sample and then selecting the 19 subsequent students by moving through the remainder of the register in equally sized steps. For example, in a class of 60, we would randomly select one of the first three students in the register as the starting point and select every third student in the register thereafter to make up a total of 20 students in the sample. This approach to sampling ensures that the final sample is evenly distributed across whatever variable is used to sort the attendance register (e.g., first letter of surname). The class teacher was also invited to participate in each assessment.

Table 3 and Table 5 show the actual number of schools that participated in each element of the 2023 NAT directed at students (i.e., the tests and questionnaires) in Grades 4 and 8 respectively. The final column of each table shows the number of schools that participated in any of these elements. As can be seen, these numbers are extremely close to the number of schools that were sampled. Incorporating responses from replacement schools, the unique totals in the tables represent response rates of 98.6% (643 out of 652 in the original sample) and 98.2% (640 out of 652) in Grades 4 and 8 respectively.

Table 4 and Table 6 show the numbers of students that actually participated in each element of the 2023 NAT.

Table 3: Number of schools that participated in each element of the 2023 NAT in Grade 4

Province	Grade 4 English	Grade 4 Foundational Literacy	Grade 4 Maths	Grade 4 Urdu and Sindhi	Grade 4 student and parent questionnaire	Unique total
AJK	48	46	48	48	48	48
Balochistan	54	49	52	53	53	54
Gilgit-Baltistan	32	29	31	32	31	32
ICT	26	23	26	26	26	26
KP	155	140	155	156	156	156
Punjab	186	173	185	186	186	188
Sindh	133	121	136	136	135	139
Total	**634**	**581**	**633**	**637**	**635**	**643**

Table 4: Number of students that participated in each element of the 2023 NAT in Grade 4

Province	Grade 4 English	Grade 4 Foundational Literacy	Grade 4 Maths	Grade 4 Urdu and Sindhi	Grade 4 student and parent questionnaire	Unique total
AJK	594	542	592	582	570	619
Balochistan	896	823	876	913	886	944
Gilgit-Baltistan	514	455	489	505	481	532
ICT	489	418	487	496	487	505
KP	2910	2514	2934	2926	2859	3038
Punjab	3567	3145	3548	3567	3480	3697
Sindh	1621	1457	1652	1673	1615	1738
Total	**10591**	**9354**	**10578**	**10662**	**10378**	**11073**

Table 5: Number of schools that participated in each element of the 2023 NAT in Grade 8

Province	Grade 8 Maths	Grade 8 science	Grade 8 student and parent questionnaire	Unique total
AJK	48	48	48	48
Balochistan	49	51	48	52
Gilgit-Baltistan	32	32	32	32
ICT	24	26	26	26
KP	155	155	156	156
Punjab	188	187	186	188
Sindh	138	136	136	138
Total	**634**	**635**	**632**	**640**

Table 6: Number of students that participated in each element of the 2023 NAT in Grade 8

Province	Grade 8 Maths	Grade 8 Science	Grade 8 student and parent questionnaire	Unique total
AJK	791	828	831	864
Balochistan	877	883	792	907
Gilgit-Baltistan	616	618	615	630
ICT	427	464	471	506
KP	3048	2982	2938	3127
Punjab	3662	3587	3514	3781
Sindh	2461	2437	2397	2568
Total	**11882**	**11799**	**11558**	**12383**

Comparability of sample to NAT 2019

One aim of analysis was to compare performance in NAT 2023 to performance in NAT 2019. Thus, it is important to note that, the way that the samples were constructed is similar enough to allow such comparisons. In particular, the samples in both cycles were designed to select representative numbers of candidates within each province. Furthermore, both samples were designed to include even numbers of boys and girls. One difference between the samples is that in NAT 2019 roughly 60% of sampled schools were from rural areas. In contrast, in NAT 2023, an equal number of schools were selected within rural and urban areas (that is, only 50% of sampled schools were from rural areas). However, since our analysis shows that differences in performance between rural and urban areas are fairly small in any

case (see Figure 8) this issue does not have any meaningful impact upon our analysis comparing performance between NAT 2019 and NAT 2023.

For consistency with NAT 2019 no weighting was applied to data before completing analysis.

Power of the sample for detecting changes to national performance over time

In statistical terminology, "power" refers to the chances of an effect of a given size being detected. In our context, it is about imagining that performance in a subject in the whole national population improved by a certain magnitude between NAT cycles and calculating the probability of this change being seen and detected as statistically significant within our samples.

Ideally, we would like the sample design to be sufficient to detect any changes in performance over time of a magnitude of at least 0.1 standard deviations. With this in mind, it is worth noting that the analysis showed that, on average across the different assessments, the standard error of the mean performance was equivalent to 0.036 (that is 3.6%) of a standard deviation. For example, in grade 8 maths, the mean performance in raw marks was 21.6 and the standard deviation was 9.3 (see later analysis in Table 9). After accounting for the clustering of students within schools (that is, the fact that students in the same school are likely to display similar levels of performance), the standard error of the mean was 0.33 which is 0.036 (=0.33/9.3) of a standard deviation.

If the standard error of the mean in any NAT cycle is 0.036 of a standard deviation, then the standard error of the *difference* between cycles will be 0.051 (calculated by multiplying 0.036 by the square root of 2 as we are comparing two samples). On this basis, we can calculate that if in the whole national population performance improved by 0.1 standard deviations, the chances of this being detected as statistically significant[8] in the comparison of two NAT samples is 50%. That is, we have only a moderate chance of detecting changes in performance of this size. Having said this, our chances of detecting a slightly larger effect size of 0.15 standard deviations would be much higher at 84%.

The relatively high standard errors in analysis are caused by the fact that variation in performance between different schools tends to be higher than variation within schools. This makes achieving a high statistical power difficult without sampling a very large numbers of schools. Sample sizes may be reviewed ahead of NAT 2025 to decide whether the power is sufficient to meet the purposes of the project.

3.3. Development of the live NAT through piloting

The 2023 NAT was developed by revising the 2021 NAT items that had previously been prepared but did not take place due to the COVID-19 pandemic. For Grade 8 Maths, all items had to be developed specifically for the 2023 NAT, as there were no Grade 8 papers written for 2021.

The revision of the items took place in December 2022. This was carried out by NAW, alongside provincial and Cambridge subject specialists. They created two pilot papers for Grade 4 English, Grade 4 Urdu (also translated into Sindhi), Grade 8 Science and Grade 8 Maths (also translated into Urdu and Sindhi).

[8] Assuming we are using a two-sided significance test and looking for significance at the 5% level.

They also created three pilot papers for Grade 4 Maths (also translated into Urdu and Sindhi). These pilot papers were taken by students in January 2023. The numbers of schools and students participating in these pilot studies is shown in Table 7 and Table 8 respectively. After they were marked, the results were analysed to provide item-level data on difficulty and discrimination.

Table 7: Pilot sample size (schools)

Pilot sample size	Number of sample schools					
	Grade 4				**Grade 8**	
	English	Urdu	Sindhi	Maths	Maths	Science
AJK	4	4		4	4	4
Balochistan	6	6		6	5	5
Gilgit-Baltistan	3	2		3	4	4
ICT	4	4		4	4	3
KP	6	6		6	5	5
Punjab	10	10		10	10	10
Sindh	4		4	4	6	6
Total	**37**	**32**	**4**	**37**	**38**	**37**

Table 8: Pilot sample size (students)

Pilot sample size	Number of students in pilot					
	Grade 4				**Grade 8**	
	English	Urdu	Sindhi	Maths	Maths	Science
AJK	61	99		76	80	80
Balochistan	103	103		99	85	95
Gilgit-Baltistan	68	39		44	70	80
ICT	75	75		71	70	60
KP	120	120		120	109	100
Punjab	207	200		200	203	200
Sindh	98	0	75	75	120	120
Total	**732**	**636**	**75**	**685**	**737**	**735**

In March 2023, Cambridge led a Test Construction workshop with NAW staff and provincial representatives to identify the best-performing pilot items to use in the live NATs. This resulted in the production of one test booklet in English for each of Grade 4 English, Maths and Urdu and Grade 8 Maths and Science. The Maths and Science booklets were also produced in Urdu and Sindhi, and the Urdu reading booklet was translated to produce the Sindhi reading test.

The translation of the booklets into Urdu and Sindhi took place at the Test Construction workshop. Subject specialists translated the texts, which were then typeset. Hard-copy checks of the translations were made by both NAW and provincial experts to ensure the authenticity and validity of the translated versions.

4. Student achievement

4.1. Main findings in this section

The results show that:

1. In terms of the proportion of items answered correctly, the highest level of performance was seen in Grade 4 Foundational Literacy (FL) with students answering 88% of items (13.2 out of 15) correctly on average.
2. Students also answered a high proportion of items correctly in Grade 4 Urdu and Sindhi (68% or 35.4 out of 52).
3. Students answered just over half of the items correctly on average in Grade 4 English (56% or 26.9 out of 48) and Grade 8 Science (51% or 26.7 out of 52); they answered slightly less than half of the items correctly in Grade 4 Maths (49% or 23.7 out of 48) and Grade 8 Maths (42% or 21.6 out of 52).
4. Each test consisted of four-option multiple choice questions. As such, on average, a student guessing at random would answer 25% of the items correctly. With this in mind, looking at the score distributions more closely reveals that a segment of students struggled to surpass scores attainable through random guessing. This is of concern as it indicates a noticeable proportion of students who either did not make any effort to complete the assessments or who genuinely lack the skills needed to engage with the test. In particular, around one in seven students performed no better than would be expected by guessing in Grade 4 English; and one in six were at or below this level in both Grade 4 and Grade 8 Maths.
5. Apart from Grade 4 FL, Punjab outperformed other provinces in all subjects, while Sindh had the second highest level of performance. The students performing at or below a level that might be expected from guessing were concentrated in the other provinces.
6. In Grade 4 FL, ICT displayed the highest level of performance, and the lowest average performance was seen within Sindh. The lower performance on this assessment in Sindh may potentially indicate that the translation of the assessment from Urdu to Sindhi had an impact on the difficulty of items.
7. For Grade 4 English, analysis of anchor items suggests that performance in this subject has improved since the 2019 NAT. However, this conclusion is based on analysis of just five anchor items included in both the 2019 and 2023 tests, all of which were in the same content domain. As such, it needs to be treated with some caution.
8. Across provinces, the largest improvement in Grade 4 English scores was in Punjab. Although Punjab was not the highest performing province in this subject in the 2019 NAT, it is now clearly outperforming other provinces.
9. For Grade 8 Maths, analysis of anchor items provided an inconsistent picture. Performance on some items had improved whereas on others it had fallen. Overall, there was no clear evidence of any national changes in average performance levels since the 2019 NAT.
10. The relative performance of different provinces in Grade 8 Maths was consistent with the pattern displayed in the 2019 NAT.

11. Across most assessments, there were no significant differences between urban and rural areas. An exception was in Grade 8 Maths, where rural areas achieved slightly higher scores on average.
12. There were also few significant differences between urban and rural areas within individual provinces. However, the proportion of students performing no better than would be expected by guessing was notably high in rural Balochistan (40% of students).
13. In all subjects except Maths in both Grade 4 and Grade 8, girls achieved significantly higher scores than boys. In Maths, girls and boys achieved similar scores on average.
14. There was no statistically significant variation in the gender gaps across different provinces.
15. On average, teachers achieved much higher scores than students.

4.2. Initial notes on analysis

At the time of writing in November 2023, no policy-linking workshops have taken place and so it is not possible to report results against Global Proficiency Framework (GPF) levels. Instead, there will be a focus on results in terms of raw test scores. This will be sufficient to gather a broad picture of performance as a proportion of the questions students could answer correctly and the number of students performing at extremely low levels.

The analysis also explores patterns of results across provinces, as well as differences between males and females, students and teachers, and between students in urban and rural locations. Subsequent sections will review performance in different content domains and links between student performance and data from the surveys.

The following sections represent the results of analysis of the versions of the data sets provided at the time of writing. It is possible that, after further inspection, some small amendments may be made to the data sets, which may lead to minor adjustments to the precise numbers within tables and figures. However, any such adjustments are not expected to make a major difference to the main findings.

4.3. Overall performance of students

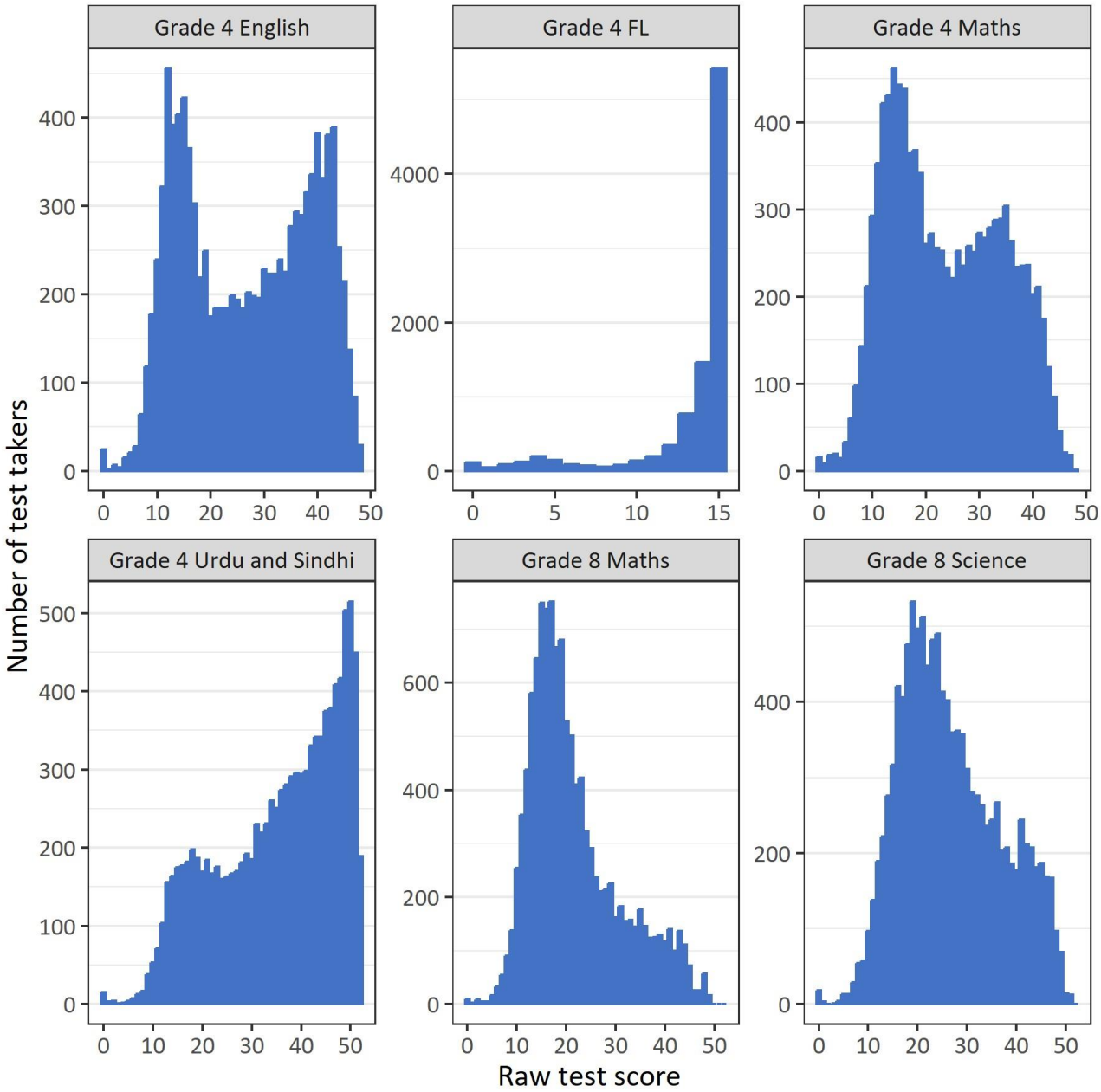


Figure 16 shows the score distribution of students on each assessment. Specifically, the chart shows the number of students achieving each of the possible raw scores out of the total number of marks available in each assessment (that is, the number of items they have answered correctly). Further details about the score distributions[9], including overall means and standard deviations, are provided in Table 9. In most subjects, students answered more than half of the items correctly. In terms of the percentage of items answered correctly, by far the highest level of achievement was in Grade 4 Foundational Literacy

[9] A score distribution means information about how many students achieved each available score. This might be shown in full (e.g., graphically) or we might give a summary of the distribution in terms of the mean and standard deviation of scores.

(FL) followed by Grade 4 Urdu and Sindhi. However, for Maths, in both Grade 4 and Grade 8, students achieved less than 50% on average.

Looking in more detail at the score distributions in the different subjects four distinct patterns can be seen:

- For Grade 4, FL (letter and word recognition) achievement is very high. 58% of the students achieved 15 out of 15.
- For Grade 4 Maths and Grade 4 English, the score distribution is clearly multi-modal (that is, it has more than one peak). One peak in the distribution occurs at the low end of achievement close to (or just above) a score of approximately one quarter of the maximum available marks (25%). This benchmark signifies the average achievement level expected when students make random guesses. Scoring at or below 25% suggests either a very low level of subject knowledge and understanding or a lack of motivation and effort during the test. It is crucial to note that interpreting performance below this benchmark without considering the student's effort level can be misleading (Crooks, Kane and Cohen, 1996[10]). The analysis shows that approximately one in six students achieved no more than a score of 25%. Grade 4 English and Grade 4 Maths also each have a second peak in the score distribution at a much higher performance level. This indicates that a segment of students can accurately answer most of the questions on each test.
- For Grade 4 Urdu and Sindhi, there is a better overall level of performance with many students achieving at or close to the maximum available score. As shown in Table 9, students achieved more than two-thirds of the available maximum score on average (68.1%). However, there is a small peak in the distribution at a point just higher than one quarter of the maximum available score, suggesting that a segment of the students performed no better than would be expected by guessing.
- For Grade 8 Maths and Science, there is only a single peak in the score distribution and this occurs at a fairly low level of performance. In contrast to Grade 4 Maths and English, this is not offset by a second peak at a higher level of performance. This indicates that relatively few students could confidently answer most of the test. The statistics in Table 9 provide further details.
- For Grade 8 Maths, students answered fewer than half of the items correctly on average and approximately one in six students performed no better than would be expected by guessing. Performance in Grade 8 Science was better, with students achieving slightly over 50% on average; only 7% of students performed no better than would be expected by guessing.

[10] Crooks, T.J., Kane, M.T. and Cohen, A.S (1996). Threats to the Valid use of Assessments. *Assessment in Education: Principles, Policy & Practice*, *3*(3), 265–286. https://doi.org/10.1080/0969594960030302

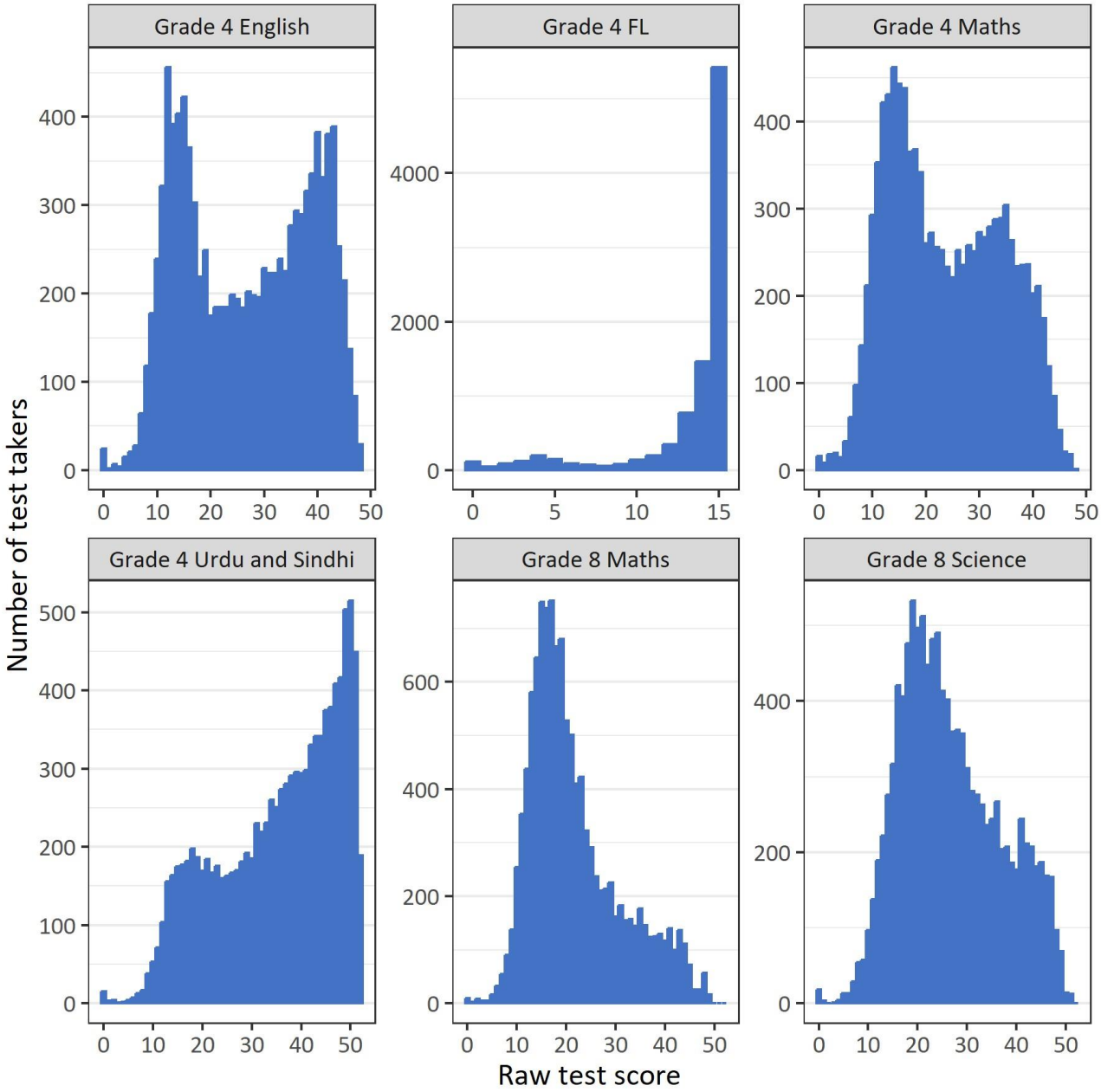


Figure 16: Raw score distributions for students on each test. The height of each bar indicates the number of students achieving each available raw score.

Table 9: Descriptive statistics of raw scores of students on each test

Subject	Max score	Number of students	Mean raw score	Mean score as % of max	Median raw score	Standard deviation (SD) of scores	% achieving at or below 25%
Grade 4 English	48	10591	26.9	56.1	27	12.0	14.0
Grade 4 FL	15	9354	13.2	88.3	15	3.4	4.0
Grade 4 Maths	48	10578	23.7	49.4	22	10.5	16.0
Grade 4 Urdu and Sindhi	52	10662	35.4	68.1	38	12.1	4.6
Grade 8 Maths	52	11882	21.6	41.6	19	9.3	16.7
Grade 8 Science	52	11799	26.7	51.3	25	10.1	7.1

Both Grade 4 English and Grade 8 Maths were tested as part of the 2019 NAT. Table 10 compares the average percentages of marks achieved in 2019 and 2023[11]. Performance in Grade 8 Maths as a proportion of the maximum available marks was similar in 2019 and 2023. For Grade 4 English, students achieved a higher percentage of the available marks on average in 2023 than in 2019. However, in both cases it is important to note that these scores are based on different tests and test scores have not been equated to be on the same scale (see Appendix 3 for further discussion of this issue). As such, the performance comparisons require some caution.

Table 10: Comparison of average percentage of marks achieved by students in 2019 and 2023 (source, 2019 draft report Table 6.1 and Figure 3.1)

Subject	Mean score as % of max		SD of scores as % of max	
	2019	**2023**	**2019**	**2023**
Grade 4 English	50.8	56.1	24.9	24.9
Grade 8 Maths	42.7	41.6	18.8	17.8

To assist with the interpretation of Table 10, Table 11 compares student performance between NAT 2019 and NAT 2023 for the anchor items that were included in both tests. Five items from 2019 were included in the 2023 Grade 4 English test and all were from the content domain 'Reading & Critical Thinking Skills' and the cognitive domain 'Understanding'. Student performance on each of these items in 2023 exceeded performance in 2019 by between six and eight percentage points. This level of difference is fairly consistent with the difference in overall percentage scores of 5.3% (56.1–50.8%) shown in Table 10. Thus, it provides some evidence that the NAT 2019 and NAT 2023 tests are of roughly the same level of difficulty overall and, therefore, that the improvement in the percentage of marks that

[11] Values for NAT 2019 are based on a draft version of the National Assessment Report 2019, which was provided to assist with the preparation of this report.

students achieved represents a genuinely improved performance. On this basis, given the standard deviation in percentage scores of 24.9 (see Table 10) this suggests an overall improvement of 0.21 standard deviations (5.3 ÷ 24.9).

For Grade 8 Maths, six items from the NAT 2019 test booklets were included in the 2023 NAT. These items were drawn from a variety of content and cognitive domains. As shown by Table 11, the changes in performance on these items were more variable. At worst, the proportion of students correctly answering an item fell by five percentage points. At best, the proportion answering an item correctly rose by nine percentage points. The variation in these results leaves us with no strong evidence of any change in overall Maths performance in either direction. Furthermore, if there has been no obvious change in student ability in Maths, and the proportion of items being answered correctly is similar in NAT 2019 and NAT 2023 (see Table 10), this suggests that the two assessments are of broadly equivalent levels of difficulty. Data from future policy-linking workshops may provide further evidence on this matter.

It is worth noting that this suggested overall improvement in Grade 4 English and lack of regression in Grade 8 Maths is a favourable outcome considering the disruptions to learning due to COVID-19 and flooding in 2022.

Table 11: Performance on common items included in both the 2019 and 2023 NAT tests

Subject	Item number in test		Domains		% correctly answered		Difference
	2019	**2023**	**Content**	**Cognitive**	**2019**	**2023**	
Grade 4 English	B1	28	Reading & Critical Thinking Skills (CTS)	Understanding	55%	61%	6%
	B2	29	Reading & CTS	Understanding	52%	60%	8%
	B4	30	Reading & CTS	Understanding	49%	57%	8%
	B5	31	Reading & CTS	Understanding	46%	54%	8%
	B7	32	Reading & CTS	Understanding	47%	53%	6%
Grade 8 Maths	A5	8	Numbers and Operations	Applying	50%	45%	-5%
	A28	14	Algebra	Applying	41%	43%	2%
	C14	16	Statistics and Probability	Knowing	45%	54%	9%
	B25	33	Measurements and Geometry	Reasoning	30%	30%	0%
	A27	43	Measurements and Geometry	Knowing	29%	36%	7%
	B24	50	Numbers and Operations	Applying	25%	26%	1%

4.4. Comparisons in performance by province, location, gender and between students and teachers

Performance by province

Table 12 shows students' performances in each subject within each province. A visual comparison of average performance in each province is provided in the executive summary in

Figure 3, Figure 4 and Figure 5. Figure 17 and Figure 18 show much more detail about the score distributions for the Grade 4 and Grade 8 tests in each province. Specifically, in each province, the bars show the number of students who achieved each available raw score in each test. The results show that Punjab had the highest level of performance in all the assessments. Except for Grade 4 FL, Sindh had the second highest level of performance.

The charts of score distributions show Sindh had a slightly different pattern of achievement in the Grade 4 FL test compared to other provinces. Many more students failed to correctly answer the majority of questions, even though they had been translated into Sindhi.

Some anomalies in the Grade 4 FL data became clear during analysis. In particular, it was clear that, in certain schools, students had uniformly given the same wrong answers to many of the questions in the FL test. Examples of this were found in Sindh and in Gilgit-Baltistan provinces. This explains some of the very low scores that are evident for these provinces in the charts. The reason for these anomalies is unclear but they suggest care should be taken not to overinterpret the patterns of results for Grade 4 FL. For this reason, Grade 4 FL was not included in analysis of subgroups within provinces (e.g., by urban and rural locations within provinces).

As noted earlier, on average, a student guessing at random would answer 25% of the items correctly. The final column of Table 12 uses this benchmark to show the percentage of students in each province performing no better than would be expected by guessing. These percentages tend to be low in Punjab with the highest value being 6.5% for Grade 8 Maths. However, the data suggest that in other provinces, for certain subjects, more than one in four students perform no better than would be expected by guessing. The highest value (36.3% of students performing no better than guessing) is for Grade 8 Maths in Balochistan.

Table 12: Descriptive statistics of raw scores of students on each test in each province

Subject	Maximum score	Province	Number of students	Mean raw score	Mean score as % of max	Median raw score	SD of scores	% achieving at or below 25%
Grade 4 English	48	Balochistan	896	24.6	51.2	25	11.0	16.4
		KP & NMD	2910	21.3	44.5	18	10.5	21.2
		Punjab	3567	34.1	71.1	37	10.2	5.1
		Sindh	1621	28.1	58.5	29	11.8	13.1
		AJK	594	23.6	49.2	21	10.8	14.8
		GB	514	18.8	39.2	16	8.8	24.3
		ICT	489	20.5	42.7	18	9.9	21.7
Grade 4 FL	15	Balochistan	823	13.0	86.9	15	3.7	4.6
		KP & NMD	2514	13.9	92.4	15	2.3	0.8
		Punjab	3145	14.0	93.5	15	2.7	2.1
		Sindh	1457	10.8	72.3	13	4.7	13.3

Subject	Maximum score	Province	Number of students	Mean raw score	Mean score as % of max	Median raw score	SD of scores	% achieving at or below 25%
		AJK	542	12.6	84.1	15	3.9	4.8
		GB	455	12.3	81.7	15	4.5	7.3
		ICT	418	14.4	95.8	15	1.4	0.0
Grade 4 Maths	48	Balochistan	876	19.4	40.4	17	8.8	24.0
		KP & NMD	2934	19.0	39.5	17	9.0	25.5
		Punjab	3548	30.6	63.7	33	9.5	4.8
		Sindh	1652	23.7	49.4	23	9.4	11.8
		AJK	592	20.9	43.5	18	9.8	21.5
		GB	489	18.5	38.6	15	8.8	26.0
		ICT	487	18.5	38.5	16	8.3	23.6
Grade 4 Urdu and Sindhi	52	Balochistan	913	33.7	64.8	35	12.0	4.6
		KP & NMD	2926	29.6	56.9	29	12.3	9.0
		Punjab	3567	42.2	81.1	45	9.2	0.9
		Sindh	1673	36.1	69.4	39	11.6	4.3
		AJK	582	33.7	64.8	36	10.8	4.6
		GB	505	29.5	56.7	30	9.4	3.6
		ICT	496	30.6	58.8	33	10.9	6.9
Grade 8 Maths	52	Balochistan	877	15.9	30.6	15	5.6	36.3
		KP & NMD	3048	18.6	35.8	18	6.5	20.8
		Punjab	3662	27.5	52.9	26	10.4	6.5
		Sindh	2461	21.2	40.8	18	9.0	18.3
		AJK	791	17.0	32.7	16	5.1	22.4
		GB	616	19.0	36.6	18	6.4	18.0
		ICT	427	19.3	37.1	19	5.6	13.6
Grade 8 Science	52	Balochistan	883	25.1	48.4	25	8.2	7.4
		KP & NMD	2982	22.8	43.9	21	8.6	11.1
		Punjab	3587	33.5	64.3	34	9.8	1.7
		Sindh	2437	25.6	49.2	24	9.6	8.7
		AJK	828	22.4	43.0	21	7.0	7.3
		GB	618	22.3	42.9	20	8.6	11.7
		ICT	464	21.4	41.2	21	6.3	8.2

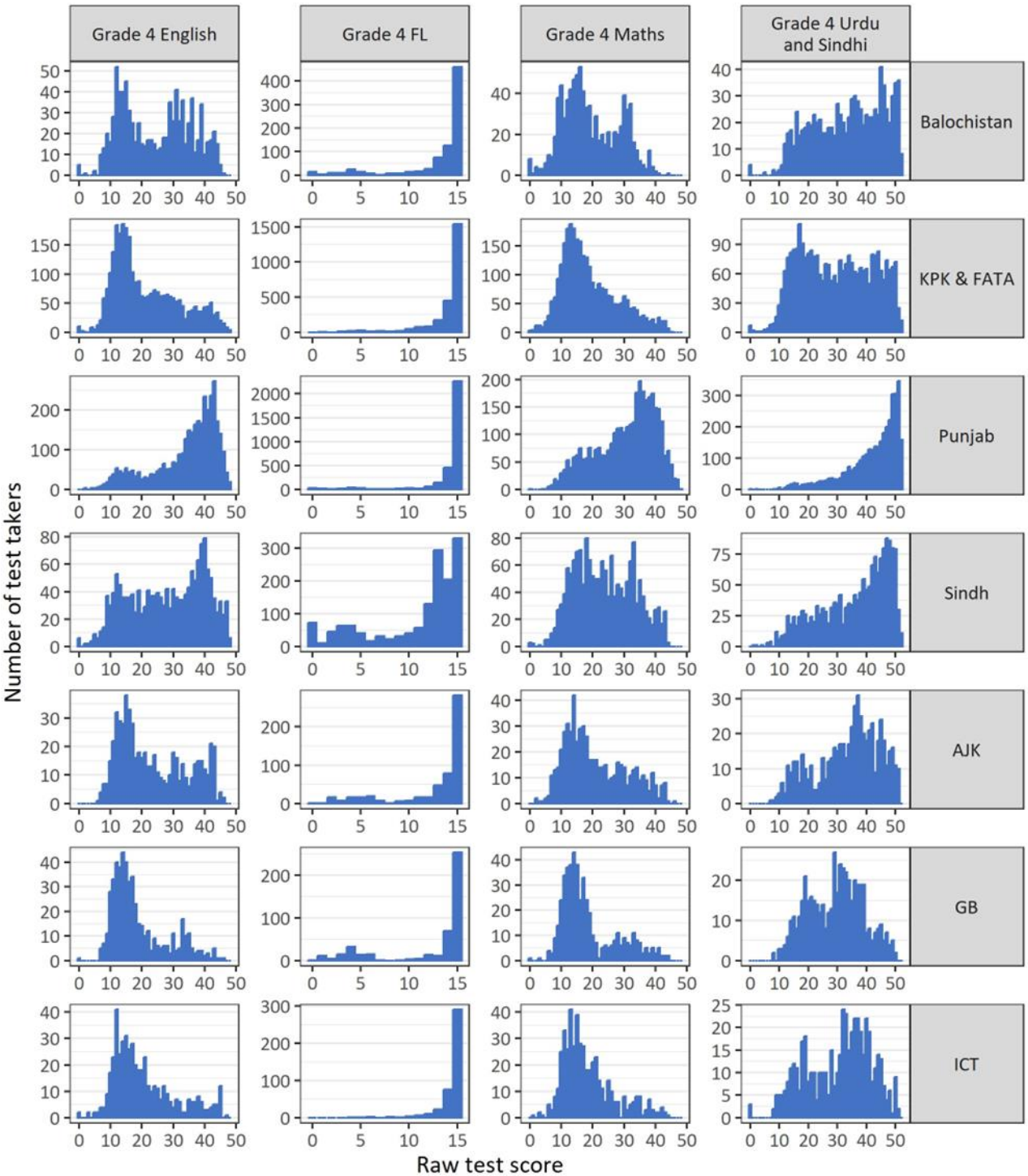


Figure 17: Student score distributions on each Grade 4 test in each province.
The heights of the bars indicate the number of students achieving each available raw score.

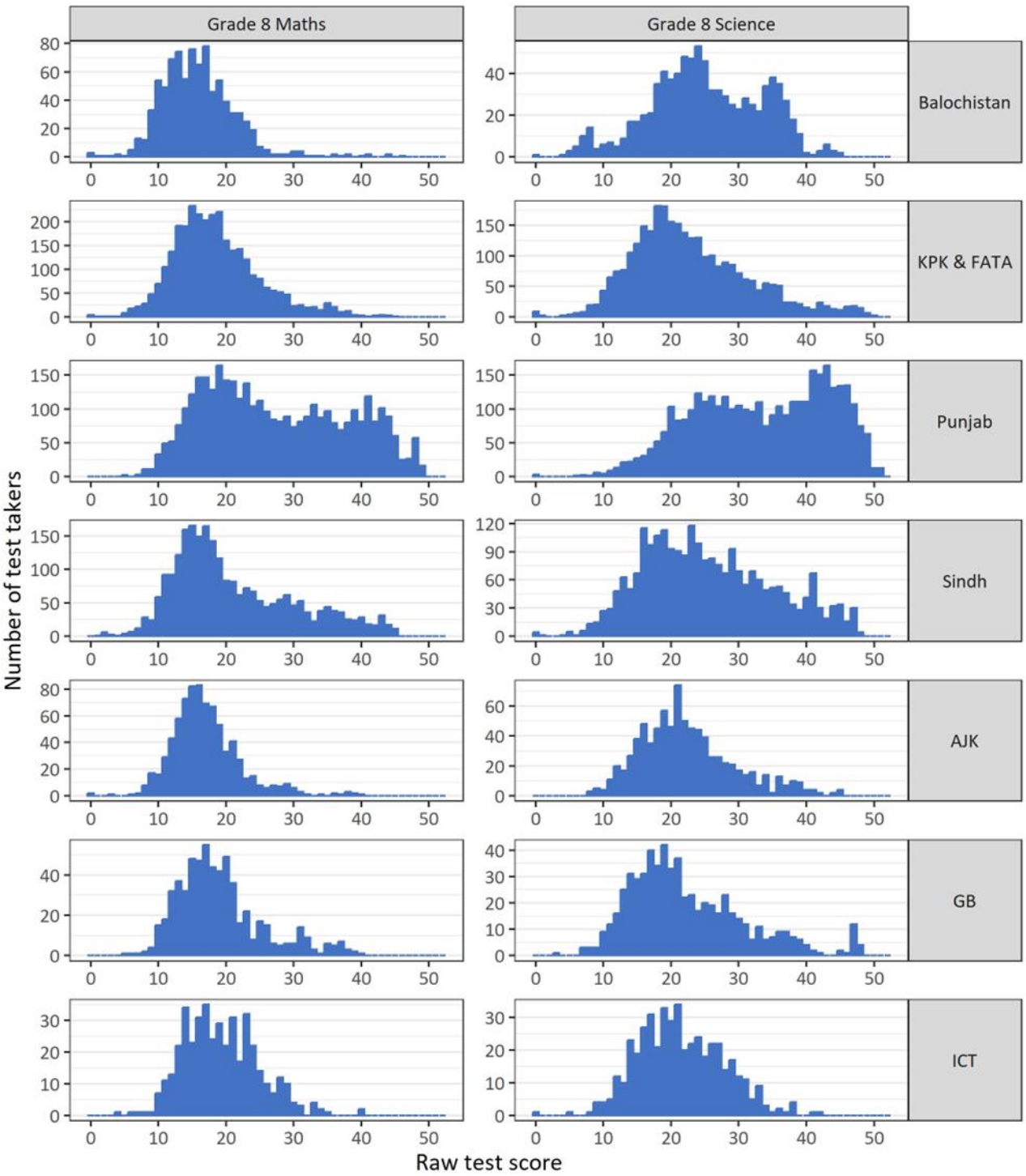


Figure 18: Student score distributions on each Grade 8 test in each province.
The heights of the bars indicate the number of students achieving each available raw score.

As with the national analysis, it is possible to compare performance in Grade 4 English and Grade 8 Maths to scores in NAT 2019 for each province. This is done in the chart in Figure 19[12] and in Table 13. Scores from the different NAT tests are not equated to the same scale but it is still possible to compare performances of different provinces relative to each other between NAT 2019 and NAT 2023. The results show that the high performance of Punjab relative to other provinces in Grade 8 Maths was also present within the NAT 2019 data. Changes in Grade 8 Maths performance since NAT 2019 are fairly small in each province.

In contrast, for Grade 4 English, it is notable that Punjab was not the highest performing province in NAT 2019. However, the percentage of items answered correctly has increased substantially in Punjab between NAT 2019 and NAT 2023, such that Punjab is now the highest performing province. Smaller increases are seen in Balochistan and AJK (although the NAT 2019 samples were fairly small). In the remaining four provinces (KP & NMD, Sindh, Gilgit-Baltistan and ICT), the proportion of items that were correctly answered in Grade 4 English fell.

These results suggest that the possible improvement in Grade 4 English achievement nationally discussed earlier in this report is largely associated with improved performance in Punjab.

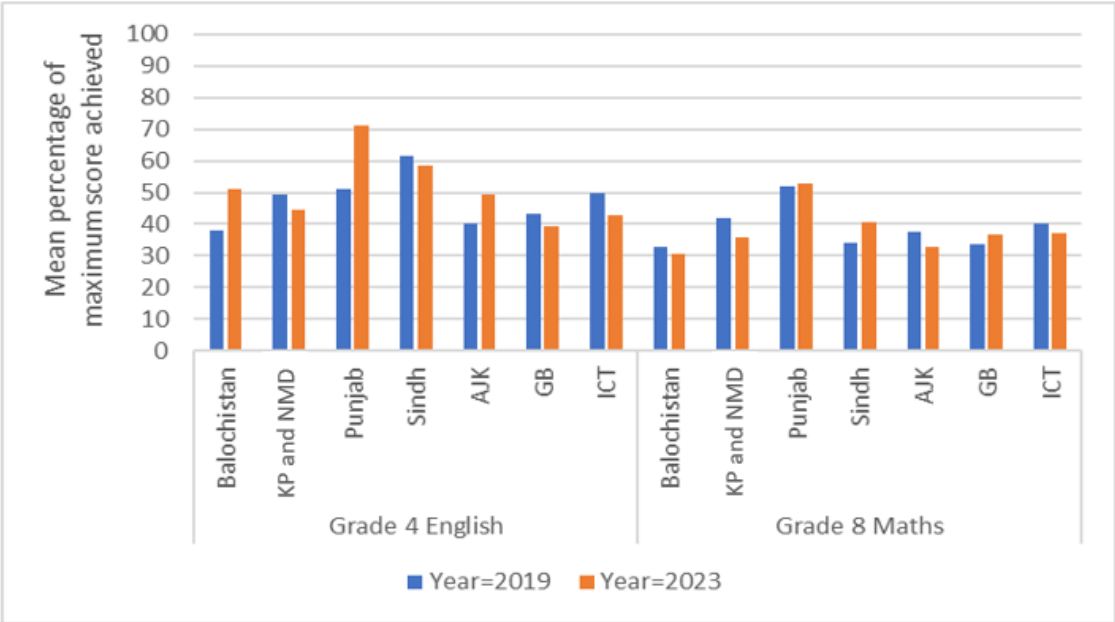


Figure 19: Changes in student performance between 2019 and 2023 by province (source of 2019 data, Tables 6.2 and 6.3 of draft 2019 report)

[12] Bar graphs are used in this instance as the original raw data files to provide complete details of the score distributions would be required. For similar reasons, it is not possible comment on the statistical significance of changes without access to the raw 2019 data. Access to the raw data is needed to assess the extent to which results vary between schools, which is an important part of assessing statistical significance.

Table 13: Changes in student performance between 2019 and 2023 by province (source of 2019 data, Tables 6.2 and 6.3 of draft 2019 report)

Subject	Province	Number of students		Mean score as % of max		SD of scores as % of max	
		2019	2023	2019	2023	2019	2023
Grade 4 English	Balochistan	181	896	38.11	51.22	17.81	22.86
	KP & NMD	1315	2910	49.17	44.45	23.33	21.76
	Punjab	1263	3567	51.11	71.07	24.58	21.19
	Sindh	553	1621	61.58	58.50	31.33	24.57
	AJK	82	594	40.22	49.17	17.33	22.42
	GB	61	514	43.25	39.17	18.92	18.25
	ICT	482	489	49.67	42.65	20.39	20.57
Grade 8 Maths	Balochistan	410	877	32.87	30.55	14.55	10.80
	KP & NMD	1172	3048	42.03	35.75	17.90	12.42
	Punjab	1578	3662	52.13	52.86	19.13	20.07
	Sindh	856	2461	33.90	40.76	15.45	17.36
	AJK	192	791	37.58	32.67	15.81	9.88
	GB	128	616	33.77	36.59	14.06	12.25
	ICT	230	427	40.32	37.13	12.90	10.68

Performance by location

Information on whether schools were urban or rural was not included in the main assessment data files. Instead, the location of schools was drawn from the teacher and headteacher surveys. Teachers and headteachers did not always agree on whether a school was urban or rural. Where available, data from the headteacher was prioritised.

Schools with differing responses regarding the location of the school that could not be resolved by using the seniority of teachers were removed from this section of analysis, as were schools where no headteacher or teacher questionnaire had been completed. For this reason, the number of students in the analysis in this section differs from those shown in earlier tables.

Table 14 compares the mean scores of students in rural and urban schools. A standard error (SE) is included for each mean, as this is helpful in evaluating whether the differences between the two groups of students are statistically significant. These standard errors account for the way the students are clustered within schools.

Using the difference in means and the two standard errors, the statistical significance of the differences can be calculated. This reveals that the only statistically significant difference ($P<0.05$) between rural and urban students is for Grade 8 Maths. In Grade 8 Maths, rural students achieved two more marks on average than those in urban areas. A visual comparison of the average performance of students in urban and rural areas is provided in Figure 20. This chart provides the same information as Table 14 but represents it in terms of the proportion of items answered correctly.

Table 14: Comparisons of the means and standard deviations (SD) of scores of students in rural and urban schools

Subject	Number of students		Mean raw score (SE)		SD		P value (difference)
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	
Grade 4 English	4988	4233	27.3 (0.7)	26.8 (0.7)	11.9	12.2	0.581
Grade 4 FL	4467	3671	13.1 (0.2)	13.5 (0.2)	3.4	3.3	0.136
Grade 4 Maths	4919	4262	24.1 (0.6)	23.6 (0.6)	10.4	10.6	0.570
Grade 4 Urdu and Sindhi	5006	4277	35.8 (0.6)	35.6 (0.6)	11.9	12.3	0.803
Grade 8 Maths	5009	5556	22.6 (0.6)	20.5 (0.6)	9.8	8.5	0.002
Grade 8 Science	4984	5529	27.3 (0.6)	26.1 (0.6)	10.5	9.8	0.112

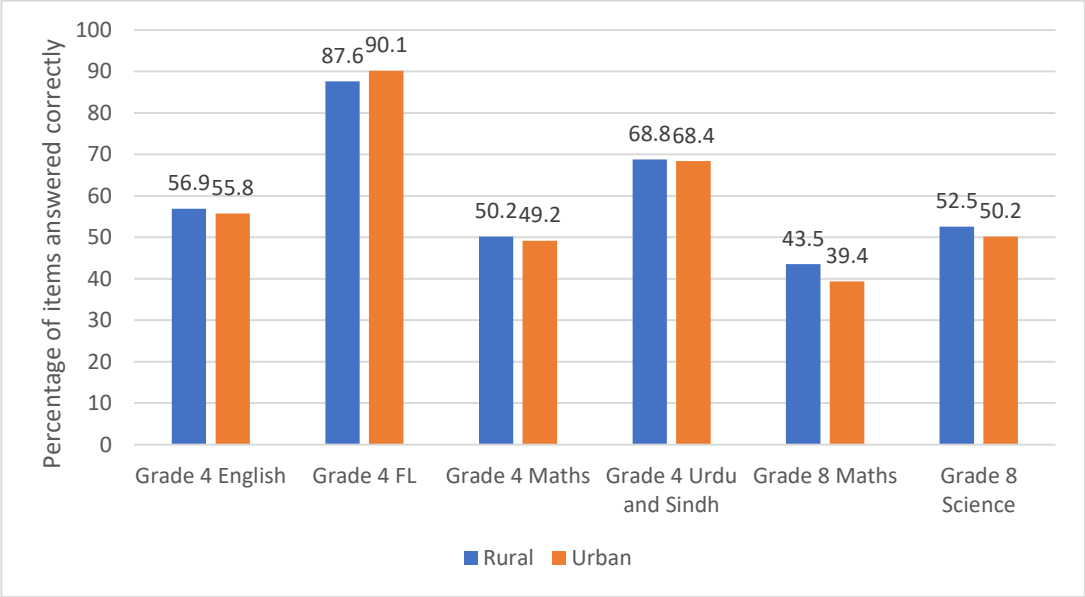


Figure 20: Mean performance of students by whether they attend an urban or a rural school. Results presented in terms of the proportion of items answered correctly.

Table 15 shows the mean performance of students in rural and urban schools in each subject and province, along with the associated standard errors (accounting for the clustering of students within schools). The standard errors are important as they give a sense of how much we might expect the reported means to differ from results if the entire population of students (as opposed to a sample) had participated in assessments. The final column of this table shows the difference in means between urban and rural areas. Further analysis has identified that a handful of the differences are statistically significant ($p<0.05$). Specifically:

- Urban schools significantly outperformed rural schools in Grade 4 Maths in Balochistan and in Grade 8 Science in ICT.
- Rural schools significantly outperformed urban schools in Grade 4 English in KP & NMD, Grade 4 Urdu and Sindhi in KP & NMD, Grade 8 Maths in Sindh (although, as described earlier, this is typical nationally), and Grade 8 Science in Sindh.

Another way to look at differences in performance between rural and urban schools in each province is to calculate the percentage of students who achieve no better than the average score that would be achieved by random guessing – that is, the percentage of students achieving no more than a quarter of marks on the test. This information is shown in

Table 16. Colour highlighting is used so that the highest values (indicating many students performing no better than guessing) are in red and the lowest values are in green. In the worst case, these values can be as high as 40%. In particular, this occurs in rural areas in Balochistan in Grade 4 Maths while only 15% of students perform at this low level within urban areas. This reflects the significant differences between rural and urban areas in Balochistan noted above. In contrast, also in Grade 4 Maths, in Gilgit-Baltistan 37% of students in urban areas perform at this low level compared to only 17% in rural areas. In other provinces, differences between rural and urban areas in Table 16 are much smaller.

Table 15: Comparisons of means and standard deviations of scores of urban and rural students in each province

Subject	Province	Number of students		Mean raw score (SE)		Standard deviation		Difference
		Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	
Grade 4 English	Balochistan	311	409	22.5 (2.3)	25.5 (2.3)	10.7	11.6	3.04
	KP & NMD	1263	1153	23.3 (1.3)	19.3 (1.0)	11.3	9.1	-4.10
	Punjab	1818	1387	33.5 (1.0)	34.7 (1.1)	10.2	10.1	1.15
	Sindh	751	658	26.6 (1.5)	30.2 (1.7)	11.9	11.6	3.64
	AJK	366	228	23.8 (1.8)	23.3 (2.3)	10.4	11.3	-0.46
	GB	218	170	21.1 (2.5)	16.9 (1.3)	10.1	7.1	-4.15
	ICT	261	228	21.1 (2.2)	19.7 (2.2)	10.5	9.1	-1.46
Grade 4 Maths	Balochistan	277	425	15.3 (1.4)	21.2 (1.6)	7.2	8.5	5.90
	KP & NMD	1268	1170	20.0 (1.0)	18.1 (1.0)	8.8	8.9	-1.93
	Punjab	1801	1370	30.5 (0.9)	30.6 (1.0)	9.6	9.3	0.02
	Sindh	756	672	23.3 (1.0)	24.4 (1.4)	8.8	9.6	1.08
	AJK	363	229	20.6 (1.7)	21.3 (2.4)	8.9	11.0	0.64
	GB	193	170	20.7 (2.2)	18.1 (2.9)	8.9	10.1	-2.63
	ICT	261	226	18.3 (2.0)	18.7 (1.9)	8.6	7.8	0.35
	Balochistan	315	419	30.4 (1.9)	35.8 (2.2)	10.5	12.3	5.41
	KP & NMD	1276	1161	31.4 (1.3)	27.7 (1.2)	12.7	11.6	-3.68

Subject	Province	Number of students		Mean raw score (SE)		Standard deviation		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 Urdu and Sindhi	Punjab	1816	1380	41.5 (0.8)	43.0 (0.8)	9.7	8.5	1.43
	Sindh	759	699	34.9 (1.3)	38.3 (1.3)	12.0	10.8	3.34
	AJK	365	217	34.9 (1.3)	31.6 (2.5)	9.3	12.7	-3.26
	GB	210	170	32.4 (2.1)	28.3 (1.3)	9.5	8.6	-4.03
	ICT	265	231	30.2 (2.3)	31.0 (2.3)	11.0	10.7	0.79
Grade 8 Maths	Balochistan	283	521	15.8 (1.2)	15.8 (0.7)	6.5	5.3	-0.08
	KP & NMD	1456	1394	18.3 (0.5)	18.4 (0.6)	6.0	6.5	0.13
	Punjab	1659	1663	28.8 (1.1)	26.0 (1.0)	10.8	10.0	-2.77
	Sindh	907	1205	23.1 (1.2)	18.9 (0.8)	9.4	7.6	-4.27
	AJK	204	364	18.6 (1.4)	16.6 (0.5)	5.9	4.3	-2.01
	GB	322	180	20.0 (1.5)	18.0 (1.0)	7.2	4.7	-2.07
	ICT	178	229	18.0 (0.5)	20.0 (1.0)	5.0	5.8	1.98
Grade 8 Science	Balochistan	285	526	26.0 (1.9)	24.5 (1.3)	8.6	8.1	-1.51
	KP & NMD	1431	1362	22.2 (0.8)	22.7 (0.8)	8.5	8.4	0.48
	Punjab	1627	1615	34.3 (0.9)	32.9 (0.9)	9.6	9.9	-1.45
	Sindh	906	1206	27.1 (1.3)	24.0 (0.9)	10.1	8.9	-3.09
	AJK	206	398	24.2 (2.1)	22.8 (1.0)	8.4	6.8	-1.41
	GB	334	173	23.1 (2.2)	21.6 (1.5)	9.8	6.8	-1.49
	ICT	195	249	19.7 (0.9)	22.6 (1.0)	5.5	6.7	2.87

Table 16: Comparing the percentages of students performing no better than guessing in rural and urban areas in each test

Subject	Province	Number of students		% students achieving 25% of marks or less	
		Rural	Urban	Rural	Urban
Grade 4 English	Balochistan	311	409	20.9%	15.9%
	KP & NMD	1263	1153	18.1%	23.9%
	Punjab	1818	1387	5.0%	5.1%
	Sindh	751	658	15.6%	9.6%
	AJK	366	228	14.2%	15.8%
	GB	218	170	26.1%	19.4%
	ICT	261	228	21.5%	21.9%
Grade 4 Maths	Balochistan	277	425	40.1%	14.8%
	KP & NMD	1268	1170	21.0%	29.5%
	Punjab	1801	1370	4.3%	4.5%
	Sindh	756	672	13.1%	8.2%
	AJK	363	229	20.4%	23.1%
	GB	193	170	17.1%	37.1%
	ICT	261	226	26.1%	20.8%
Grade 4 Urdu and Sindhi	Balochistan	315	419	3.2%	5.0%
	KP & NMD	1276	1161	7.9%	10.6%
	Punjab	1816	1380	1.0%	0.7%
	Sindh	759	699	5.1%	2.7%

	AJK	365	217	1.6%	9.7%
	GB	210	170	1.4%	2.9%
	ICT	265	231	8.3%	5.2%
Grade 8 Maths	Balochistan	283	521	38.9%	36.9%
	KP & NMD	1456	1394	21.2%	21.4%
	Punjab	1659	1663	5.9%	6.9%
	Sindh	907	1205	13.2%	23.6%
	AJK	204	364	16.7%	22.0%
	GB	322	180	16.8%	17.2%
	ICT	178	229	15.2%	13.5%
Grade 8 Science	Balochistan	285	526	8.8%	7.2%
	KP & NMD	1431	1362	12.1%	10.6%
	Punjab	1627	1615	1.6%	1.4%
	Sindh	906	1206	8.4%	10.2%
	AJK	206	398	6.3%	6.5%
	GB	334	173	12.9%	12.7%
	ICT	195	249	11.8%	6.0%

Performance by gender

The gender of each student was included in the data sets for each assessment. On rare occasions, the gender data were inconsistent for the same student across different data files. Where this occurred, the gender recorded against the specific assessment of interest was used in analysis. For example, the results for Grade 4 English were based on gender as recorded in the Grade 4 English data file.

Table 17 shows the performance on each test split by gender. A standard error (SE) is included for each mean. This indicates the extent to which the results would be expected to differ from their true population values, given the number of students and schools included in the sample (i.e., not all students took these tests). The SE is helpful in evaluating whether differences between males and females are statistically significant – that is, whether the size of difference seen would be likely to occur even if, at population level, males and females performed identically. Note that these standard errors account for the way in which students are clustered within schools.

The mean scores of female students were higher than those for male students in every subject. Statistically significant results can be identified by any rows where the P value is less than 0.05. Statistically significant differences were clearly evident in Grade 4 English, Grade 4 Urdu and Sindhi, and Grade 8 Science. A smaller, but still statistically significant difference, was seen for Grade 4 FL. The difference in performance between boys and girls was not significant for either of the Maths assessments. A visual comparison of the average performance of boys and girls is provided in Figure 21. This chart provides the same information as Table 17 but represents it in terms of the proportion of items answered correctly.

Table 17: Comparisons of the means and standard deviations (SDs) of scores of male and female students

Subject	Max	Number of students		Mean raw score (SE)		Standard deviation		P value (difference)
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	5386	5205	25.46 (0.63)	28.41 (0.65)	11.80	11.97	0.001
Grade 4 FL	15	4713	4641	12.97 (0.19)	13.52 (0.16)	3.69	3.09	0.023
Grade 4 Maths	48	5435	5143	23.33 (0.55)	24.10 (0.58)	10.47	10.60	0.338
Grade 4 Urdu and Sindhi	52	5425	5237	33.88 (0.62)	37.05 (0.56)	12.60	11.37	0.000
Grade 8 Maths	52	5900	5982	21.10 (0.44)	22.12 (0.50)	8.91	9.58	0.124
Grade 8 Science	52	5789	6010	25.17 (0.48)	28.14 (0.52)	9.77	10.24	0.000

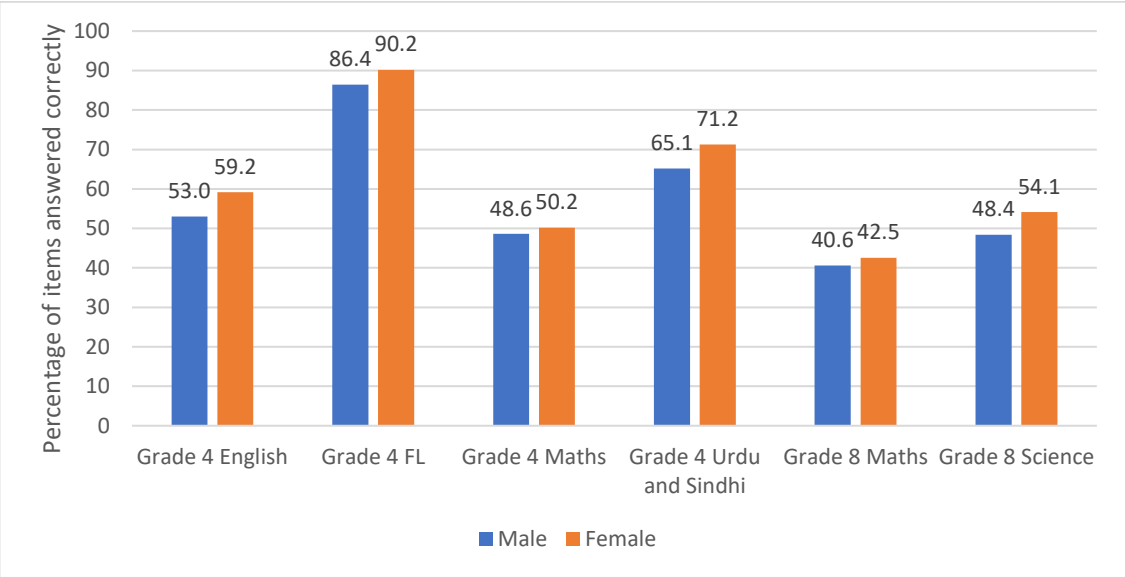


Figure 21: Mean performance of boys and girls in each test. Results presented in terms of the proportion of items answered correctly.

Table 18 shows the means and standard deviations of raw scores by gender within each province. However, it should be noted that a review of this data using multilevel modelling analysis found that, after accounting for the way students are clustered within schools (that is, accounting for the fact that students in the same school often display similar performance), there was no significant variation in the size of the gender gap across different provinces. As such, the values in Table 18 should not be taken to imply that any province has a particularly positive or negative impact on female students compared to male students.

Table 18: Comparisons of means and standard deviations of scores of male and female students in each province

Subject	Max	Province	Number of students		Mean raw score (SE)		SD		Difference in means
			Male	Female	Male	Female	Male	Female	
Grade 4 English	48	Balochistan	436	460	23.52 (1.67)	25.60 (2.15)	9.99	11.75	2.07
		KP & NMD	1477	1433	19.36 (0.93)	23.37 (1.18)	9.27	11.18	4.02
		Punjab	1762	1805	32.37 (1.02)	35.81 (0.87)	10.63	9.40	3.44
		Sindh	875	746	26.51 (1.48)	29.91 (1.35)	12.30	10.90	3.40
		AJK	318	276	24.25 (2.13)	22.85 (1.88)	10.98	10.47	-1.41
		GB	258	256	17.81 (1.89)	19.80 (1.97)	8.68	8.74	1.99
		ICT	260	229	22.03 (2.39)	18.69 (1.79)	10.83	8.34	-3.34
Grade 4 Maths	48	Balochistan	448	428	19.61 (1.34)	19.13 (1.76)	8.52	9.15	-0.48
		KP & NMD	1510	1424	18.62 (0.93)	19.35 (0.87)	9.35	8.48	0.74
		Punjab	1776	1772	29.52 (0.94)	31.64 (0.82)	9.98	8.80	2.12
		Sindh	879	773	23.28 (1.06)	24.24 (1.21)	8.89	9.93	0.96
		AJK	320	272	22.84 (1.90)	18.54 (1.79)	9.53	9.55	-4.30
		GB	243	246	18.88 (1.71)	18.18 (2.27)	8.26	9.32	-0.70
		ICT	259	228	19.75 (2.30)	17.04 (1.23)	9.54	6.26	-2.72
Grade 4 Urdu and Sindhi	52	Balochistan	444	469	32.01 (1.82)	35.27 (2.07)	11.63	12.06	3.26
		KP & NMD	1499	1427	26.54 (1.16)	32.78 (1.03)	12.29	11.36	6.23
		Punjab	1754	1813	41.00 (0.82)	43.32 (0.71)	9.74	8.48	2.32
		Sindh	896	777	35.54 (1.27)	36.75 (1.27)	11.70	11.46	1.21
		AJK	317	265	34.32 (1.95)	32.92 (1.43)	11.51	9.92	-1.39
		GB	250	255	27.49 (1.66)	31.40 (1.70)	9.72	8.69	3.91

Subject	Max	Province	Number of students		Mean raw score (SE)		SD		Difference in means
			Male	Female	Male	Female	Male	Female	
Grade 8 Maths	52	Balochistan	379	498	16.11 (0.95)	15.72 (0.76)	6.15	5.18	-0.38
		KP & NMD	1527	1521	18.15 (0.52)	19.04 (0.64)	6.08	6.79	0.89
		Punjab	1818	1844	26.16 (0.94)	28.80 (0.99)	10.28	10.43	2.64
		Sindh	1225	1236	21.26 (0.99)	21.12 (1.03)	8.95	9.11	-0.14
		AJK	398	393	17.53 (0.82)	16.44 (0.41)	5.38	4.83	-1.09
		GB	330	286	18.40 (1.18)	19.75 (1.46)	5.98	6.73	1.35
		ICT	223	204	18.02 (0.48)	20.71 (1.07)	4.83	5.96	2.69
Grade 8 Science	52	Balochistan	352	531	24.78 (1.15)	25.38 (1.48)	7.30	8.78	0.60
		KP & NMD	1479	1503	21.27 (0.75)	24.38 (0.85)	8.18	8.80	3.11
		Punjab	1761	1826	31.21 (0.83)	35.62 (0.87)	9.58	9.54	4.41
		Sindh	1226	1211	24.25 (1.06)	26.91 (0.98)	9.89	9.05	2.66
		AJK	425	403	22.65 (1.27)	22.03 (0.88)	7.58	6.36	-0.61
		GB	329	289	20.88 (1.67)	23.95 (1.99)	7.99	8.87	3.08
		ICT	217	247	20.14 (0.62)	22.57 (1.16)	5.52	6.72	2.42

Comparison of results between students and teachers

Teachers were identified within our data as those with roll numbers equal to 21 or 22. A comparison of the mean raw scores achieved by students and teachers is shown in Table 19. Figure 22 compares the score distributions. Specifically, the figure contains two panels for each assessment – one for students and one for teachers. Within each panel, the bars indicate the number of people achieving each available score on each test. The panels relating to students actually convey the same information as shown earlier in Figure 16. Only the comparison with teachers is new. Grade 4 FL is not included in this analysis as it was not usually taken by teachers.

Teachers' scores were well above students' scores on average. The lowest average scores for teachers occurred in the Maths tests, where the mean scores of 35.5 for Grade 4 and 36.9 for Grade 8 represent achieving 74% and 71% of the available maximum. In percentage terms, teachers' highest average scores occurred in Grade 4 English (39.8 or 83% of the maximum) and Grade 4 Urdu and Sindhi (46.9 or 90% of the maximum). In every subject, about 90% of teachers' scores were greater than or equal to the average student score in their school[13].

[13] Based on schools with test scores from at least five students.

More detailed analysis of the relationship between teacher performance and the average scores of students in the same school will be provided later in this report (see Section 6.2). The score distributions show that some test takers identified as teachers by their roll number had extremely low-test scores. The reasons for this are not known. However, this issue should not detract from the overall pattern, which indicates generally high teacher performance.

Table 19: Means and standard deviations of scores in each subject for students and teachers

Subject	Max	Number of students		Mean raw score (SE)		Standard deviation		Difference
		Student	Teacher	Student	Teacher	Student	Teacher	
Grade 4 English	48	10591	229	26.91 (0.46)	39.79 (0.58)	11.97	8.83	12.9
Grade 4 Maths	48	10578	236	23.70 (0.40)	35.47 (0.62)	10.54	9.48	11.8
Grade 4 Urdu and Sindhi	52	10662	228	35.43 (0.42)	46.92 (0.57)	12.11	8.40	11.5
Grade 8 Maths	52	11882	310	21.61 (0.33)	36.95 (0.53)	9.26	9.49	15.3
Grade 8 Science	52	11799	302	26.68 (0.36)	40.20 (0.53)	10.12	9.17	13.5

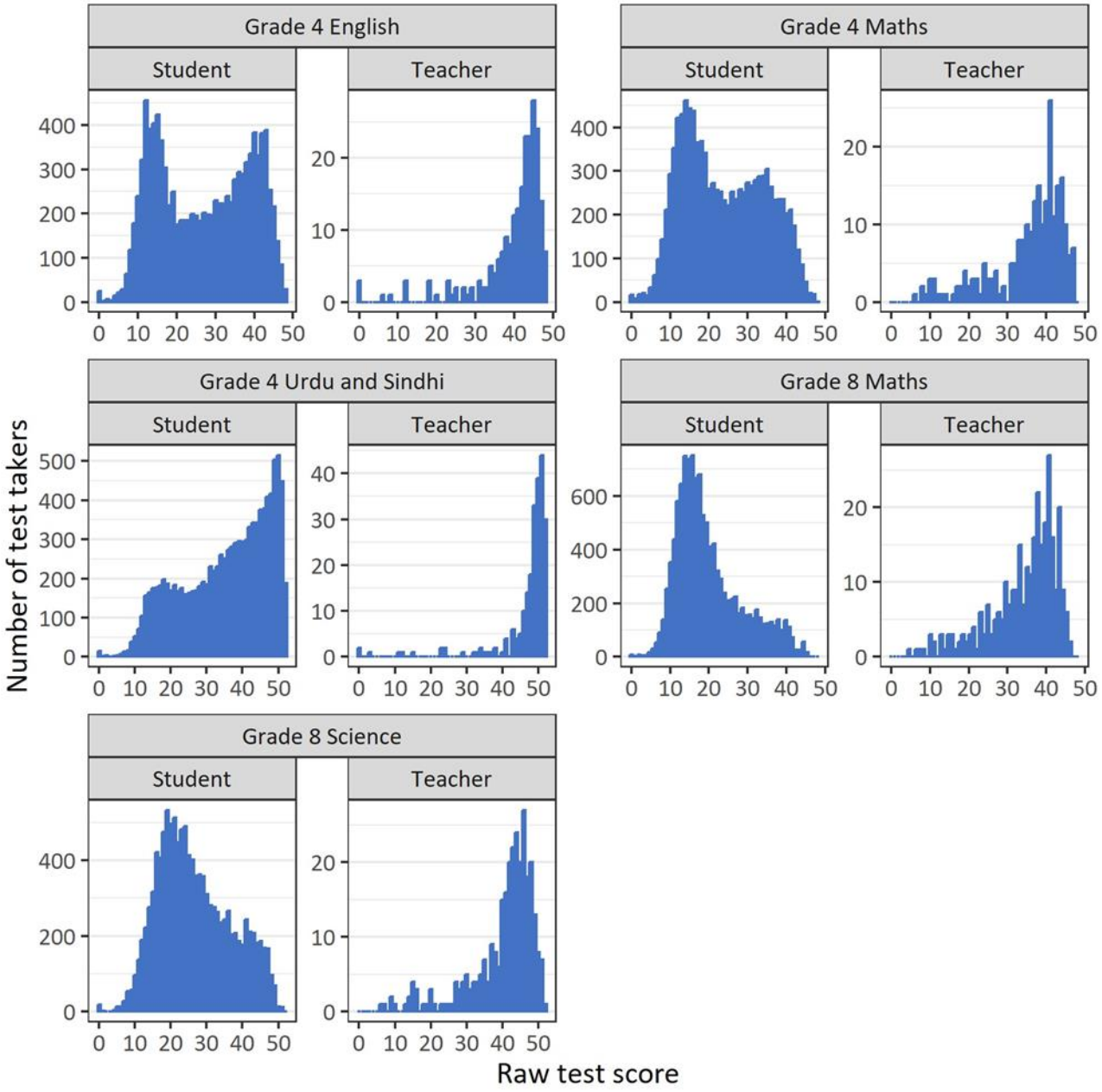


Figure 22: A comparison of raw score distributions for students and teachers

Note: No analysis of the performance of teachers across different provinces is included as the numbers of teachers in certain provinces was too low for analysis to be meaningful.

5. Student achievement in different content and cognitive domains

5.1. Main findings in this section

The results in this section show:

1. For both students and teachers, performance in each assessment was fairly consistent across the different content and cognitive domains.
2. However, it was clear that students and teachers both answered a smaller proportion of items correctly in the Grade 8 Maths content domain of statistics and probability. This may relate to this content domain being introduced to the curriculum since the 2019 NAT and may suggest that teachers require additional support in understanding and teaching this area.
3. Performance for students and teachers in Grade 8 tended to be higher in Life Sciences than in either Earth and Space Sciences or Physical Sciences.

5.2. Detailed method and results

This section compares assessment performance across the different content and cognitive domains covered by each test. Full details of what these domains mean and how they are measured is provided in Appendix 1. Please note that this analysis is not relevant for Grade 4 Foundational Literacy and so this subject is not included in the results.

The results for each content and cognitive domain are presented in terms of the proportion of items that are answered correctly on average. At the moment, it is not clear whether differences in performance represent differences in the difficulty of items (e.g., relative to relevant learning frameworks) or genuine strengths and weaknesses of students and teachers in different areas. For example, a low proportion of students answering an item incorrectly might indicate that the items within a particular domain were qualitatively more difficult or it might indicate a gap in student knowledge. As such, some caution needs to be exercised in the interpretation of these results.

In order to provide some indication of statistical significance, the extent to which results vary across different individual items within each cognitive and content domain was analysed and used to create indicative standard errors and confidence intervals[14]. Indicative standard errors are included to give a sense of how much average performance in each domain might change if it was measured with a larger number of items.

[14] Specifically, as well as calculating the average proportion of items that are answered correctly, the standard deviation of the proportion of correct responses across each item was calculated. Indicative standard errors were then estimated by dividing this value by the square root of the number of items in the domain. Indicative confidence intervals were derived from these standard errors. These confidence intervals were only indicative as they do not account for sampling error across students and schools. However, given the fairly small number of items in each domain, variation across items is likely to be the major source of sampling error. The confidence intervals provide a range of values where we are 95% certain the percentage of items answered correctly in each domain would lie if this was measured with a larger sample of items.

Figure 23 shows the results by cognitive domain for students and teachers. The exact same information is provided in tabular form in Table 20. For both students and teachers, the figure shows that the indicative confidence intervals for performance overlap across the different domains. This implies there is no significant difference in the level of performance across different cognitive domains.

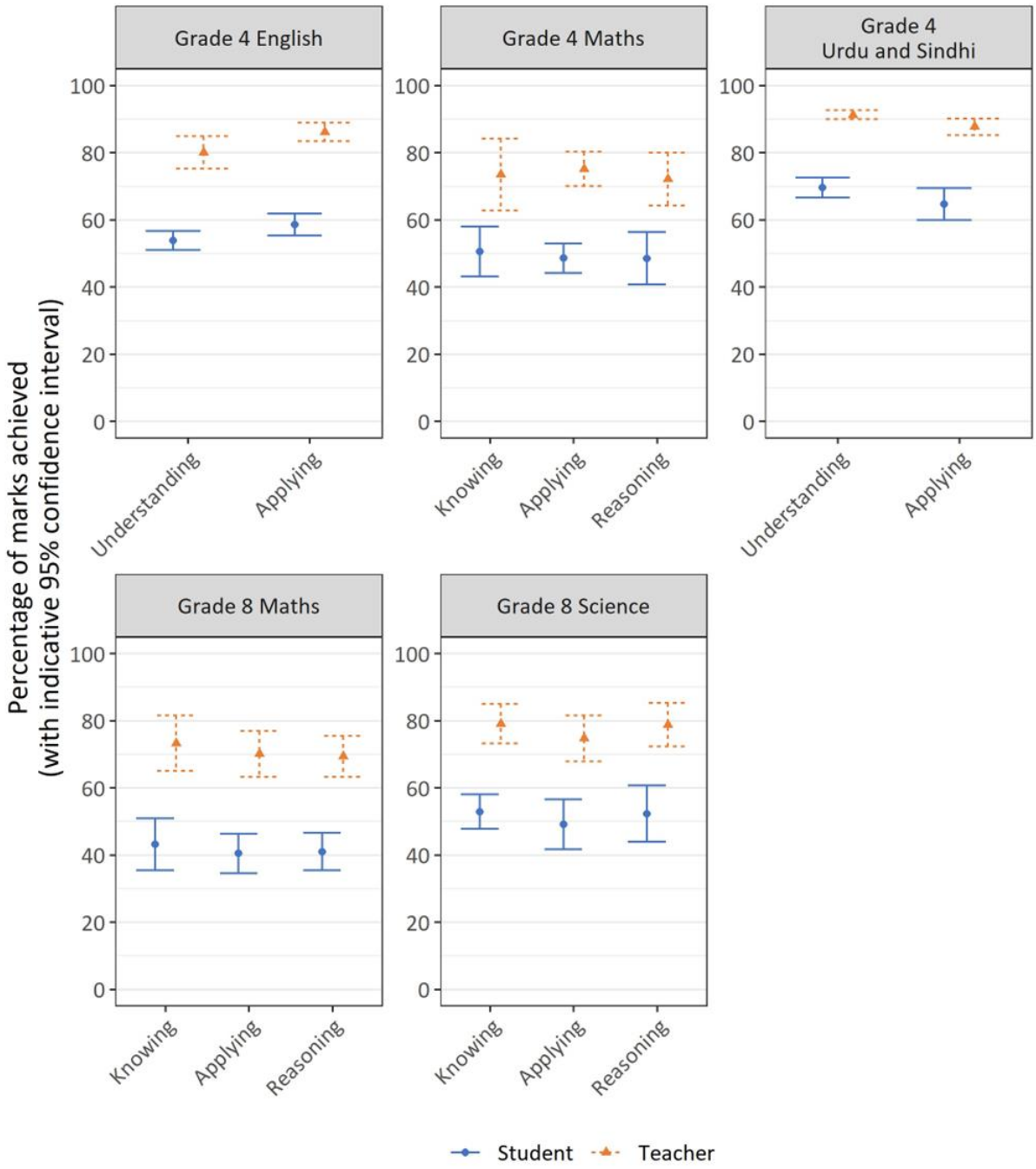
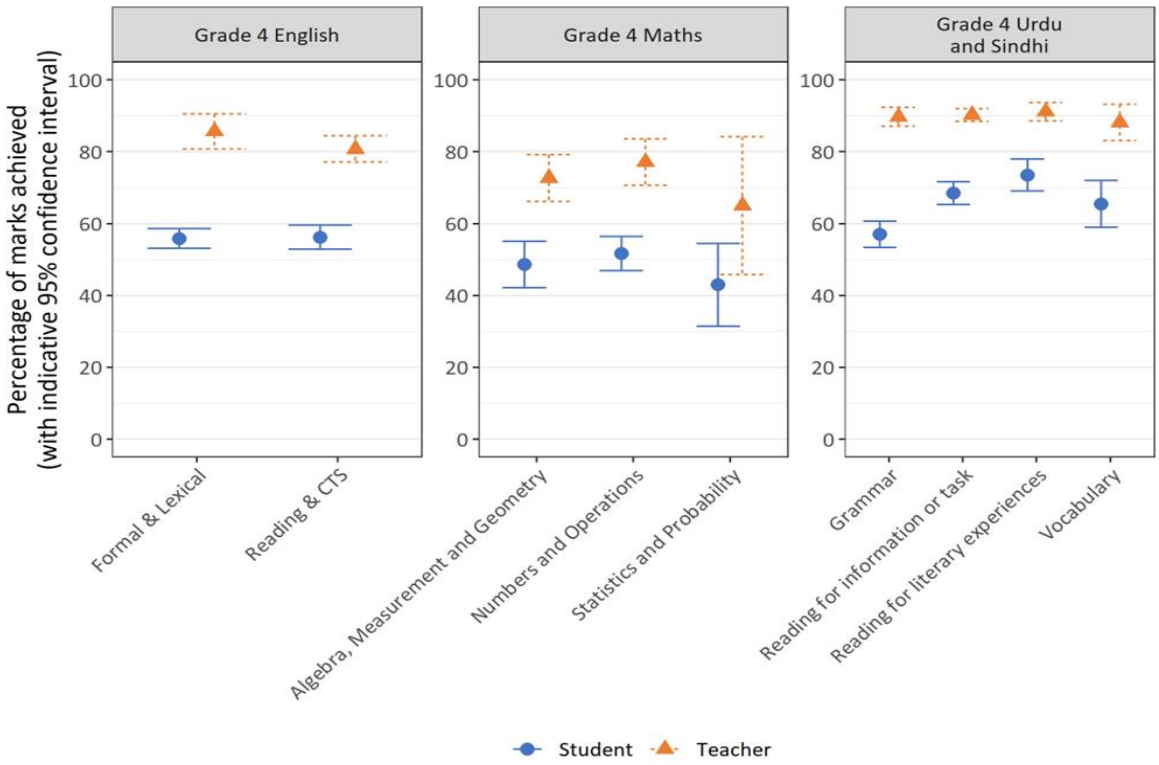


Figure 23: Average percentage of marks achieved by students and teachers within each cognitive domain in each subject

Table 20: Comparisons of the average percentage of marks achieved by students and teachers within each cognitive domain in each subject

Subject	Cognitive domain	Percentage of possible marks achieved (indicative standard error)	
		Student	Teacher
Grade 4 English	Understanding	53.9 (1.4)	80.1 (2.5)
	Applying	58.6 (1.7)	86.2 (1.4)
Grade 4 Maths	Knowing	50.6 (3.8)	73.5 (5.4)
	Applying	48.6 (2.2)	75.2 (2.6)
	Reasoning	48.6 (4.0)	72.2 (4.0)
Grade 4 Urdu and Sindhi	Understanding	69.6 (1.5)	91.3 (0.7)
	Applying	64.8 (2.4)	87.8 (1.3)
Grade 8 Maths	Knowing	43.2 (3.9)	73.3 (4.2)
	Applying	40.5 (3.0)	70.1 (3.5)
	Reasoning	41.0 (2.9)	69.4 (3.1)
Grade 8 Science	Knowing	52.9 (2.6)	79.2 (3.0)
	Applying	49.2 (3.8)	74.7 (3.5)
	Reasoning	52.4 (4.3)	78.8 (3.3)



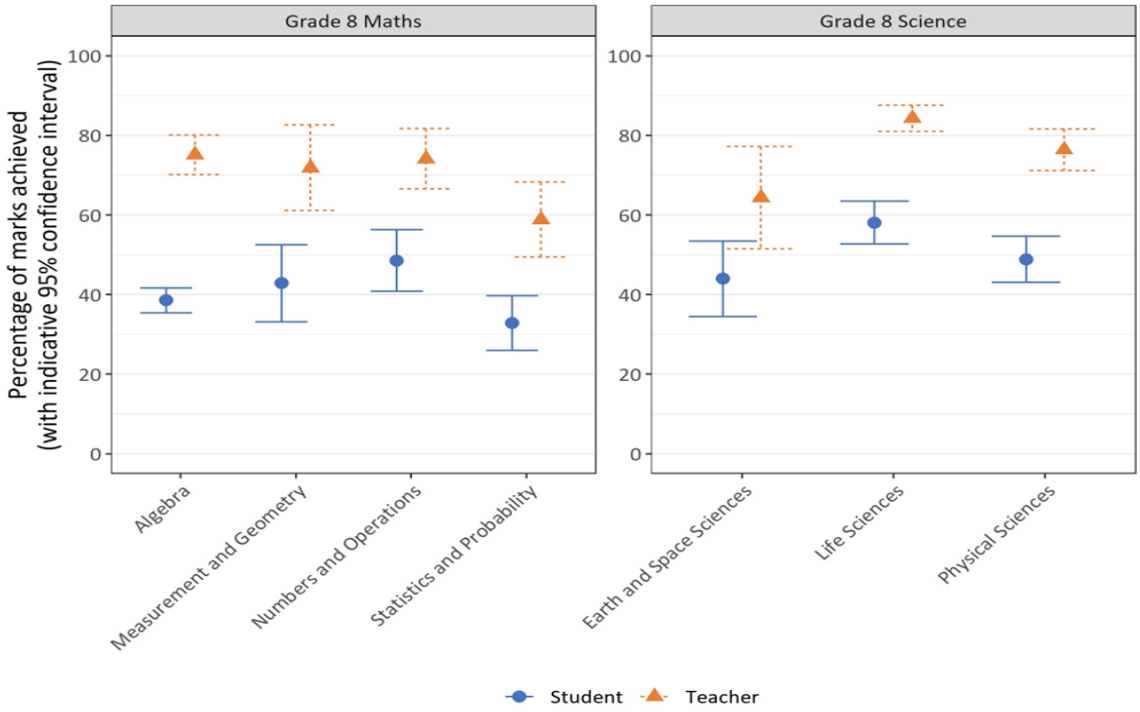


Figure 24,

Figure 25 and Table 21 repeat this same analysis for content domains. The figures are split into results for Grade 4 and Grade 8 subjects respectively. In Grade 4, there is no evidence of any significant difference in teacher performance across the different content domains. However, students achieved a significantly smaller proportion of available marks on grammar questions within Grade 4 Urdu and Sindhi.

For Grade 8 Science, both students and teachers achieved a significantly higher proportion of marks in Life Sciences items than in either of the other content domains. In Grade 8 Maths, both students and teachers achieved fewer marks in Statistics and Probability items than in other content domains.

Grade 8 Statistics and Probability displayed the lowest level of performance across all subjects and content domains, both for students (only 33% of items answered correctly on average) and for teachers (only 59% of items answered correctly on average). This fact may be related to this being a new area of content that has been introduced in the 2022 National Curriculum. Student performance in Algebra was also a little below performance in other Grade 8 Maths content domains other than Statistics and Probability.

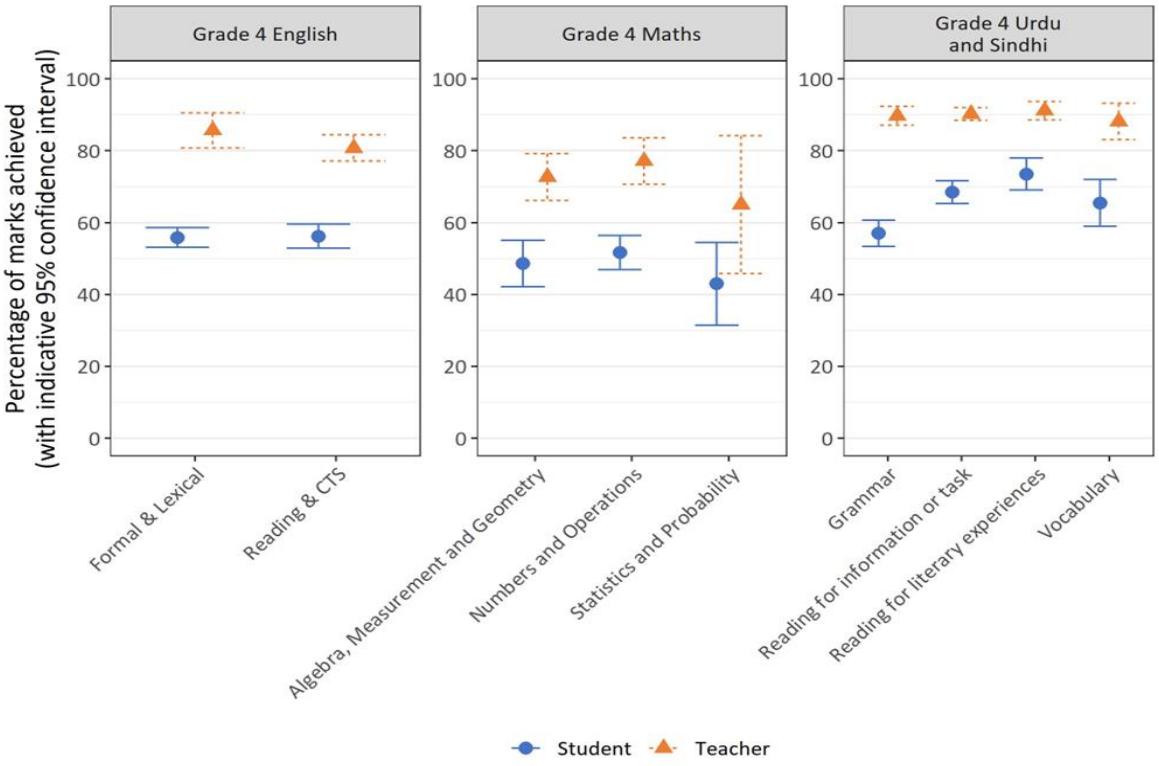


Figure 24: Average percentage of marks achieved by students and teachers within each content domain in each Grade 4 subject

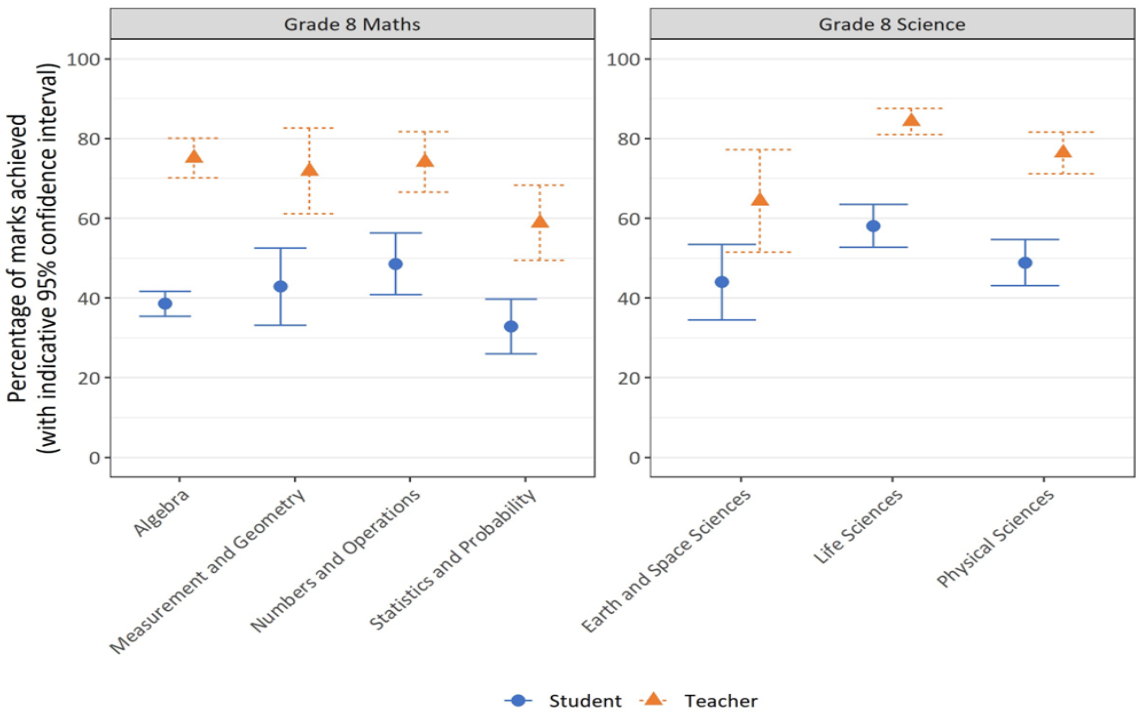


Figure 25: Average percentage of marks achieved by students and teachers within each content domain in each Grade 8 subject

Table 21: Percentage of items answered correctly by students and teachers within each content domain in each subject

Subject	Content domain	Percentage of possible marks achieved (indicative standard error)	
		Student	Teacher
Grade 4 English	Formal & Lexical	55.9 (1.4)	85.7 (2.5)
	Reading & CTS	56.2 (1.7)	80.8 (1.9)
Grade 4 Maths	Algebra, Measurement and Geometry	48.6 (3.3)	72.7 (3.3)
	Numbers and Operations	51.7 (2.4)	77.2 (3.3)
	Statistics and Probability	43.0 (5.9)	65.0 (9.8)
Grade 4 Urdu and Sindhi	Grammar	57.0 (1.9)	89.7 (1.3)
	Reading for information or task	68.5 (1.6)	90.3 (0.9)
	Reading for literary experiences	73.5 (2.3)	91.2 (1.3)
	Vocabulary	65.5 (3.3)	88.2 (2.5)
Grade 8 Maths	Algebra	38.6 (1.6)	75.1 (2.5)
	Measurement and Geometry	42.9 (4.9)	71.9 (5.5)
	Numbers and Operations	48.6 (4.0)	74.1 (3.9)
	Statistics and Probability	32.9 (3.5)	58.9 (4.8)
Grade 8 Science	Earth and Space Sciences	44.0 (4.8)	64.4 (6.6)
	Life Sciences	58.1 (2.7)	84.3 (1.7)
	Physical Sciences	48.9 (3.0)	76.4 (2.7)

Table 22 and Table 23 show the performance of students on each cognitive and content domain across different provinces. Indicative standard errors are included to give a sense of how much average performance in each domain might change in each province if it was measured with a larger number of items. For every subject, for every content and cognitive domain, the highest proportion of correctly answered items was seen in Punjab. The only other major difference that could be seen was for Grammar in Grade 4 Urdu and Sindhi, where Gilgit-Baltistan and ICT displayed substantially lower performance in this content domain compared to the others.

Table 22: Mean and indicative standard error for the percentage of possible marks achieved by students within each cognitive domain in each province

Subject	Cognitive domain	Percentage of possible marks achieved (indicative standard error)						
		Balochistan	KP & NMD	Punjab	Sindh	AJK	GB	ICT
Grade 4 English	Applying	53.0 (2.0)	46.6 (1.9)	74.5 (1.4)	59.6 (1.7)	51.5 (1.9)	42.7 (2.4)	46.5 (2.6)
	Understanding	49.7 (1.5)	42.6 (1.3)	68.2 (2.0)	57.5 (1.5)	47.2 (1.7)	36.2 (1.8)	39.4 (1.6)
Grade 4 Maths	Knowing	43.3 (3.6)	42.0 (3.1)	63.2 (4.9)	50.5 (4.2)	44.1 (3.1)	40.1 (3.5)	42.1 (3.8)
	Applying	38.8 (2.0)	37.3 (2.0)	64.6 (2.8)	49.0 (2.8)	43.2 (2.0)	35.9 (2.1)	36.0 (2.9)

Subject	Cognitive domain	Percentage of possible marks achieved (indicative standard error)						
		Balochistan	KP & NMD	Punjab	Sindh	AJK	GB	ICT
Grade 4 Urdu and Sindh	Reasoning	37.9 (4.3)	39.2 (3.5)	62.9 (4.6)	48.2 (4.9)	42.7 (4.5)	40.8 (3.7)	36.6 (4.6)
	Understanding	66.4 (1.6)	58.0 (1.6)	82.5 (1.3)	71.5 (1.3)	66.1 (2.0)	59.1 (2.7)	60.7 (2.5)
	Applying	61.2 (3.2)	54.4 (2.3)	78.0 (2.1)	64.7 (3.1)	61.9 (2.9)	51.1 (3.9)	54.6 (3.7)
Grade 8 Maths	Knowing	32.2 (3.8)	37.3 (4.5)	55.3 (4.3)	41.4 (3.6)	33.7 (4.7)	38.7 (4.7)	39.4 (5.1)
	Applying	29.0 (2.6)	35.0 (2.9)	51.1 (3.5)	40.1 (3.4)	33.2 (3.3)	35.1 (3.4)	35.7 (3.5)
	Reasoning	30.8 (3.0)	34.9 (2.7)	52.4 (3.5)	40.9 (2.6)	30.3 (2.8)	36.0 (3.5)	36.3 (4.4)
Grade 8 Science	Knowing	50.7 (3.1)	44.7 (3.0)	66.1 (2.9)	51.0 (2.4)	46.6 (3.6)	44.6 (3.2)	41.2 (3.3)
	Applying	45.6 (5.2)	41.9 (3.3)	62.2 (4.2)	47.7 (3.7)	38.7 (4.2)	40.9 (3.6)	40.2 (4.3)
	Reasoning	49.2 (4.9)	46.5 (4.1)	65.0 (4.5)	48.6 (5.0)	44.4 (4.7)	43.5 (3.8)	43.5 (4.8)

Table 23: Mean and indicative standard error for the percentage of possible marks achieved by students within each content domain in each province

Subject	Content domain	Percentage of possible marks achieved (indicative standard error)						
		Balochistan	KP & NMD	Punjab	Sindh	AJK	GB	ICT
Grade 4 English	Formal & Lexical	49.9 (1.6)	43.2 (1.3)	72.1 (2.1)	59.1 (1.5)	47.5 (1.7)	37.6 (1.5)	42.9 (1.8)
	Reading & Critical Thinking Skills	52.2 (1.8)	45.4 (1.8)	70.3 (1.8)	58.0 (1.7)	50.5 (1.9)	40.4 (2.4)	42.4 (2.4)
Grade 4 Maths	Algebra, Measurement and Geometry	39.7 (3.3)	40.0 (3.2)	61.5 (3.7)	48.9 (3.9)	43.3 (2.9)	38.4 (3.6)	38.2 (4.3)
	Numbers and Operations	42.3 (2.4)	41.0 (2.0)	67.0 (3.2)	51.5 (2.7)	45.3 (2.2)	40.2 (2.3)	40.4 (2.7)
	Statistics and Probability	34.7 (5.6)	33.2 (4.2)	56.8 (7.9)	43.1 (7.7)	37.4 (5.4)	33.3 (4.4)	32.5 (5.1)
Grade 4 Urdu and Sindh	Grammar	51.9 (1.5)	45.1 (2.7)	73.2 (2.4)	61.9 (1.7)	49.8 (3.2)	33.6 (3.6)	36.8 (3.6)
	Reading for information or task	65.5 (1.9)	56.7 (1.6)	81.6 (1.5)	68.8 (2.2)	65.7 (1.9)	58.2 (2.7)	61.7 (2.1)
	Reading for literary experiences	70.3 (2.7)	62.9 (2.2)	85.2 (2.0)	74.1 (2.4)	71.5 (2.8)	65.6 (3.5)	66.4 (3.4)
	Vocabulary	62.4 (4.1)	55.3 (2.2)	77.5 (3.6)	68.0 (3.1)	60.9 (4.5)	53.9 (5.3)	53.0 (6.1)

Subject	Content domain	Percentage of possible marks achieved (indicative standard error)						
		Balochistan	KP & NMD	Punjab	Sindh	AJK	GB	ICT
Grade 8 Maths	Algebra	27.7 (2.0)	31.3 (1.8)	51.5 (1.9)	38.9 (1.9)	26.9 (2.2)	31.1 (2.6)	31.4 (2.5)
	Measurement and Geometry	32.7 (5.7)	37.2 (6.2)	55.8 (5.4)	39.7 (3.6)	33.7 (5.8)	37.3 (6.4)	36.6 (7.0)
	Numbers and Operations	36.4 (3.5)	43.6 (3.7)	58.5 (4.8)	47.7 (4.1)	41.5 (4.5)	45.3 (4.2)	45.7 (4.9)
	Statistics and Probability	22.8 (2.2)	27.6 (3.4)	42.3 (4.1)	32.7 (4.4)	25.2 (2.9)	29.2 (3.9)	31.6 (4.6)
Grade 8 Science	Earth and Space Sciences	42.7 (6.3)	37.3 (4.5)	54.0 (5.0)	43.4 (5.2)	37.5 (5.5)	37.2 (4.5)	36.2 (5.6)
	Life Sciences	57.7 (2.9)	49.3 (3.2)	72.4 (2.9)	55.6 (3.0)	48.0 (3.9)	49.5 (3.5)	47.3 (3.7)
	Physical Sciences	43.6 (4.0)	42.2 (2.7)	61.9 (3.2)	46.5 (2.8)	41.2 (3.6)	40.1 (2.8)	38.6 (3.4)

6. Analysis of factors associated with student performance

6.1. Main findings in this section

The results in this section show:

1. On average, teachers achieved much higher scores than students. More interestingly, the analysis shows a clear link between teacher and student performance. It was rare for the average performance of students in a school to exceed the performance of their teacher. This highlights that the knowledge of teachers acts as a limit on the possible achievement of their students and reiterates the obvious need for teachers to be well trained in the subjects they teach.
2. The strongest link with student performance from the teacher and headteacher questionnaires was the extent to which Grade 4 teachers report that the course had been completed in time. In schools where teachers reported that the course was 'always completed in time', students achieved higher test scores on average and were far less likely to have scores at or below a level that would be expected by guessing. It is worth noting that teachers were most likely to report that the course always ended in time in Punjab. As such, this may explain a small part of the differences in performance across provinces. Note that the importance of course completion was also highlighted after NAT 2016[15].
3. A large number of factors from the student and parent questionnaires were significantly associated with student performance. In particular, questions associated with the assignment, completion and checking of homework by teachers consistently emerged as being significantly associated with attainment. The analysis shows that students who are assigned homework, complete homework and receive feedback from their teachers tended to achieve higher scores on average than those who did not. They were also far less likely to display scores that were no better than would be expected by guessing. While caution should be taken in assuming that these associations represent causal links, it would be worth considering the impact of homework as an area for further research. Note that the importance of homework being assigned, checked, and corrected was also noted in reporting on NAT 2016.
4. Analysis of links between the student and parent surveys and attainment also revealed that students tended to achieve higher scores in English when this was taught in their local or mother tongue at least for some of the time.
5. Where students felt they could express their ideas in class and their parents stated their child had self-confidence, the students tended to achieve higher scores. Considering the impacts of self-expression and self-confidence may also be a useful area for further research.

[15] See Dissemination of National Achievement Test Findings 2016–2017, Fifth Stakeholders Conference, National Education Assessment System Ministry of Federal Education and Professional Training Islamabad: https://allchildrenlearning.org/wp-content/uploads/2020/01/Presentation-disseminating-findings-from-Nat-2016-17.pdf.

6.2. Overview of factors associated with student performance

This section describes some of the main results from the analysis, exploring the association between student performance and the information gathered from students, parents, teachers and headteachers within the questionnaires. The aim of this analysis was to find and report on factors that had the largest and most obvious associations with student performance.

The analyses are limited to students where relevant matched questionnaire data could be found. Within each grade, data on student performance was linked to student and parent questionnaires based on the school identification details (ID) and the roll number of the student. Linkage to teacher and headteacher questionnaires was achieved using school IDs.

For the vast majority of schools, only one teacher completed a questionnaire. To simplify the analysis, the small minority of schools with more than one response from a teacher were removed from the analysis of the association between student performance and the teacher questionnaires. A similar approach was taken for the headteacher questionnaires.

This section initially explores the relationship between achievement and certain factors that were determined to be of initial interest, whether or not the relationship was statistically significant. Specifically, the analysis first presents some descriptive statistics on the association of performance with teacher qualifications, parental education, students' socio-economic status as captured by the possessions in their home and student attendance at school.

After this, the analysis explored the statistical significance of the association between performance on each assessment (excluding Grade 4 FL[16]) and the responses to every question in every questionnaire where there was clean data. Any questions that allowed respondents to answer freely have not been included. At the time of writing, the responses to such open-ended questions (including all questions asking respondents to specify precisely "how much" or "how many") were answered in a format that was too inconsistent to allow easy inclusion in analysis within the available timescales.

The statistical significance of associations was estimated using multilevel modelling. Specifically, for every question in the surveys and every subject, a separate model was fitted looking at how student performance varied according to responses to the survey question. The multilevel models had two levels relating to schools and students within schools. In this way, the models accounted for the way student assessment data was collected from samples of schools rather than completely at random. As noted earlier, in each of these schools, 20 students were selected for inclusion in the NAT. Every multilevel model also included fixed effects for the impact of provinces. This information was included as, from the analysis in previous sections, it was clear that there were major variations in performance between provinces and it is necessary to ensure that findings in this section moved beyond this. There is no interest in associations between performance and survey responses if these can be explained purely in terms of certain survey responses being more likely in some provinces than in others.

The main focus for this section is on the factors that are statistically significantly associated with all assessments in the relevant grade (that is, with all of Maths, English and Urdu in Grade 4 or both Maths and Science in Grade 8), with the aim of identifying the most obvious and largest associations. Due to the very large number of factors (hundreds) that were tested for their association with performance, in

[16] This was considered less relevant for this section due to the generally very high performance on this test and the possible influence of some of the anomalies in the data that were described earlier. Factors influencing performance on Grade 4 FL could be an area for future research.

order to prevent reporting results that have occurred due to random chance, the analysis uses a slightly stricter definition of statistical significance $p<0.01$ (rather than $p<0.05$ as in earlier sections).

In addition to the above, this section also provides an analysis of the association between the performance of teachers on each test and the performance of students in the same school.

Across all sections, it is extremely important to bear in mind that a statistically significant association does not imply causality. It is not known whether the behaviours and attitudes recorded in questionnaires cause high performance, whether they reflect things that may be a result of high performance (e.g. self-confidence), or whether they are attributes of students and schools that performed well for other reasons. Also worth noting is that there is no guarantee that policies aimed at changing the behaviours associated with high performance will lead to improved outcomes for students. Rather the results in this section should be taken as a stimulus for thinking about what approaches to improving education could be considered, researched or trialled.

6.3. Initial exploration of relationship between performance and factors of interest

This section explores the relationship between NAT performance and some factors that are often explored as a preliminary step in educational research.

Student performance and teacher qualifications

To begin with, Table 24 and Table 25 show the relationship between the academic qualifications of teachers and the raw test scores achieved by their students. The majority of students were within classes taught by teachers with either a BA/BSc/AD qualification or an MA/MSc/BS (Hons). In every subject, there was less than a single score point difference between these two groups. Students taught by a teacher with an MPhil/PhD had noticeably higher scores in each Grade 4 test, but this pattern was not repeated in Grade 8. This differs from other studies on students of similar ages. For example, in a cross-national study specifically looking at Grade 4 performance, there was no evidence that higher level qualifications, beyond a first degree, had any impact on student achievement in Trends in International Maths and Science Study (TIMSS) (Luschei and Chudgar, 2011[17]). There was no clear pattern for teachers with lower-level academic qualifications.

A fairly consistent finding in the literature in this field is that teacher knowledge (and qualification) in the relevant subject does impact on achievement (Monk, 1994[18]; Goldhaber and Brewer, 2000[19]). It must be noted that the number of teachers with lower-level qualifications in the current study was fairly small, meaning caution should be taken about reading too much into this fact. Teachers' academic background warrants further investigation in the Pakistani context.

[17] Luschei, T.F. and Chudgar, A. (2011). Teachers, student achievement and national income: a cross-national examination of relationships and interactions. *Prospects*, *41*(4), 507–533.

[18] Monk, D.H. (1994). Subject area preparation of secondary Maths and science teachers and student achievement. *Economics of Education Review*, *13*(2), 125–145.

[19] Goldhaber, D.D. and Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, *22*, 129–145.

Table 24: The relationship between teachers' academic qualifications and student performance in Grade 4

What is your academic qualification?	Number of students in relevant schools	Mean raw score of Grade 4 students			Standard deviation		
		English	Maths	Urdu & Sindhi	English	Maths	Urdu & Sindhi
Matriculation (Grade 10)	271	26.1	23.3	35.8	13.5	10.2	11.8
Intermediate (Grade 12)	443	26.7	21.5	36.5	12.1	10.3	12.3
Diploma	89	26.1	19.7	39.9	11.9	7.4	9.4
BA/BSc/AD	1847	26.1	22.7	35.0	11.7	10.1	11.8
MA/MSc/BS (Hons)	3899	26.8	23.4	35.5	12.0	10.3	12.0
MPhil/PhD	818	30.0	27.2	37.3	11.7	10.7	12.0

Table 25: The relationship between teachers' academic qualifications and student performance in Grade 8

What is your academic qualification?	Number of students in relevant schools	Mean raw score of Grade 8 students		Standard deviation	
		Maths	Science	Maths	Science
Matriculation (Grade 10)	412	20.2	25.3	9.8	10.0
Intermediate (Grade 12)	80	18.9	26.1	8.0	12.8
Diploma	85	18.7	19.9	7.5	8.0
BA/BSc/AD	1138	20.7	25.4	9.5	10.5
MA/MSc/BS (Hons)	5316	19.9	25.1	8.0	9.2
MPhil/PhD	883	21.3	25.9	7.7	9.6

Table 26 and Table 27 show the relationship between student performance and teachers' professional qualifications in Grade 4 and Grade 8 respectively. In Grade 4, most students were taught by teachers with either a BEd/Based, an MEd/MSEd or a Primary Teaching Certificate (PTC). Although the scores for students of teachers with the PTC qualification was slightly lower than the other two most widely held qualifications, further investigation revealed this difference was not significant after accounting for the clustering of students within schools.

In Grade 8, most students were taught by teachers with either a BEd/BSEd or an MEd/MSEd. There was very little obvious difference in their performance.

Table 26: The relationship between teachers' professional qualifications and student performance in Grade 4

What is your professional educational qualification?	Number of students in relevant schools	Mean raw score of Grade 4 students			Standard deviation		
		English	Maths	Urdu & Sindhi	English	Maths	Urdu & Sindhi
Primary Teaching Certificate	1394	25.0	21.6	33.8	12.2	9.9	12.5
Certificate of Teaching	377	24.3	20.8	34.6	11.3	8.7	12.2
Diploma in Education	126	28.6	28.4	37.7	11.3	12.4	13.6
BEd/BSEd	2974	27.7	24.3	36.3	12.2	10.3	11.6
MEd/MSEd	1916	27.0	24.0	36.7	11.4	10.6	11.3
Other	339	28.3	21.5	33.8	11.6	9.7	12.9

Table 27: The relationship between teachers' professional qualifications and student performance in Grade 8

What is your professional educational qualification?	Number of students in relevant schools	Mean raw score of Grade 8 students		Standard deviation	
		Maths	Science	Maths	Science
Primary Teaching Certificate	331	20.4	26.4	7.3	9.7
Certificate of Teaching	597	21.4	26.5	10.7	11.9
Diploma in Education	40	12.2	18.9	3.2	5.4
BEd/BSEd	2509	19.8	25.1	7.7	9.5
MEd/MSEd	3793	20.1	25.1	8.2	9.2
Other	549	22.3	25.3	9.1	9.1

Student performance and parental education

The questionnaire contained several questions about the academic qualifications of parents or guardians: one for the father, one for the mother and one for a guardian. Surprisingly, analysis found no obvious link between the level of education recorded for parents and the students' performance. This finding is at odds with most academic research on this issue, including results from NAT 2016[20]. Having reflected on these results and explored various ways of analysing the data, we believe the lack of any association may indicate an issue with the way this piece of data was collected. The question relating to parental education had a fairly complex format, with questions about the educational levels of parents and guardians condensed within a fairly small space. Furthermore, the survey question appeared to require parents to indicate whether they were illiterate based upon a written questionnaire. We suggest

[20] See Dissemination of National Achievement Test Findings 2016–2017, Fifth Stakeholders Conference, National Education Assessment System Ministry of Federal Education and Professional Training Islamabad: https://allchildrenlearning.org/wp-content/uploads/2020/01/Presentation-disseminating-findings-from-Nat-2016-17.pdf.

that the format of this question is reviewed and, ideally, piloted, ahead of NAT 2025 to ensure that we can trust we have accurate information on the educational levels of parents.

Student performance and (use of) possessions in home

Table 28 and Table 29 display the relationship between the number of books parents report that they have in their home (excluding textbooks) and student performance. The majority of parents reported having no more than 20 books in their home. The performance of students in homes with 1 to 20 books was slightly higher on average than that of students with no books in their home[21]. Given the small numbers of students in homes with a larger number of books, it is not possible to draw firm conclusions from the remainder of the data in these tables. It is also important to note that there were no questions asking whether parents read these books with their children (or any books), or what type of books they owned.

Table 28: The relationship between number of books in the home (excluding textbooks) reported by parents and student performance in Grade 4

How many books are there in your house apart from textbooks?	Number of students	Mean raw score of Grade 4 students			Standard deviation		
		English	Maths	Urdu & Sindhi	English	Maths	Urdu & Sindhi
Not even one	4440	26.3	23.1	34.7	12.0	10.5	12.4
From 1 to 20	4987	27.7	24.4	36.2	11.8	10.5	11.8
From 21 to 40	366	25.0	21.9	34.1	11.8	9.9	11.6
From 41 to 60	106	24.7	22.4	35.4	11.7	10.3	11.3
Over 60	389	25.7	21.6	34.8	12.0	10.0	11.9

Table 29: The relationship between number of books in the home (excluding textbooks) reported by parents and student performance in Grade 8

How many books are there in your house apart from textbooks?	Number of students	Mean raw score Grade 8 of students		Standard deviation	
		Maths	Science	Maths	Science
Not even one	3398	20.8	25.4	8.8	9.7
From 1 to 20	6004	22.2	27.4	9.5	10.2
From 21 to 40	899	21.1	26.7	8.8	10.0
From 41 to 60	302	22.0	27.4	9.6	10.2
Over 60	863	21.3	26.3	8.7	10.0

[21] This difference was statistically significant (after accounting for the impacts of provinces and clustering within schools) for both Maths assessments and Grade 8 Science but not for Grade 4 English or Urdu/Sindhi.

Table 30 and Table 31 consider the relationship of performance with possessions beyond books in the home. Each student questionnaire contained a list of nine items where students could indicate whether they "use" them at home. Although questions of this nature are commonly used to capture a proxy for socio-economic status, it is important to note that the questionnaire does not specifically ask whether students own each item, but rather whether they "use" them. As such, this is not a pure measure of socio-economic status but also a measure of behaviour.

The tables compare the mean raw scores of candidates who have indicated they use each item against the mean for those who do not. To make it easier to see the main pattern in this data, the same information is shown in visual form in the subsequent figure.

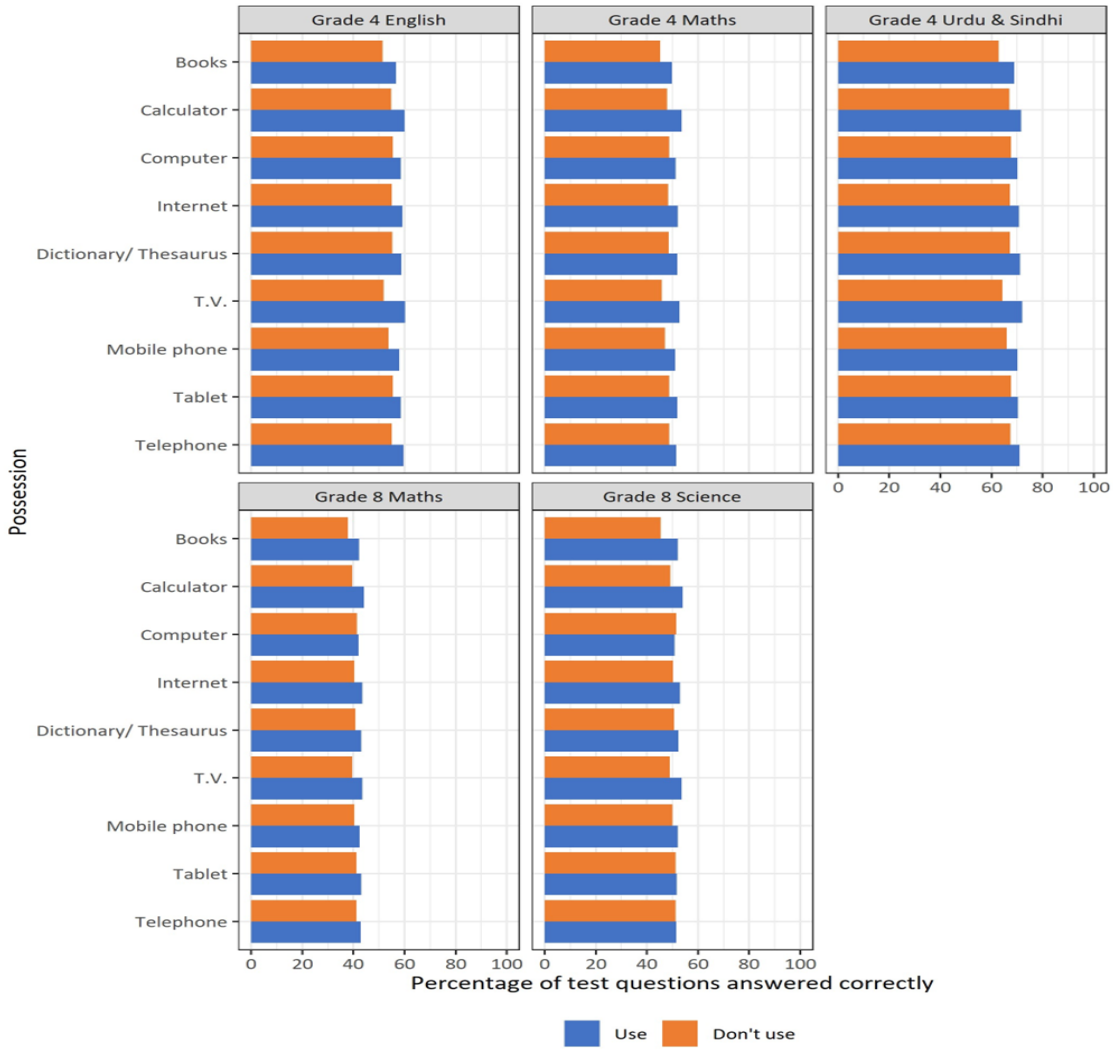


Figure 26 shows the average percentage of test questions answered correctly by students who use each item (the blue bars) against the average percentage for those who do not (the orange bars). As can be seen, in all but one case (Grade 8 Science and using a computer at home), students who own and use each item tend to perform better than those who do not.

Table 30: Mean raw scores in each Grade 4 assessment by whether or not students use various items in their home

		Mean raw score of Grade 4 students	Standard deviation

Do you use the following things at your home?	Number of students		English		Maths		Urdu & Sindhi		English		Maths		Urdu & Sindhi	
	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Books	8927	1320	27.2	24.7	23.9	21.7	35.8	32.7	11.9	11.9	10.5	10.2	11.9	12.4
Calculator	2408	7671	28.8	26.3	25.7	23.0	37.2	34.9	12.0	11.8	10.7	10.3	12.0	12.0
Computer	1964	8105	28.1	26.6	24.6	23.4	36.4	35.2	12.3	11.8	10.9	10.4	12.0	12.0
Internet	2507	7564	28.4	26.4	25.0	23.2	36.8	35.0	11.9	11.9	10.6	10.4	12.0	12.0
Dictionary/ Thesaurus	2230	7845	28.2	26.5	24.9	23.3	37.0	35.0	11.8	11.9	10.5	10.5	11.6	12.1
TV	5076	5070	28.9	24.9	25.4	22.0	37.5	33.4	11.7	11.8	10.5	10.2	11.5	12.3
Mobile phone	5444	4723	27.8	25.8	24.5	22.6	36.5	34.3	11.9	11.8	10.6	10.2	11.8	12.2
Tablet	1733	8345	28.1	26.6	24.9	23.4	36.6	35.2	11.7	11.9	10.7	10.4	12.2	12.0
Telephone	2177	7886	28.6	26.4	24.7	23.4	36.9	35.0	11.9	11.9	10.5	10.5	11.9	12.1

Table 31: Mean raw scores in each Grade 8 assessment by whether or not students use various items in their home

Do you use the following things at your home?	Number of students		Mean raw score of Grade 8 students				Standard deviation			
			Maths		Science		Maths		Science	
	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Books	9972	1450	21.9	19.6	27.1	23.6	9.3	8.2	10.1	9.5
Calculator	5034	6240	22.9	20.6	28.1	25.6	9.4	8.9	10.1	9.8
Computer	2999	8210	21.9	21.5	26.5	26.7	9.3	9.2	10.1	10.1
Internet	4799	6477	22.6	20.9	27.5	26.1	9.4	9.0	10.2	9.9
Dictionary/ Thesaurus	4310	6937	22.4	21.1	27.2	26.3	9.2	9.2	9.9	10.1
TV	6074	5232	22.6	20.5	27.8	25.4	9.6	8.6	10.2	9.7
Mobile phone	6851	4478	22.0	21.0	27.1	26.0	9.4	8.9	10.1	10.0
Tablet	2142	9036	22.4	21.4	26.8	26.6	9.7	9.1	10.2	10.0
Telephone	2666	8528	22.2	21.4	26.8	26.6	9.3	9.2	10.1	10.0

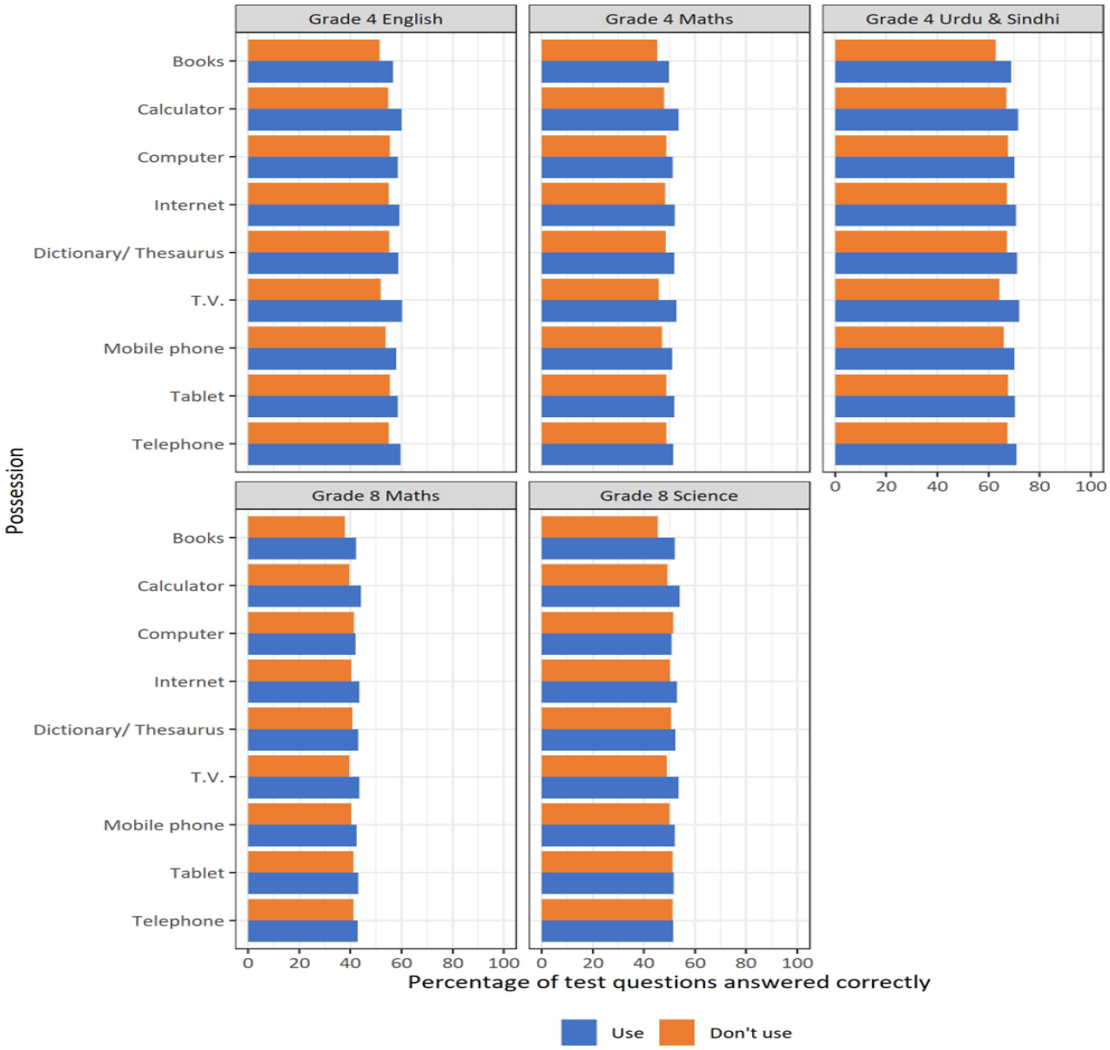


Figure 26: Average percentage of test questions answered correctly in each assessment by whether or not students use various items in their home

To look at this same issue another way, Table 32 and Table 33 show the relationship between the *number* of listed items that students indicate having and using (up to nine) and the average scores of students. As can be seen, in general, the greater the number of listed items the student owns and uses, the higher their average score. This trend is more prominent in Grade 4 than in Grade 8. Furthermore, the apparent impact of each additional item appears less obvious once students own and use more than four of the listed items[22].

Table 32: The relationship between student scores in Grade 4 and the number of listed items they report having and using in their home

		Mean raw score of students	Standard deviation

[22] The highest mean scores occur for students who use eight out of nine and seven out of nine items in Grades 4 and 8 respectively, with slightly lower mean scores for students in the very highest category. However, given the very small percentage of students saying they use all nine of the listed items, it would be dangerous to read too much into this fact.

Number of listed items that student indicates using in home	Number of students	English	Maths	Urdu & Sindhi	English	Maths	Urdu & Sindhi
0	240	22.5	19.8	29.7	12.3	9.6	12.4
1	2089	23.2	20.8	31.9	11.2	9.8	11.9
2	1954	26.4	22.8	35.2	11.7	9.9	11.9
3	2119	27.6	24.0	35.8	11.9	10.7	11.8
4	1622	28.5	25.4	37.0	11.7	10.7	12.0
5	1095	29.4	25.4	37.9	11.9	10.6	11.3
6	548	29.1	26.2	38.2	11.7	10.5	11.3
7	302	30.2	26.4	39.4	11.9	10.5	11.0
8	169	30.6	28.3	40.1	11.8	10.4	10.9
9	155	26.5	23.0	33.0	12.3	10.4	12.7

Table 33: The relationship between student scores in Grade 8 and the number of listed items they report having and using in their home

Number of listed items that student indicates using in home	Number of students	Mean raw score of students		Standard deviation	
		Maths	Science	Maths	Science
0	204	19.0	22.4	7.7	9.4
1	1422	19.3	24.1	8.8	9.7
2	1481	20.3	25.8	8.4	9.6
3	1944	21.0	26.2	8.6	9.6
4	2080	22.2	27.5	9.2	10.1
5	1863	22.6	27.8	9.6	10.3
6	1228	23.1	27.9	9.7	10.5
7	638	23.8	28.4	9.7	10.3
8	292	22.3	26.7	9.4	10.0
9	314	23.7	27.3	9.7	9.6

Student performance and student attendance

To conclude these initial explorations, Table 34 and Table 35 show the mean raw scores of students according to the self-reported extent to which they are absent from school. The majority of students stated that either they were never absent, or they were absent just once or twice each month. There was no sign of students who were absent once or twice a month performing worse than those who were never absent. However, in Grade 8, the minority of students who were absent even more often than this displayed performance levels below that of those who were present at school more frequently.

Table 34: The relationship between student performance in Grade 4 and self-reported absence

How often are you absent from school?	Number of students	Mean raw score of Grade 4 students			Standard deviation		
		English	Maths	Urdu & Sindhi	English	Maths	Urdu & Sindhi
Not at all	3023	25.6	22.9	34.2	11.8	10.4	12.2
1 to 2 times a month	5942	27.6	24.2	36.3	11.9	10.5	11.8
(Up to)[23] 5 times a month	861	26.7	23.1	34.8	11.7	10.4	12.4
More than 5 times a month	467	26.0	21.9	33.6	11.6	10.1	12.4

Table 35: The relationship between student performance in Grade 8 and self-reported absence

How often are you absent from school?	Number of students	Mean raw score Grade 8 of students		Standard deviation	
		Maths	Science	Maths	Science
Not at all	3769	21.8	26.4	9.2	10.1
1 to 2 times a month	6524	21.9	27.2	9.3	10.1
(Up to) 5 times a month	786	19.9	24.6	8.8	9.4
More than 5 times a month	387	19.1	24.0	7.9	9.6

6.4. Exploring other factors significantly related to performance

The previous section showed the relationship between specific factors that are often measured within educational research and performance. This section moves on to explore the largest and most obvious relationships between performance and the remainder of factors measured within the NAT background questionnaires.

The relationship between teacher and student performance

Figure 27 shows the relationship between the average scores of teachers in each school and the average scores of students (as noted above, there is usually only one teacher in the data for each school). Schools

[23] The words "Up to" do not appear in the English translation of the student questionnaire. However, based on the surrounding test, that would appear to be the intended meaning. The same is true for the Grade 8 questionnaire.

with fewer than five students were removed from the analysis and each point on each chart represents a school. A dashed line of equality has been added to each chart. A point exactly on this line indicates that students in the given school have achieved the same score, on average, as their teacher. Roughly 90% of all points are below this diagonal line, indicating that teachers tend to perform better than students.

More importantly, aside from obvious outliers (relating to extremely low teacher performance), it is very rare for students to noticeably outperform their teacher, regardless of the level of teacher performance. This pattern is most clearly visible for Grade 8 Maths. For this scatterplot (in the bottom left panel of Figure 27), while students may have performed poorly even if their teacher had achieved a high score, no students achieved a high average score when their teacher had a low score. In fact, the performance of teachers almost forms a hard limit on the performance of students.

In very broad terms the analysis here suggests that students cannot know more than their teachers. This reinforces the importance of teachers being properly trained in their subjects. It also emphasises the value of checking that teaching applicants have a good level of understanding of the subjects they will be teaching. This tallies with a consistent finding in the literature that teachers' knowledge appears to be positively correlated with student achievement (Luschei and Chudgar, 2011[24]).

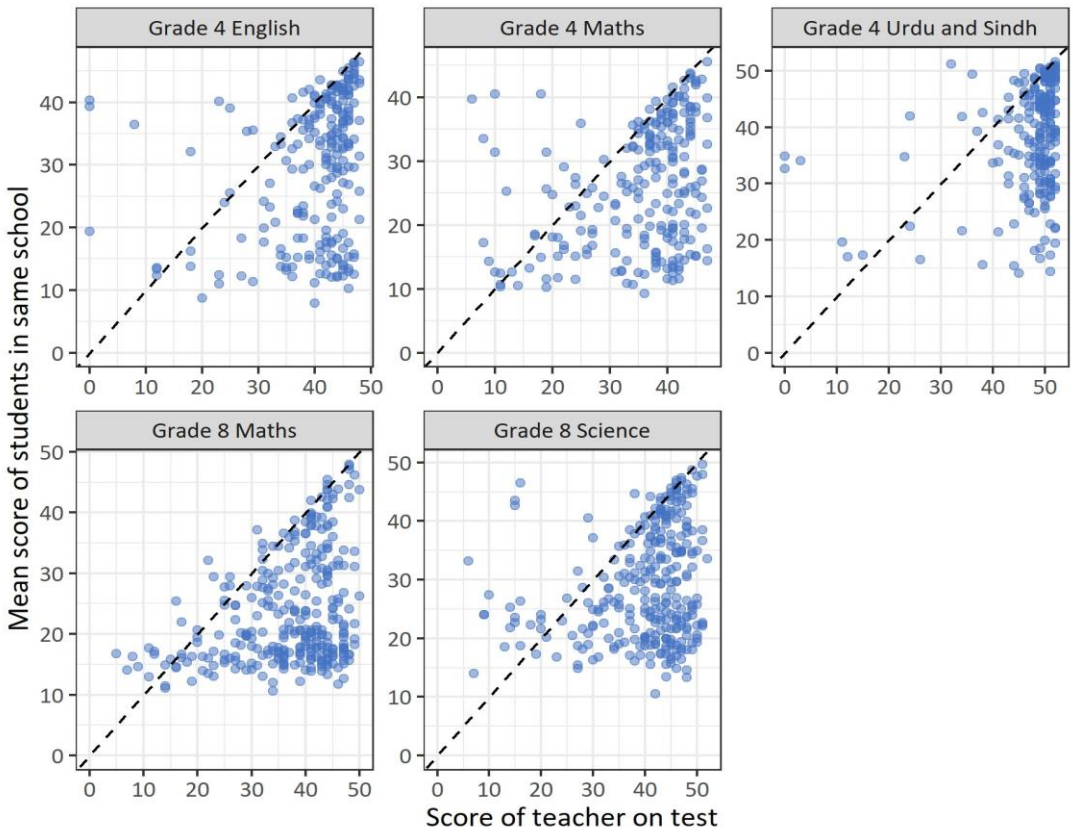


Figure 27: A scatterplot of the average scores of teachers in each school against the average scores achieved by students.

[24] Luschei, T.F. and Chudgar, A. (2011). Teachers, student achievement and national income: a cross-national examination of relationships and interactions. *Prospects*, *41*(4), 507–533.

School and teacher questionnaire items related to performance

There were just two questions in the Grade 4 teacher survey that were significantly ($p<0.01$) associated with performance in all three subjects: English, Maths and Urdu/Sindhi. The first of these (and the one with the highest level of significance) was question 25b from the teacher survey, which asked teachers to state the extent to which they feel "the course ends in time". Teachers could choose the responses "Never", "Occasionally", "Quite often" or "Always".

Table 36 shows how performance in each Grade 4 subject is associated with teachers' responses to this question. As can be seen, while the majority of students are in schools where teachers say that the course "Always" ends in time, many are not. It can also be seen that, as the extent of completion reported by teachers increases, the mean scores of students increase in all three subjects. Figure 28 provides further data on the proportion of students achieving no more than 25% of the available marks (i.e. the average amount achievable through pure guessing) within each category. For example, Figure 28 shows that, in Grade 4 Maths, students in schools where teachers say the course "Never" finishes on time are three times more likely to perform at or below a level commensurate with guessing than those where the course "Always" finishes on time. Strong associations can also be seen in the other two Grade 4 subjects.

Table 36: The relationship between the extent to which Grade 4 teachers say "the course ends in time" and performance in each Grade 4 subject

To what extent would you say the course finishes on time?	Number of students in relevant schools	Mean raw score of students			Standard deviation		
		Maths	English	Urdu and Sindhi	Maths	English	Urdu and Sindhi
Never	518	15.5	18.8	27.8	6.3	9.2	11.2
Occasionally	884	20.3	23.9	33.4	9.5	11.7	12.2
Quite often	2026	23.0	26.3	34.4	10.3	11.5	12.0
Always	4012	25.3	28.9	37.8	10.4	12.1	11.4

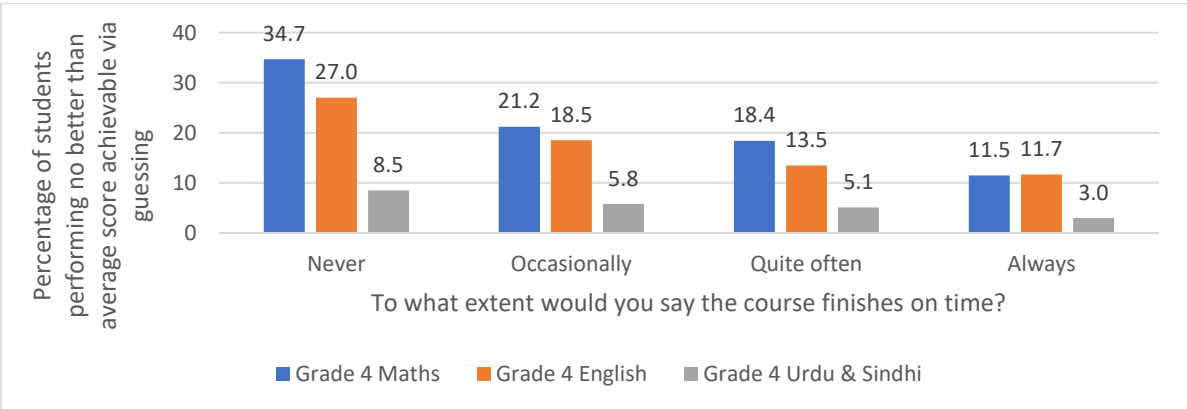


Figure 28: The proportion of Grade 4 students performing at or below the average score achievable via pure guessing (i.e. 25% of the total available in each assessment) by the extent to which Grade 4 teachers say "the course ends in time"

To explore this topic further, Table 37 shows the percentage of teachers giving each response to this question in each province. As can be seen, Grade 4 teachers in Punjab were more likely than those in any other province to say that the course "Always" ends in time. They were also by far the least likely to say that it "Never" ends in time. This may explain a small proportion of the difference in performance between provinces seen in the previous section.

Table 37: Extent to which Grade 4 teachers report that "the course ends in time" in each province

The course ends in time	% of Grade 4 teachers giving each response in…						
	Balochistan	KP & NMD	Punjab	Sindh	AJK	GB	ICT
Never	19%	6%	1%	6%	13%	8%	9%
Occasionally	19%	15%	10%	14%	5%	0%	9%
Quite often	31%	38%	21%	27%	28%	32%	22%
Always	31%	41%	68%	53%	55%	60%	61%
Number of teachers	42	93	139	93	40	25	23

The second question in the Grade 4 teacher questionnaire that was significantly associated with performance in all three subjects asked English teachers the extent to which they use the "Grammar translation method" to teach English to Grade 4 students. Since this question relates to English teaching, the association with English performance will be the area of focus.

Table 38 shows how performance in Grade 4 English varies according to teachers' answers to this question. Note that, since this question was only relevant where there were responses from English teachers, the numbers of students in this analysis is smaller than in the earlier tables. While the association is not as dramatic as for the course ending on time, the table shows that students in schools where teachers "Often" or "Always" use the grammar translation method perform better, on average, than students in schools where this is not the case.

Table 38: Student performance on Grade 4 English by the extent to which they use the Grammar translation method

To what extent do you use the "Grammar translation method" to teach English to Grade 4	Number of students in relevant schools	Mean raw score	SD	% students at or below guessing level
Never	631	24.4	10.8	17.6
Sometimes	578	22.8	12.6	24.0
Often	1516	28.9	11.5	9.5
Always	1505	27.2	12.0	12.8

The only question for the Grade 4 headteacher questionnaire that was significantly associated with performance in all three subjects was a question (Q9) regarding the nature of the headteacher's job and

whether it was permanent, contract or other. Only two headteachers mentioned "Other" as a possibility so the focus will be on the comparison between headteachers with permanent and contract jobs.

Table 39 shows the number of students in each type of school, along with the means and standard deviations of test scores. As can be seen, although only a small number of students attended schools where the headteacher was not permanent, scores in these schools were substantially higher than elsewhere. Further research and more detailed information would be required to better understand why this might be the case. Note that the cause is not due to the impact of provinces, as this was taken into account for the original models that tested significance.

Table 39: Student performance in Grade 4 by nature of the headteacher's job

Nature of [headteacher's] job	Number of students in relevant schools	Mean raw score of students			Standard deviation		
		Maths	English	Urdu and Sindhi	Maths	English	Urdu and Sindhi
Permanent	7474	23.3	26.4	35.3	10.4	11.8	12.0
Contract	382	31.5	35.7	44.0	8.9	9.6	8.2

For the Grade 8 data, no survey questions in either the teacher or the headteacher questionnaire were statistically significantly ($p<0.01$) associated with performance in both Maths and Science. However, there were four survey questions in the headteacher questionnaire that displayed a significant relationship with science. One of these related to the gender of the headteacher, another to whether the school was for boys, girls or provided co-education. There was also another question about whether headteachers communicated with parents to provide information about punitive activities.

More detailed analysis revealed that all three of these significant relationships could be explained by further accounting for the gender of the students within the schools. Since gender differences have already been explored earlier (section 4.4), there is no need to expand upon these effects any further. The final question that displayed a significant association with science achievement asked about the frequency with which headteachers face challenges due to "Un-interesting teaching learning material". Table 40 provides more details and shows that students in schools where headteachers stated such challenges "Never" occurred had higher test scores.

Table 40: Student performance in Grade 8 by frequency of challenges with un-interesting teaching and learning material

How frequently do you have challenges with un-interesting teaching learning material?	Number of students in relevant schools	Mean raw score of students		Standard deviation	
Never	3959	23.3	29.0	9.9	10.6
Seldom	3264	20.9	25.5	8.9	9.8

How frequently do you have challenges with un-interesting teaching learning material?	Number of students in relevant schools	Mean raw score of students		Standard deviation	
Often	1855	20.4	25.1	8.6	9.4
Always	496	19.7	24.8	8.0	9.1

There were also two survey questions in the teacher questionnaire that were significantly associated with performance in Maths (but not Science). The two questions regarded:

- the extent to which teachers said they were given subjects for teaching according to their interest
- the extent to which teachers said they evaluated students based on participation in class activities.

Further analysis also revealed that, for the first question in the list, the statistical significance disappeared after further accounting for the impact of student gender. Analysis of the second question appeared to relate to the slightly higher performance of students in schools where teachers stated that they "Never" evaluate their students based on participation in class (mean score of 22.8 in Maths compared to a mean of roughly 20 in all other groups). Given the small size of the effect and the small number of students it relates to (less than 700), we will not expand on this result any further here.

Student and parent questionnaire items related to performance

Within the Grade 4 student and parent questionnaire, 32 questions were found to be significantly associated ($p<0.01$) with performance in all of English, Maths and Urdu/Sindhi. Ninety-eight questions were found to be significantly associated with performance in at least one subject. There may be more significant effects in this analysis than in any investigation of the relationship with teacher and headteacher questionnaires, as it is possible to take advantage of exploring differences between students within the same school. This is unlike the analysis of association with teacher and headteacher questionnaires where it is limited to looking at variation between schools only. This indicates plenty of scope for further research into the student and parent factors that have an impact upon performance. For the purposes of this section, reporting has been limited to the top 10 most significant factors, both in the student questionnaire and in the parent questionnaire. These are listed in

Table 41[25].

A large proportion of the most significant associations are related to questions about homework. These are highlighted in bold in

Table 41. Questions about the extent to which students were assigned homework, completed it, had it checked and had mistakes' identified and corrected were all significantly associated with performance in the assessments. Questions about homework accounted for nine of the top 10 associations with performance in the student questionnaire and four of the top 10 in the parent questionnaire. These

[25] A crude method of ranking by overall significance was developed by multiplying the p-values for the associations with the different subjects together and then sorting from lowest to highest values. Survey questions that did not have a significant association with performance in all subjects were removed from the list.

findings that relate to the importance of homework in student achievement are in line with the literature in this field (Cooper, Robinson and Patall, 2006[26]; Marzano and Pickering, 2007[27]), as well as with findings reported after NAT 2016[28].

Table 41: The top 10 survey questions from the Grade 4 student and parent questionnaires with the most significant associations with performance

Student questionnaire		Parent questionnaire	
Number	**Question**	**Number**	**Question**
15_b	**Do you complete the assigned homework? – English**	**18_b**	**Free textbooks are provided on time**
16_c	**Do your teachers check your assigned homework? – Maths**	**18_d**	**You are satisfied with school's performance regarding your child's education**
15_c	**Do you complete the assigned homework? – Maths**	**18_e**	**Academic assessment (tests, assignments, etc.) of the child is conducted from time to time**
14_b	**Do your teachers assign you homework? – English**	**5**	**Does your child do homework?**
14_c	**Do your teachers assign you homework? – Maths**	**10_c**	**Do the teachers check the homework of your child? – Maths**
17_c	**Do your teachers identify and correct mistakes of your assigned homework? – Maths**	**11_c**	**Do the teachers identify and correct the mistakes in the homework of your child? – Maths**
16_b	**Do your teachers check your assigned homework? – English**	**12_b**	**Teachers use local language to teach? – English**
17_b	**Do your teachers identify and correct mistakes of your assigned homework? – English**	**25_d**	**Does your child participate in academic activities at school?**
18_b	**To what extent do your teachers use local or mother language to explain the lesson during teaching learning process? – English**	**25_g**	**Does your child have self-confidence?**
15_a	**Do you complete the assigned homework? – Urdu**	**10_b**	**Do the teachers check the homework of your child? – English**

[26] Cooper, H., Robinson, J.C. and Patall, E.A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, *76*(1), 1–62.
[27] Marzano, R.J., and Pickering, D.J. (2007). Special topic: The case for and against homework. *Educational Leadership*, *64*(6), 74–79.
[28] See Dissemination of National Achievement Test Findings 2016–2017, Fifth Stakeholders Conference, National Education Assessment System Ministry of Federal Education and Professional Training Islamabad: https://allchildrenlearning.org/wp-content/uploads/2020/01/Presentation-disseminating-findings-from-Nat-2016-17.pdf.

In order to provide a succinct illustration of the relationship between homework and performance, the analysis focused on the association between Maths homework and Maths performance. The various survey questions in the student questionnaire about homework were combined together to classify students into the following six categories:

1) The student states that they are "seldom" or "never" assigned any Maths homework.
2) The student is "often" or "always" assigned Maths homework but "never" or "seldom" completes it.
3) The student is "often" or "always" assigned Maths homework and completes it, but their teacher "never" or "seldom" checks it.
4) The student is "often" or "always" assigned Maths homework, completes it and it is checked, but their teacher "never" or "seldom" identifies and corrects their mistakes.
5) The student is "often" or "always" assigned Maths homework, completes it, it is checked, and the teacher identifies and corrects their mistakes. However, at least one step in the process is only labelled "often".
6) The student is "always" assigned Maths homework, completes it, it is checked, and the teacher identifies and corrects their mistakes.

Table 42 shows the relationship between the six categories of engagement in homework and performance in Maths. In broad terms, the higher the level of engagement in Maths homework, the higher the average level of Maths performance. Based on the final column of the table, it is clear that students who are "always" assigned homework, complete it, have it checked, and whose mistakes are corrected were half as likely to perform no better than guessing. This additive effect of not simply assigning homework but feeding back fully, closing the "homework loop", is also seen in other research in this field (for example, Murillo and Martinez-Garrido, 2014[29]; Dettmers, Trautwein, Lüdtke, Kunter and Baumert, 2010[30]).

Table 42: The association between Grade 4 students' self-reported levels of engagement with homework and performance in Maths

Engagement with Maths homework	Performance in Maths			
	Number of students	Mean raw score	SD	% students achieving less than 25% of maximum
Homework seldom or never assigned	2614	20.2	9.6	22.5%
Often assigned but not completed	853	23.2	10.3	17.0%

[29] Murillo, F. and Martinez-Garrido, C. (2014). Homework and primary-school students' academic achievement in Latin America. *International Review of Education*, *60*(5), 661–681.
[30] Dettmers, S., Trautwein, U., Lüdtke, O., Kunter, M. and Baumert, J. (2010). Homework works if homework quality is high: Using multilevel modelling to predict the development of achievement in Maths. *Journal of Educational Psychology*, *102*(2), 467–482.

Engagement with Maths homework	Performance in Maths			
	Number of students	Mean raw score	SD	% students achieving less than 25% of maximum
Often assigned and completed but not checked by teacher	398	21.9	9.8	18.1%
Often assigned, completed and checked but not corrected by teacher	708	22.4	10.3	18.6%
Often assigned, completed and corrected by teacher	1333	25.2	10.6	13.6%
ALWAYS assigned, completed and corrected by teacher	4111	25.8	10.5	11.7%

The only question in the top 10 for Grade 4 students that did not relate to homework asked whether English teaching occurred in the students' own local language. A similar question is also part of the top 10 most significant results from the parent questionnaire. The association between teaching English to students in their own language and performance in English is illustrated in Table 43. In particular, this shows that students where teaching is "never" done in their own language achieved lower scores on average than those where their own language was used at least some of the time.

Table 43: The association between teaching in a Grade 4 student's own language and performance in English

To what extent do your teachers use local or mother language to explain the lesson during teaching learning process? – English	Performance in English			
	Number of students	Mean raw score	SD	% students achieving less than 25% of maximum
Never	2604	24.6	11.9	18.6%
Seldom	2176	27.6	11.8	11.4%
Often	1514	29.5	11.4	10.2%
Always	3722	26.9	11.9	13.5%

Focusing on the Grade 4 parent questionnaire, a number of the significant associations come from Q18. This block of questions in the survey asked parents to express their opinion on a range of statements about their child's school, including whether free textbooks are provided on time and whether they are satisfied with the school's performance.

For illustrative purposes, Table 44 shows the relationship between parents' levels of satisfaction with the school and performance in English. As can be seen, students with parents who were "always" or "often" satisfied with the performance of the school tended to achieve higher marks than those with parents who were "never" or "seldom" satisfied. It is not clear whether these results occur because

parents' satisfaction stems from their children being able to do well in tests, or whether schools that have good relationships with parents has a directly beneficial impact for students.

The pattern of the relationships of the other significant Q18 questions with performance was very similar to that shown below.

Table 44: The association between Grade 4 parents' satisfaction with a school's performance and students' performance in English

To what extent are you satisfied with school's performance regarding your child's education?	Performance in English			
	Number of students	Mean raw score	SD	% students achieving less than 25% of maximum
Never	1677	22.6	11.5	22.2%
Seldom	817	24.6	11.7	19.7%
Often	1519	28.5	11.8	10.5%
Always	5998	28.0	11.8	11.6%

Turning to Grade 8, a total of 49 questions were significantly ($p<0.01$) associated with achievement in both Maths and Science and 81 were significantly associated with at least one of these. Table 45 shows the top 10 survey questions in the student and parent questionnaire that are most significantly associated with performance in Maths and Science. As before, many of these questions relate to homework. As was found with the analysis of Grade 4, a number of the most significant questions in the parent survey come from Q18.

For brevity, we will not expand upon either of these relationships any further in this report. Instead, the analysis will provide further illustrative details of the association with performance of the most significant questions not related to homework in each survey.

Table 45: The top 10 survey questions from the Grade 8 student and parent questionnaires with the most significant associations with performance

Student questionnaire		Parent questionnaire	
Number	Question	Number	Question
14_b	Do your teachers assign you homework of the following subjects? – Maths	25_g	Does your child have self-confidence?
13	How much time do you study at your home?	18_a	Teachers provide helpful guidance by checking homework
24_c	Can you express your ideas in class?	25_d	Does your child participate in academic activities at school?
11_e	Do you use the following things at your home? – Dictionary/ Thesaurus	18_e	Academic assessment (tests, assignments, etc.) of the child is conducted from time to time

Student questionnaire		Parent questionnaire	
21_d	To what extent does the school provide a conducive environment for learning?	18_d	You are satisfied with school's performance regarding your child's education
15_b	Do you complete the assigned homework? – Maths	11_a	Do the teachers identify and correct the mistakes in the homework of your child? – Science
21_g	To what extent do your teachers use the blackboard during teaching?	18_f	You are satisfied with the educational environment of the school
16_b	Do your teachers check your assigned homework? – Maths	11_b	Do the teachers identify and correct the mistakes in the homework of your child? – Maths
17_b	Do your teachers identify and correct your mistakes in your assigned homework? – Maths	18_c	You contact the school regarding the child's performance
21_f	To what extent do your teachers see your homework and give useful guidance?	24_e	To what extent does your child have problems relating to remembering lessons?

Table 46 shows the relationship between Grade 8 students stating that they can express their opinion in class and performance in each subject. Students who said they could do this achieved higher scores on average in each assessment and were much less likely to have performed at or below a level that might be expected by guessing.

Table 47 shows the relationship between performance and whether parents say their child has self-confidence. There is clearly a thematic link between this table and the previous one. Table 46 is about self-expression and Table 47 is about self-confidence, both of which are clearly associated with achievement in the tests. As with all of the results in this section, the causality of these relationships is unknown; in particular, it is not clear whether self-confidence is a driver of high performance or comes as a result of it. Nonetheless, the results in this section may provide a useful starting point for future research into identifying ways of improving student outcomes.

Table 46: The association between performance and whether Grade 8 students say they can express their opinion in class

Can you express your opinion in class?	Number of students	Mean raw score		SD		% students achieving less than 25% of maximum	
		Maths	Science	Maths	Science	Maths	Science
Yes	8971	22.1	27.3	9.3	10.1	14.8%	6.1%
No	2495	19.8	24.3	8.6	9.6	22.2%	9.8%

Table 47: The association between performance and Grade 8 parents reporting whether their child has self-confidence

Does your child have self-confidence?	Number of students	Mean raw score		SD		% students achieving less than 25% of maximum	
		Maths	Science	Maths	Science	Maths	Science
Yes	8624	22.1	27.2	9.3	10.1	14.7%	6.1%
No	2842	20.2	25.0	8.8	9.9	21.7%	9.1%

6.5. Summary of the main factors associated with student learning

Much of the analysis in this section can be summarised in the following three key elements of teaching that seem to be associated with high performance in the NAT:

1. **"Know it":** The data shows that, on average, students rarely performed better than their own teacher. To put this another way, in broad terms, a student cannot know more than their teacher. This emphasises the crucial importance of ensuring that every teacher has a full understanding of the content they are teaching.
2. **"Teach it":** The data displayed a very strong association between the extent to which teachers stated that "The course ends in time" and student performance. For example, in Grade 4 Maths, students whose teacher said the course "Never" ends in time were three times more likely to perform at or below a level that might be expected by guessing than those students where the teacher stated the course "Always" ends in time. Although the importance of completing teaching is obvious, many teachers report that this is not achieved.
3. **"Check it":** Questions on the student and parent questionnaires relating to homework consistently displayed a highly significant association with performance. This includes questions relating to whether homework was assigned at all, whether it was completed, whether it was checked, and whether mistakes were identified and corrected. This suggests that using homework to check that students have understood what is being taught and that misunderstandings are dealt with is an important part of teaching. Another thread of the "check it" theme is that of periodic assessment. In both the Grade 4 and Grade 8 parent surveys, a significant association was found between children being assessed periodically and their achievement in the NAT.

While it is not possible to infer from the NAT data alone whether the associations summarised above are causal, all of the above factors might be considered as standard recommendations relating to high-quality teaching. However, the significance of the results in this report is that it is clear that these simple ideals are not always being achieved and the analysis provides quantitative evidence of the impact on children's learning.

The data also showed that Grade 4 students performed better (particularly in English) where English teachers frequently used the Grammar translation method. This correlation should be considered carefully as the Grammar translation method is not robustly supported by educational research. The data also suggested that Grade 4 students performed better in schools where the headteacher did not

have a permanent job. However, this finding was only relevant to a relatively small number of students and should be treated with some caution.

Based on data from the student and parent surveys and attainment, it could also be seen that students tended to achieve higher scores in English when this subject was taught in their local or mother tongue at least some of the time. Students also tended to achieve higher scores in schools where parents were satisfied with the performance of the school.

Finally, the data suggested that, where students felt they could express their ideas in class and their parents stated they had self-confidence, students tended to achieve higher scores. When considering these results, as mentioned above, caution is needed in assuming that identified relationships are causal. However, the results in this section will be useful for stimulating thinking and suggesting areas where further research may be valuable.

7. Policy recommendations

Section 7 focuses on the findings of the 2023 NAT and the implications of this for policy and practice in Pakistan. A series of recommendations is presented under key areas of focus. It should be noted that many of the recommendations centre around the need for further investigation at local level. These should be taken as a starting point for context-specific inquiry and decision making.

As part of the implementation of the recommendations, it is important that a layered monitoring framework is developed. This must have clearly defined parameters at each level to ensure that recommendations are acted upon and their impact is evaluated. As part of this framework, it is important to establish a clear escalation path to ensure that any issues are addressed in a timely manner.

7.1. Specific suggestions for each of the key areas

The 2023 NAT has generated a rich data set that has illuminated many aspects of the Pakistan schooling system. As with any investigation, it has also given rise to further areas of study and further questions that need to be asked. Many of these areas can be investigated specifically via targeted questions in the 2025 NAT. However, to build robust policy decisions for current students and to prevent further delay, it is recommended that a further survey is sent to all headteachers as a direct result of the 2023 NAT.

Ideally this would be an online tool that would facilitate easy dissemination, data capture and analysis. Without additional more targeted data, recommendations for policy and practice can only be general. For example, it is not clear why courses were not completed. Understanding why a teacher was unable to finish the course may help to identify trends, either at national or provincial level, to enable the development of a targeted intervention. The 2023 NAT results clearly demonstrate the importance of course completion. This is not a surprising result. A more important consideration in terms of policy would be to identify the challenges that hindered course completion. The NAT data cannot answer that, hence the need for a further round of data collection.

A recurring theme of the recommendations that follow is the role of headteachers and school leaders. It is very important that the insights and experiences of these key stakeholders are taken into consideration when decisions on training and development needs are taken. The NAT is a sample-based monitoring tool. As such, the recommendations or findings will not be relevant to all schools. It is important for headteachers and other senior leaders in schools to be briefed on the key findings of the report and then supported to develop their own lines of enquiry from the findings to better understand which aspects are relevant to their context. This can form the basis of good school and leadership self-evaluation upon which school leaders can develop a framework that works best for their context.

Through building evaluative practices from within a school, rather than simply relying on dictated monitoring from external agencies, school leaders can ensure that they focus on and develop the areas that are of most relevance to them and will be able to gain the approval of other key stakeholders in their school community and drive successful change (Kyriakides and Campbell, 2004[31]). Using data to develop evaluative enquiry and to make subsequent decisions on areas for change is a highly skilled set of activities. It is essential that headteachers and other school leaders are supported and provided with

[31] Kyriakides, L. and Campbell, R.J. (2004). School self-evaluation and school improvement: a critique of values and procedures. *Studies in Educational Evaluation*, *30*, 23–36.

training and guidance. By supplying schools with a follow-up survey, headteachers can use this as a basic scaffold upon which to build their evaluative framework.

7.2. Curriculum and instruction

Recommendations relating to course completion

An important finding of the 2023 NAT is the strong association between the course being completed on time and students' achievements. It is essential that this topic is explored further. School leaders will understand more fully the context of their schools and the reasons behind the lack of course completion. If there have been significant periods of school closure (for example, as a result of COVID-19 or flooding), this will clearly impact on a teacher's ability to cover the entire course. However, if a school has not been subject periods of closure, it is very important for headteachers and other school leaders to determine the reasons behind courses not being completed. Obviously, this is a sample-based survey and the issues reflected may not be relevant to individual schools. However, school leaders can use any findings as a prompt for further discussion and inquiry within their school.

If course completion is shown to be an issue within a school, then consideration needs to be given as to why this might be the case and appropriate action should be taken. For example, do teachers need support with curriculum planning? Are there issues with the positioning of holidays or topic breaks that hinder course completion? Is it a consistent picture across a school or just for certain teachers? The findings of the NAT are a powerful tool for headteachers to use as the basis of enquiry and self-evaluation.

It is also important for system-wide monitoring that the issue of course completion is more fully understood. A key recommendation is that all schools are invited to participate in an online survey that picks up on some of the key questions that have arisen as a result of the NAT, including course completion. This may indicate a wider issue around course completion that stems from the volume of material that needs to be covered. It could also indicate that the curriculum for a particular subject is too full for the available number of teaching hours.

A further consideration that may be related to course completion is the requirement of teachers to undertake additional duties that could be carried out by other staff members. It is important that teaching time is not lost through teachers having to undertake such duties.

Recommendations relating to homework

Homework featured highly in the top 10 associated factors with student achievement in both the Grade 4 and Grade 8 parent and student questionnaires. The results of the NAT highlight the critical role that homework plays in a student's progress and achievement. While the assignment of homework alone was associated with increased achievement, checking and then feeding back to the student through identifying and correcting mistakes unsurprisingly had the greatest impact. This highlights the need for quality homework provision. Homework should be a *planned* part of the curriculum and schools need clear homework policies in place to ensure that the full benefits of appropriate homework are realised,

through thorough checking and constructive, individual feedback (Murillo et al, 2014[32]; Dettmers et al, 2010[33]).

While it is acknowledged that many schools will have effective homework policies in place and teachers are skilled in providing meaningful homework and feedback, the data from the NAT 2023 results clearly show that this is not universally the case. School leaders may therefore benefit from support in developing rigorous homework policies for their settings that balance the needs of the students, the degree of parental support of their 'typical' families, the skills of their teachers and the time pressures that they are under. It is important that homework is authentic, relevant and practically manageable. The NAT results show that providing individual and timely feedback is essential to maximise the benefits of homework. Where a rigorous policy needs development, teachers may also need support with its implementation; for example, in being able to give useful formative feedback from homework tasks. Area Education Assessment Centres (AEACs) could help with cascading training related to effective formative feedback for teachers.

This should form the focus of in-service training for schools to ensure that all teachers are equipped with the skills necessary to provide this vital service to their students. The results from the NAT demonstrate that access to quality homework and feedback is not consistent across the country. Initial teacher training providers should use these results to reinforce the need for timely and effective formative feedback. While the evidence from the NAT is almost exclusively in the English and Maths subject domains, all subjects will benefit from an agreed whole-school approach to homework.

One route that may be appropriate for reaching all schools and ensuring a basic, consistent approach to homework is through the use of dedicated homework booklets that track progress alongside the student textbooks. While it is acknowledged that such a prescriptive approach would not fully embrace the concept of personalising the tasks to students' needs, it may be a helpful tool for teachers and schools. Despite previous policy directives, homework setting, marking and effective feedback from it remains an area of concern. The use of workbooks could help in this regard. The use of a dedicated workbook to accompany textbooks would also help parents to support their children and provide a manageable and structured approach to ensure basic provision for all students. An additional tool could be the use of a homework diary/journal in which students record their homework and which parents countersign to indicate that the work has been completed. This extra level of focus may help to raise the profile of homework in both the classroom and at home.

While this approach may be an appropriate way of ensuring all students receive a basic, common level of homework provision and that this may be the most feasible way forward in the context of many of Pakistan's schools, a note of caution is necessary. To maximise the benefit of homework, not only must all steps be carried out (including feeding back and integrating into the teaching, as evidenced by the findings of NAT 2023) but also homework should ideally be tailored to the student (Baker, LeTendre and Akiba 2005[34]). Clearly, a generic homework workbook will not achieve this aim. However, evidence from the 2023 NAT highlights the importance of homework, even if the ideal scenario of individualised tasks

[32] Murillo, F. and Martinez-Garrido, C. (2014). Homework and primary-school students' academic achievement in Latin America. *International Review of Education*, *60*(5), 661–681.

[33] Dettmers, S., Trautwein, U., Lüdtke, O., Kunter, M. and Baumert, J. (2010). Homework works if homework quality is high: Using multilevel modelling to predict the development of achievement in Maths. *Journal of Educational Psychology*, *102*(2), 467–482.

[34] Baker, D. P., LeTendre, G.K. and Akiba, M. (2005). Schoolwork at home? Low quality schooling and homework. In D.P. Baker and G.K. LeTendre (Eds.), *National differences, global similarities: World culture and the future of schooling* (pp.117–133). Stanford, CA: Stanford University Press.

is not achievable in the current context. It is therefore recommended that a more comprehensive strategy on homework and its feedback is included in the future course of action for schools.

Recommendations relating to assessment

Closely aligned with further development and implementation of homework policies in schools is the topic of assessment. In both the parent and student questionnaires, conducting periodic academic assessments was positively associated with student performance in the NAT at both Grade 4 and Grade 8 levels. The importance of assessment in good teaching practice is well documented and this links with the very strong theme of the importance of homework and the "know it, teach it, check it" headline findings (section 6.5).

Some school leaders may need support in developing appropriate assessment policies for their school based on contextually relevant best practice. This is an area in which the AEACs could play a significant role in showcasing examples of good practice and providing school leaders with the skills required to support successful implementation in their schools. As with implementing a homework policy, it is likely that teachers will also need support in this area. Timely and effective assessment is not a new priority and there will be many schools in which suitable assessment policies are being successfully implemented. However, the evidence from the NAT results is that the picture is inconsistent, and this must therefore remain an area of focus for policy makers.

The use of homework workbooks could form part of the assessment process in a school, by providing timely formative feedback that teachers can use to target their lessons to the needs of their students. Assessment must be planned into the curriculum and school year to ensure a minimum acceptable frequency in this regard. Despite previous directives, not all schools are implementing a regular effective assessment regime. Developing and implementing a rigorous but student-centred assessment policy is not an easy task and it may require a shift in mindset, a willingness at the leadership level and the acquisition of new skills for all stakeholders. Time and resources should be given to support schools to build their expertise in assessment.

Recommendations relating to the language of instruction

Headteachers should be briefed on the significant association between performance and the use of local or mother tongue language while teaching English. Teachers should be encouraged to use the local language to support their students in English lessons when students find it difficult to grasp concepts. This is not to take the place of the majority of the dialogue in the classroom being in English. Simply that if students require help but cannot understand that support if delivered in English, they are not going to make the desired progress. A centrally generated policy recommendation encapsulating this suggestion may help headteachers to ensure acceptance by teachers.

Recommendations relating to the provision of library facilities

The majority of parents in both Grade 4 and Grade 8 reported having 20 books or fewer in the home. At Grade 4, 4440 students came from homes without any books and at Grade 8 this figure decreased to 3398. The figures for those homes without books are not dissimilar to those parents who reported being illiterate or below primary level education. The critical importance of reading to and listening to children read in terms of their development and subsequent academic performance is well documented.

The data on homes without books and parents who are illiterate/below primary level education indicate that a large proportion of children in Pakistan are not getting their basic reading needs met. This makes the significance of reading and library provision in schools of absolute importance, especially at primary level when children are learning to read. The findings from the 2023 NAT indicate that only 11.6% of primary schools (and only 73.3% at Grade 8) have a functioning library that can be used by teachers and students. It is recommended that improving the library provision in schools should be a priority. Access to the library is also a key area of focus, with the recommendation of a dedicated library period each week. It is hoped that this will dovetail with the recent launch of the Foundational Learning Hub and the reading hour entitlement. The findings of the NAT highlight the crucial importance of this new policy, given the need to ensure quality provision within schools.

7.3. Teacher preparation, recruitment, and professional development

Recommendations relating to teacher subject knowledge

One of the strongest findings (statistically) of the 2023 NAT was the impact of teacher achievement on student achievement; the performance of the teacher acted as a ceiling for the performance of the student. This effect was most strongly seen at Grade 8, presumably due to the more demanding nature of the content. This highlights the critical importance of teachers being appropriately qualified and updated in the specific subjects that they will be required to teach, especially at higher grades. Teachers must have the required depth of knowledge (Luschei and Chudgar, 2011[35]). This is an important message for senior leaders in schools as they will understand the subject background of their staff. If teachers require extra training, this must be prioritised, particularly on new aspects of the 2022 National Curriculum. A wider review of this situation should be undertaken to determine whether a targeted recruitment campaign could be the solution.

The distributions for student performance in the Grade 4 subjects (Foundational Literacy aside) are multimodal, with a clear peak at or around the rate expected due to guessing, but there is also a peak at a much higher level of performance. This indicates that there are some students who are achieving very well and have a firm grasp of the subject matter. However, the distribution of performance in Grade 8 subjects is notably different. There is still a strong peak at or around the guessing rate but there is not a clear peak of high achievement. While the average student performance is greater for Grade 8 Science than for Maths, this lack of higher performance peak indicates that the teaching and learning at this level in these subjects needs to be improved.

The reasons for this difference could be multifaceted, from the provision of resources and engagement of students to the confidence and practical competence of the teachers, not just their command of the subject matter. It is important to determine what the underlying reasons are to enable the issue to be addressed, as there are often interactions between the different elements (Monk, 1994[36]).

Given the lower performance of teachers in the Grade 8 subjects when compared to those at Grade 4, it is likely that teachers' skillsets need to be improved. A key policy recommendation is for school leaders to undertake a survey of the background and skills of their teaching staff to identify any staff that require additional training to meet the standard of knowledge necessary to teach their subject at the required

[35] Luschei, T.F. and Chudgar, A. (2011). Teachers, student achievement and national income: a cross-national examination of relationships and interactions. *Prospects*, *41*(4), 507–533.
[36] Monk, D.H. (1994). Subject area preparation of secondary Maths and science teachers and student achievement. *Economics of Education Review*, *13*(2), 125–145.

level (and also to identify other factors that could be influencing the effectiveness of their teaching, including changes in curriculum). This information should be passed to AEACs to enable appropriate planning of interventions. For example, it may be that a targeted series of face-to-face or video tutorials is needed. It is likely that there are inter- as well as intra-provincial similarities in training needs. Producing a central training package that could either be delivered remotely or cascaded through the local/provincial assessment centres is a cost-effective way of ensuring that all areas receive the same core materials, training and development opportunities.

It is important that training is targeted to the specific needs of the teaching staff, especially when time and/or funds are limited. The need for school leaders to undertake a thorough training-needs analysis of their staff should not be underestimated. When interviewing representatives from the different provincial assessment organisations in June 2022, Cambridge found clear evidence of a lack of teacher training on the new National Curriculum, despite training programmes having been directed at federal level. The results of the 2023 NAT suggest that there are still teachers who need training and support.

Recommendations relating to 'Grammar'

In Grade 4 Urdu and Sindhi, a significantly smaller proportion of the available marks for Grammar were answered correctly than for other content domains. It is not clear from the data set whether this lower performance is due to a lack of emphasis in the teaching or the teaching pedagogy, or the level of teacher skill in this area. It would be beneficial for headteachers to investigate the provision of grammar in Urdu and Sindhi within their schools. This enquiry could be supported via an online survey that addresses many of the further queries that the results of the NAT have highlighted.

The Grammar translation method is a traditional language teaching methodology that focuses on students learning grammatical rules and applying them through translation. In Grade 4 English, the "use of the Grammar translation method" for teaching was shown to impact positively on achievement.

Both findings highlight the need for a focus on the teaching of grammar. These findings warrant further investigation regarding current practice in schools around grammar instruction. A suggested topic for training to be delivered from an AEAC (or via an online tutorial) could be the sharing of ideas and good practice in the teaching of grammar. Teachers who proficiently and explicitly teach grammar could be asked to present at one of these events to showcase their practice and the impacts it has on their students.

Alongside communicative language teaching methodologies that promote students' ability to use the target language in authentic situations, methods for teaching grammar explicitly should be covered as part of initial teacher training and any Teacher Induction Programmes (TIPs) currently in place.

Recommendations relating to Maths

The performance of students and teachers in Maths, at Grade 8 in particular, is below that of the other subjects. Interestingly, girls perform significantly better than boys in all subjects, except for Maths (where there is no statistically significant difference in performance). These two findings may be interrelated and linked to the quality of Maths provision through teacher skills or confidence. It is not clear whether the poor performance of students and teachers (compared to other subjects) in Maths is due to teachers lacking the requisite knowledge through being non-subject specialists (teaching specialism was not included as part of the teacher questionnaires) or due to the style and approach of Maths teaching. Further investigation regarding the teaching specialism/subject background of teachers

is necessary to determine whether it is necessary to recruit and train more Maths specialists, or train existing staff.

The overall performance in Maths adds weight to the assumption that there may be insufficient Maths specialist teachers. It also supports a policy suggestion that older students should be taught the core subjects by teachers who are subject specialists in those areas. If there is a shortage of teachers in those areas, then a targeted recruitment campaign may be required, alongside onboarding and training needs assessment. Schools may need support for teachers if they have to teach outside their specialist areas for older students. There could be a role for the AEACs to act as learning hubs, hosting subject-specific workshops. The development of subject-specific online tutorials or webinars may also help to reach teachers in more remote areas.

The very poor performance of students and teachers in the Statistics and Probability content domain, a recent addition to the curriculum following the introduction of the National Curriculum in 2022, indicates that teachers need further training to teach this subject. If it was only students who scored badly in this domain, then it could be concluded that they had simply not been taught the material. However, the statistically significant underperformance of the teachers highlights the need for better subject knowledge and building capability in this area.

It is essential that teachers of Grade 8 Maths are given targeted support to teach this area of the content domain. For example, a specific bulletin that covers the key facts, common misconceptions and lesson activities could be circulated. Ideally, teachers would be offered the opportunity for face-to-face training (for example, at an AEAC). If this is not possible, the production of a series of online video tutorials for teachers to watch at a time convenient for them could be considered. It is important that school leaders understand the need for this additional training to help improve results so that it can be facilitated and monitored as required.

Recommendations relating to the implementation of the new National Curriculum

To make sure the NAT is suitable in length for the age of the students, it only samples from the Student Learning Outcomes (SLOs). This is not an uncommon feature of tests and should not be viewed as a shortcoming. In addition, the NAT necessarily focused on a few core subjects at each grade. However, given the very poor performance in Statistics and Probability in Maths, it may be that other areas of the 2022 curriculum also need to be an area of focus for training and development of teachers.

It is very important to conduct a thorough review of whether the 2022 National Curriculum is being implemented. This could feature as part of the 2025 NAT. However, given that this is two years away, it is recommended that this should be investigated sooner so that the aims of the new curriculum can be realised. Any shortfalls can then be addressed and monitored via the next NAT. Targeted questions relating to the implementation of the 2022 National Curriculum should form part of the suggested online survey for school leaders.

Recommendations relating to recruitment and retention of teachers

The NAT data set does not explicitly hold information on the subject specialisms of teachers. In order to determine whether a targeted recruitment and retention campaign is needed, this information should be gathered via the suggested online survey for school leaders. For science subject specialisms outside of the Life Sciences and Maths, if it is shown that there are insufficient teachers with the appropriate subject background, a targeted recruitment campaign should be considered. This could include reaching

out to the relevant subject-specific professional body. In addition, recruitment incentives (for example, in the form of a bonus or university-fee coverage) could be considered if recruitment of target numbers is not reached.

Recommendations relating to headteachers

When considering training and professional development needs, headteachers' needs must not be forgotten. Many headteachers will have had this type of training previously, but there are always people new to the role and refresher workshops can be very helpful for those who are more experienced too. It is recommended that headteachers receive a briefing in how to interpret the results of the NAT and how to best apply them to their setting. Regional headteacher network events could be an effective way of sharing this message and providing support. In addition, through collaboration at such events and looking outwards, headteachers can learn a lot about their own schools. The results of the NAT could be used as the stimulus for discussion and the basis for school self-evaluation. The senior leaders of the schools could use the results and questions from the NAT as a scaffold to evaluate their setting. Where self-evaluation is driven from within and developed by the leaders of a school, for their school, it is more effective at improving learning than rigid accountability measures imposed by external agencies. Headteachers could be provided with a series of prompts, with themes drawn from the results of the NAT to act as a starting point for building a culture of self-evaluation and improvement in their school. If this is a new area for headteachers, then support should be prioritised.

An interesting finding of the 2023 NAT was that, in the Grade 4 headteacher questionnaire, the only factor that was significantly associated with increased achievement in all three subjects was the nature of the headteacher's employment contract. While only a small number of students attended schools where the headteacher's contract of employment was not permanent, scores in these schools were substantially higher than elsewhere. This finding requires further investigation, and it is recommended that additional research is carried out to try to understand the reasons behind this. For example, has the new headteacher brought with them more recent approaches to pedagogy or are they encouraging more effective parental engagement?

7.4. Parental involvement and student engagement

Recommendations relating to parent involvement

It is important to engage parents as key stakeholders as soon as possible after their child joins a school, and this should be re-affirmed with each year and new class of instruction. Hosting meetings or workshops to introduce/remind parents of their role in supporting their children and to provide them with some practical tools and ideas is very valuable. Curriculum briefings or newsletters that summarise the key themes and topics to be covered each term can help parents to feel involved in their child's schooling.

Ensuring that parents are given feedback about their children on a regular but sustainable basis is fundamental to help maintain the link between home and school. Having flexibility around the timings of these meetings (for example, offering daytime and evening slots) will help more parents to participate. If infrastructure allows, schools should also consider using technology to facilitate meetings between parents and their child's teacher (through the use of video-calling software). Establishing a regular and open dialogue with parents is critical for their understanding of the school's life and to build

a culture of trust and engagement. Learning-focused discussions can help all members of the school community.

Providing non-judgemental and targeted support for parents with core literacy and numeracy skills is one way to improve parental engagement, as well as equipping parents with the core skills necessary to support their children and to be able to engage with their learning. Research shows that a parent's level of education will influence their views on whether they have sufficient skills to engage with school and support their children (Green, Walker, Hoover-Dempsey and Sandler, 2007[37]).

The data from the 2023 NAT indicates that there is a need for this parental support. A large number of parents/guardians of both Grade 4 and Grade 8 students only reported having education levels up to primary level (3972 at Grade 4 and 2902 at Grade 8). Of these, 1871 Grade 4 parents/guardians were illiterate/below primary, and 1493 reported a similar education level at Grade 8. Only 10% of parents have reached the educational level of Diploma or above. Parental booster classes or workshops could be an effective way of helping parents to support their children at school. Effective schools focus on the learning of the whole community of stakeholders. Parents must not be forgotten as an integral part of this learning journey, even though it can be a challenging exercise in some communities.

Recommendations relating to student engagement

Children's self-confidence, as indicated by their parents, was positively associated with achievement at both Grade 4 and Grade 8. Self-confidence at school or about schooling can indicate good engagement with students' learning. It cannot be determined from the current datasets whether the increased achievement is a result of this self-confidence or whether the self-confidence has grown through increased achievement. However, an important area for further research, as a direct result of the findings of the NAT, should be to investigate this link between self-confidence and achievement, and to determine whether taking self-confidence as a proxy for engagement is also a valid assumption.

In the interim, policies directed at improving students' confidence are recommended. For example, through building a focus on personal reflection, and supporting children to develop this important metacognitive skill, schools can help children to become more self-aware and more able to think about their strengths and areas they need support with. Through open and supportive dialogue with their teachers (and parents), reflective students can engage more fully with their work. They also have increased confidence by knowing where they are currently and where they need to be regarding their academic and social skills. The positive association between self-expression and performance for Grade 8 students is also relevant to this discussion and may indicate that schools operating a more open and reflective practice with their students achieve better results. This is an avenue that warrants further research.

Students need support when it comes to developing metacognitive skills. Teachers can provide them with the necessary scaffolding through modelling in the classroom and adequate opportunities for such activities. It is recommended that this is a planned and coordinated approach throughout the school and, as such, this can form part of a whole-school commitment and policy. Increasing school leaders' and teachers' confidence in this area is important and there is likely a need for increased emphasis in initial teacher training and workshops specifically focused on teaching children how to learn and how to

[37] Green, C.L., Walker, J.M.T, Hoover-Dempsey, K.V. and Sandler, H.M. (2007). Parents' motivations for involvement in children's education: An empirical test of a theoretical model of parental involvement. *Journal of Educational Psychology*, *99*(3), 532–544.

reflect on their learning. Continuing the shift away from rote learning, which was a core objective of the new curriculum and SLOs, and focusing on deepening learning through engaging students in meaningful activities and exciting conversations will be useful approaches in developing students' metacognitive skills.

While no significant association between school absence and performance was found at Grade 4, for Grade 8 this was not the case. For students who identified as missing five or more school days per month, there was a significant decrease in performance in both subjects. While the absolute numbers of students are small (as the vast majority reported being present in school far more regularly), this does highlight the need to emphasise the importance of school attendance.

The correlation between higher rates of absence and lower achievement has been shown in multiple and varying countries (for example, Carroll, 2010[38]; Paredes and Ugarte, 2011[39]), highlighting the importance of regular school attendance. A recommendation is to ensure that schools make their expectations on attendance clear to all children and families. It is essential that everyone understands the link between attendance and performance.

This information should form part of the initial briefings upon entering a school and at the beginning of each school year. A targeted poster or social media campaign might be a useful way to encourage students to attend school as much as possible and to help reinforce this importance with parents too. It is important that all students are targeted, not just those from schools with a low socio-economic intake. The impacts of absenteeism on student achievement have been shown to be similar regardless of students' socio-economic status and, as such, all schools should focus on reducing absenteeism (Hancock, Lawrence, Shepherd, Mitrou and Zubrick, 2017[40]).

[38] Carroll, H.C.M. (2010). The Effect of Pupil Absenteeism on Literacy and Numeracy in the Primary School. *School Psychology International*, *31*(2), 115–130.

[39] Paredes, R.D. and Ugarte, G.A. (2011). Should students be allowed to miss? *Journal of Educational Research*, *104*(3), 194–201.

[40] Hancock, K.J., Lawrence, D., Shepherd, C.J.C., Mitrou, F. and Zubrick, S.R. (2017). Associations between school absence and academic achievement: Do socioeconomics matter? *British Educational Research Journal*, *43*(3), 415–40.

8. Limitations of the 2023 NAT and recommendations for future iterations

The preparation and successful administration of the 2023 NAT should be celebrated given the prevailing circumstances and very recent formation of NAW. It has provided much useful data that can be used to improve the school system in Pakistan. However, there are limitations with the current data set and, as such, there are areas that can be improved upon for the 2025 NAT. Section 8 outlines the most significant limitations of the current study by key areas and makes suggestions as to how to remedy these for the 2025 iteration of the NAT.

Limitations and recommendations relating to sampling

1) Increased sample size

The 2023 NAT sample size was greater than the previous test in 2019, which was a great achievement given time and other constraints. The number of schools involved, however, is still smaller than many of the provincial-level large-scale assessments. This may result in different results at national versus provincial level. Where there are conflicting results between provincial large-scale assessment results and those of the NAT, the provincial-level results should be used to inform provincial priorities. A greater number of sample schools, and therefore a greater number of learners, is ideally required for the 2025 NAT.

Increasing the number of sample schools will also allow for teacher- and headteacher-effects to be analysed. Only one teacher (usually) per school completes the teacher questionnaire and there is only one headteacher questionnaire submitted. This means that the sample size is limited to that of the number of schools. This is a limitation of the current survey and makes it very difficult to draw statistically valid inferences and recommendations at this level. To fully understand, for example, the impact of years of experience or time since the most recent teacher training, there needs to be a larger data set. The current data cannot be used for this purpose.

2) Inclusion of private schools

Approximately one-third of students in Pakistan attend private schools but these learners did not form part of the sampling strategy (due to logistical and time constraints) for the 2023 NAT. This means that the results offer an incomplete picture of the state of education in Pakistan and cannot be used to make decisions for the private sector. Moving forward, both private and public schools should form the sample population for the NAT. This will enable a more accurate representation of the educational landscape across the country.

As well as including private schools in the sample population, specific questions targeting parents' decisions around sending their children to private schools should be included in the next NAT parent questionnaire to address queries arising from limitations with the 2023 data set. Understanding the reasons why parents send their children to private school, or the reasons that they attend public school, are important for a fuller interpretation of the 2023 NAT results. This could help with the interpretation of the differences in performance of students from rural versus urban schools. It was surprising to find that (with the exception of Grade 8 Maths, where rural performance was better), there was no significant difference in the performance in these two very different contexts.

This pattern holds true at both national and provincial level (with the exception of rural Balochistan where performance was lower). It could be that in urban centres, where access to private schools is easier, more parents who can afford to are choosing to educate their children privately. In rural areas, where access to private schools may be more limited, it may be that parents who could afford to send their children to private school are not able to do so. Given that household income is a well-known correlator with academic performance, the intake of students in rural areas without easy access to private schools may be a more comprehensive cross-section of society. This could explain why these schools performed as well as their urban counterparts (contrary to many other studies). Without further data on the private versus public school landscape and parental wishes in this regard, it is difficult to fully understand these findings. The possible explanations remain a hypothesis rather than a conclusion.

Limitations and recommendations relating to the Foundational Literacy items

The 2023 NAT was the first time that Foundational Literacy items have been included. The decision to include them was taken after the main Test Administrator (TA) training had taken place and the training manuals had been produced. As a result, there was a risk that the administration of these items (as a separate mini test after the main NAT) could be adversely affected. There is some evidence to suggest that there was mal-administration of these items in some schools. This may limit the usefulness of this data in certain provinces. It is recommended that if foundational literacy items are to be administered in the 2025 NAT the instructions on how they should be administered should form part of the main TA training and TA manuals.

Limitations and recommendations relating to the contextual questionnaires

1) Ensuring the content reflects current policy initiatives

It is necessary to review the questionnaires for each NAT cycle to ensure that they are focused on the current issues and collect specific data to inform current or future policy. The educational landscape evolves and so too do the indicators that are useful to policy makers and school leaders. It is therefore critical that the questionnaires are all reviewed before each NAT in light of the current landscape to ensure that they collect the type of data needed. For example, in September 2023, the Foundational Literacy Hub was launched and the reading hour entitlement for all primary pupils was introduced. Given the timing of the 2023 NAT, specific questions relating to this were not included. However, the results could still be used as a baseline to look at impact of this initiative in future years. The 2025 NAT must ensure that questions relating to the implementation of the Foundational Literacy Hub are included so that any links to performance can be analysed. This will helped to determine if the policy is a) being implemented as intended and b) having the desired impact.

2) Triangulation of data from key stakeholder perspectives

Not all questions appeared on each version of the contextual questionnaire. This meant that it was not possible to triangulate the evidence from the different stakeholders. This is an important step when validating results and helps to create a fuller, richer picture. It is acknowledged that not all questions would be appropriate on all questionnaires. However, all questions that are relevant and appropriate to each group of stakeholders should appear on each questionnaire version.

3) Interpretation of open/free response items

The contextual questionnaires contain some open/free responses. Currently, none of these responses have been analysed. Thematic analysis by key topic or area of focus could be used to process these items and conclusions could be drawn. However, this is a time-consuming process and was not possible within the analysis timeframe available. These responses are potentially a rich source of information that should be analysed in due course. If this is not possible due to the time and skill-sets required, inclusion of such items in the 2025 questionnaires must be considered carefully. Assessments and research instruments should not contain questions that cannot or will not be analysed.

4) Specific additional lines of inquiry

a) Educational level of teachers

While the teacher questionnaire asked about teachers' educational level, it did not ask specifically about subject background. The results of the teacher questionnaire shows that there is a positive association between teachers being able to teach the subjects they enjoy and student achievement. This may indicate that the greater performance in the Life Sciences is due to teacher expertise, but it is not certain. A recommendation for the 2025 NAT is to ask specifically about the subject background of teachers. This would enable a better understanding of the relationship between performance in certain subjects/content domains and the impact of teacher knowledge.

b) Course completion

The strong correlation between course completion and student achievement is unsurprising. Unfortunately, it is not possible to make firm recommendations or policy decisions on this topic without further information. It is necessary to understand precisely in which subject and which content domains course completion is an issue, as well as whether this was a whole-school issue as a result of periods of closure. It has not been possible to discern this from the current data set. For example, to understand the performance in Life Sciences compared to the other science content domains, or the Grammar domain in G4 Urdu and Sindhi, it is necessary to confirm whether or not those content areas have been covered and to what degree compared to others in a subject. This will enable more targeted support and decisions to be made. For example, it may be that there is indeed an issue with the teaching of Grammar and potentially costly intervention strategies should be implemented. However, it could simply be that this area of the curriculum had not finished being taught.

c) Level of familiarity with the new National Curriculum

While there has been national outreach regarding the new National Curriculum, it is not possible to discern for certain how successful this has been from the current data set. The lack of teachers' knowledge regarding Statistics and Probability highlights the need to determine whether there are other areas of the 2022 National Curriculum that are either not being taught or are areas for teacher training and development. It is necessary for the 2025 NAT to ensure coverage of SLOs and content areas that are new to the subjects for the 2022 curriculum. In addition, there needs to be targeted questioning of

this in the teacher and headteacher questionnaires[41]. The questionnaires should seek to understand not only whether the topics are being taught but also the level of confidence that teachers feel they have in delivering the material and the amount of training that they have had on its implementation. There should also be targeted questions on the provision of resources for the new curriculum and any new teaching strategies and approaches that are being used in lessons. This will enable more meaningful data to be generated, from which focused policy recommendations can be made.

d) Number of books in the home

The contextual questionnaires collect data on the number of books in the home. However, such data does not necessarily imply that reading to/with the children is occurring in the household. It is not possible to make any conclusions or recommendations on reading in the home through the current data set. Rather than asking simply for the number of books in the home, the questionnaire should additionally focus on the types of books in the home and whether children are reading with or being read to by their caregivers.

e) Nature of academic assessments

Regular academic assessment was positively correlated with achievement in both Grade 4 and Grade 8 students. While all schools should be implementing regular assessments, it is clear from the data that this is not currently the case. However, a limitation of the data set is that it is not possible to determine the nature, frequency and curriculum subject of these assessments. This type of information would enable more concrete conclusions and specific recommendations to be made and should be incorporated into the revised 2025 contextual questionnaires.

Limitations and recommendations relating to engagement with the NAT

There was a large proportion of learners performing at or below the expected guessing rate (25% of the total number of marks). It is not clear from the current data set whether this is a reflection of ability in the subjects or due to learners (and teachers in some instances) simply guessing to finish the assessments in the shortest time possible. It is important that the NAT is seen as low stakes and does not lead to "teaching to the test". However, it is equally important that the significance of the results of this assessment is understood. This will help to ensure that everyone completes the tests to the best of their ability.

With this in mind, it is necessary to run a series of media/social media campaigns prior to the 2025 NAT to raise the level of awareness of and engagement with the NAT. This process can begin with the dissemination of the results of the 2023 NAT. Caution should be used to ensure that the tests are not seen as high stakes. However, they must be viewed as important for the understanding of the national educational landscape. Through careful media handling, the profile of the NAT can be raised with all stakeholders, including parents. This will hopefully help to increase the prestige of the test and may help to reduce the proportion of learners at or below the guessing rate.

[41] Note that it is suggested to question coverage and confidence at broad topic level, not at the level of SLOs. Topic level will give sufficient information to inform next steps without becoming too onerous for the respondents to complete nor for the analysis stage.

In addition to the suggested media campaign, accurate timings for each learner to complete the tests for the 2025 NAT must be recorded and shared as part of the data analysis. If the length of time a learner (or teacher) takes is significantly shorter than expected, this can be taken as a strong indicator that they have not engaged with the test as expected. The results from these learners could be excluded from the final analysis to enable the generation of a more accurate picture and conclusions. A further method for monitoring the degree of guessing would be to include some items that are very easy. A balance must be struck, however, with the usefulness of the information gained versus the impact on the learners for answering additional items. It is essential that such items form part of the piloting phase if this approach is to be used.

Appendix 1: NAT methodology

1. Instrument development

Item development and finalisation is the technical building block on which the entire assessment activity rests. Major inputs into the process of item development are student learning outcomes (SLOs) outlined in the national curriculum.

For the 2023 pilot items the items prepared for the 2021 NAT which did not take place were reviewed. The items were reviewed by subject specialists and assessment experts. This group of experts include NAW team members as well as nominated personnel from PEACs/AEACs, qualified working teachers nominated by provincial departments, along with experts and academics from education universities and Cambridge specialists.

The pilot items are taken in a sample of schools and students that equivalent to the target population for the LSA. Analysis of the results of the pilot items is then used to select the items with optimal psychometric properties for the LSA.

Items are selected based on:

I. Item difficulty analysis
II. Item discrimination analysis
III. Distractor analysis

The ToS for each subject is used to inform the distribution of questions to each content and cognitive domain for all subjects in the NAT.

2. Activities and processes for the NAT

The assessment exercise includes four phases. Each phase includes a set of activities. The four phases and their respective activities are outlined below:

Phase 1. Finalisation of assessment tools and training workshops

- Item validation – the team selects psychometrically optimal items for the assessment using items from the pilot items.
- Test construction - informed by the ToS, which is based on word weightage given in the national curriculum with reference to content and cognitive domains.
- Training of Lead Master Trainers (LMTs) in Islamabad and Test Administrators (TAs) in respective provincial centres.

Phase 2. Successful execution of the large-scale assessment

- Printing of assessment tools (including test booklets, background questionnaires, TA manuals and charts).
- Provision of stationery including pencil, eraser, protractor, compass and ruler to all sample students.
- Test administration (in selected sample schools across the country).

Phase 3. Marking and coding, data entry and analysis

- Development of a marking scheme to ensure consistency in the assessment experts' marking processes.
- Marking test booklets by trained experts from the Federal Board of Intermediate and Secondary Education (FBISE).
- Data entry using FBSIE data entry software for further analysis.
- Data cleaning and data management to ensure a reliable dataset for analysis.
- Statistical and psychometric analysis using the cleaned dataset to develop salient findings and insights.

Phase 4. Preparation of the assessment report including salient findings from the data analysis

- The findings from the data analysis are presented in a coherent manner in the assessment report.
- The report focuses on the following themes:
    - The performance of students in each assessment and how this varied by province, gender and by urban and rural locations.
    - Trends in performance in Grade 4 English and Grade 4 Maths since the 2019 NAT.
    - Comparisons of the performances of students and teachers including examining the relationship between these
    - Exploring variations in student and teacher performance across different cognitive and construct domains to identify any apparent strengths and weaknesses
    - Identifying the factors, both at an institutional level, and those relating to individual students that were most strongly linked with performance across the different tests
- Students' performance with reference to variation in instructional context, students' backgrounds and other factors affecting their achievements.
- Strong and weak areas of students' learning with reference to the curriculum and target competencies.
- Teacher and educational institutions' performance by relating it to the learning achievements of students.
- Actionable feedback on curricula, print materials such as textbooks, lesson plans and teacher manuals, and teacher training.

3. Pilot item analysis

The pilot items were reviewed focusing on the following:

- item facility (the difficulty level)
- item discrimination
- option discrimination
- percentage of students choosing each response option for each item
- percentage of students not reaching the item
- percentage of students omitting the item.

Table 48: Item Facility Ranges with the live NAT papers

Range (% of total)	Difficulty level
≤20	Very difficult
21 - 40	Difficult
41 - 60	Average
61 - 80	Easy
≥81	Very Easy

Table 49: Facility Proportions within the live NAT papers

Easy	Medium	Challenge
35 – 40%	35 – 40%	20 – 25%

Table 50: Item Discrimination Ranges within the live NAT papers

Range	Discrimination
≤0.19	Poor
0.20 – 0.29	Marginal
0.30 – 0.39	Good
≥0.40	Very good

Table 51: Psychometric indicators for items included in the live NAT papers

	Range	Percentage of paper
Facility (difficulty level)	≤ 40% (difficult)	20 – 25%
	41 – 60% (average)	35 – 40%

	Range	Percentage of paper
	61 – 84% (easy)	35 – 40%
Discrimination	≥ 0.2	100%
Distractors	Only the key has a positive point biserial correlation	

Table 52: Length of tests and number of marks in the live NAT papers

Subject	Length of test		Number of marks	
	Grade 4	Grade 8	Grade 4	Grade 8
English reading	90 minutes	n/a	46	n/a
Urdu reading	90 minutes	n/a	52	n/a
Sindhi reading	90 minutes	n/a	52	n/a
Maths	90 minutes	90 minutes	48	52
Science	n/a	90 minutes	n/a	52

4. Framework for Maths

The framework for Maths for Grade 4 and Grade 8 is based on learning outcomes outlined in the 2022 National Curriculum that has been implemented across all of Grades 1 to 8.

This framework specifies the purpose, format, content, and cognitive domains of the Maths tests. It is not designed to be used as a guide for teaching and learning.

NAW designed the framework to facilitate the development of test items and guide subsequent test construction. It includes learning outcomes from the national curricula 2022 for Grade 4 and Grade 8. These will be covered in the test as content domains. The cognitive processes associated with the measurement of Maths are also detailed under cognitive domains. The framework includes ToS from which valid, reliable, and comparable tests can be constructed each year for pilot testing and large-scale assessment across the country.

Cognitive domains for Grade 4 and Grade 8 Maths

For content knowledge to be internalised effectively, and applied to diverse situations, it is important to ensure simultaneous focus on students' cognitive domains and skills. Under the national curriculum, the cognitive domains have been made consistent with the TIMSS assessment framework to ensure greater alignment of Pakistan's learning assessments with international standards. The cognitive domains outlined in the national curriculum are:

Knowing

In this domain, students are expected to have knowledge of words and symbols and understand the basic ideas behind them. It covers the careful use of the concepts, definitions, relations or representation of either.

Applying

In this domain, students should be able to select and apply appropriate mathematical concepts and procedure while solving real-life situations. It covers pure mathematical questions, including numeric or algebraic expressions, equations, geometric figures and statistical data sets.

Reasoning

In this domain, students are required to use their prior knowledge of Maths in new situations. It recognises and formulates a situation by analysing, synthesising and evaluating to solve real-life situations, while considering whether there is sufficient and consistent data.

The NAT papers for 2023 reflect the same cognitive domain proportions as the TIMSS papers for both Grade 4 and Grade 8.

Table 53: Proportion of cognitive domains

Cognitive Domain	Percentages	
	Grade 4	Grade 8
Knowing	40%	35%
Applying	40%	40%
Reasoning	20%	25%

Table of specification for Maths Grade 4

The ToS provides the blueprint for the number of items that should be included in the test and the proportion of the content and cognitive domains across the paper for the LSA.

The content domains are accompanied by cognitive domains which align with those of TIMSS. The content domain percentages broadly reflect the proportion of SLOs in each sub-domain in the curriculum.

Cognitive / **Content**	**Knowing (40%)**	**Applying (40%)**	**Reasoning (20%)**	**Total**
Number and Algebra **(58%)**	11 items	11 items	6 items	**28 items**
Measurement and Geometry **(27%)**	6 items	5 items	2 items	**13 items**

Cognitive / Content	Knowing (40%)	Applying (40%)	Reasoning (20%)	Total
Statistics and Probability **(15%)**	2 items	3 items	2 items	**7 items**
Total	**19 items**	**19 items**	**10 items**	**48 items**

Table of specification for Grade 8 Maths

Cognitive / Content	Knowing (35%)	Applying (40%)	Reasoning (25%)	Total
Numbers and Operations (31%)	4 items	7 items	5 items	16 items
Algebra (31%)	5 items	8 items	3 items	16 items
Geometry and Measurement (19%)	5 items	2 items	3 items	10 items
Statistics and Probability (19%)	4 items	4 items	2 items	10 items
Total	18 items	20 items	14 items	52 items

5. Framework for Science

The framework for Science emphasises knowledge, skills, and competencies needed for a high degree of scientific understanding among students. It is constructed in the form of practical problem-solving tasks that involve design and use of materials while considering the students' developmental level. It is based on the understanding of desirable elements of science education against which student attainment ought to be measured. It covers content domains and learning outcomes for science as

outlined in the national curriculum 2006 for Grade 8. The cognitive dimension specifying the thinking processes to be assessed are also detailed under cognitive domains.

Cognitive domains for Grade 8 Science

For content knowledge to be internalised effectively, and applied to diverse situations, it is important to ensure simultaneous focus on students' cognitive domains and skills. Under the national curriculum, the cognitive domains have been made consistent with the TIMSS assessment framework to ensure greater alignment of Pakistan's learning assessments with international standards. The cognitive domains outlined in the national curriculum are:

Knowing

In this domain, students are expected to have knowledge of concepts and facts and understand the basic ideas behind them. It covers the careful use of the concepts, definitions, relations or representations of either.

Applying

In this domain, students should be able to select and apply appropriate scientific concepts and procedures while solving real-life situations.

Reasoning

In this domain, students are required to use their prior knowledge of science in new situations. It recognises and formulates a situation by analysing, synthesising and evaluating to solve real-life situations.

The NAT papers for 2023 reflect the same cognitive domain proportions as the TIMSS papers for Grade 8 Science.

Table 54: Percentage of cognitive domains

Cognitive Domain	Percentages
	Grade 8
Knowing	40%
Applying	40%
Reasoning	20%

Table of specification for Grade 8 Science

The ToS provides the blueprint for the number of items that should be included in the test and the proportion of the content and cognitive domains across the paper for the LSA.

The content domains are accompanied by cognitive domains which align with TIMSS for cognitive domain. The content domain percentages broadly reflect the proportion of SLOs in each sub-domain in the curriculum.

Cognitive / Content	Knowing (40%)	Applying (40%)	Reasoning (20%)	Total
Life Sciences **(35%)**	7 items	7 items	4 items	**18 items**
Physical Sciences **(50%)**	11 items	10 items	5 items	**26 items**
Earth Sciences **(15%)**	3 items	4 items	1 item	**8 items**
Total	**21 items**	**21 items**	**10 items**	**52 items**

6. Framework for languages

The framework for languages (English, Urdu and Sindhi) focuses on the knowledge, skills and competencies required to understand, respond to and use a range of written texts, and the aspects of grammar and vocabulary that underpin this. It specifies the purpose, format, and content and cognitive domains for the language's tests from which valid, reliable and comparable tests can be constructed each year. It also provides detailed tables of specifications drawing on the relevant SLOs in the National Curriculum.

Currently, the framework for Grade 4 is based on learning outcomes outlined in the 2020 National Curriculum.

The medium of instruction in most schools around the country is Urdu. In some schools, English is also the official medium of instruction. Both are treated as first language subjects under this framework.

Curriculum aims, content and cognitive domains for Grade 4

The 2020 national curriculum outlines relevant objectives for students at the Grade 4 level. The Ministry of Education designed the curriculum to enable students to understand different forms and functions of language and equip them with the skills needed to handle ideas and information. In this way the Grade 4 curriculum provides foundational knowledge and skills for study at subsequent levels. The curriculum is also designed to allow students to develop the logical and critical approaches required to deal with different ideas and abstractions. Finally, the curriculum builds and enhances the students' ability to communicate effectively through a focus on grammar, comprehension skills and vocabulary building.

To support these objectives the national curriculum for Grade 4 is structured around key competencies. Each competency is further categorised into sub-domains that include specific expected learning outcomes. The language competencies outlined in the national curriculum are:

- Oral communication skills (listening and speaking)
- Reading and critical thinking skills
- Formal and lexical aspects of language
- Writing skills
- Appropriate ethical and social development.

The test model for the languages NATs, consisting only of multiple-choice questions to ensure reliability of outcomes and enable the potential for automated marking, means that not all competencies can be assessed in a valid way. Therefore, the NAT framework only includes:

- Reading and critical thinking skills
- Formal and lexical aspects of language
- Reading to perform a task.

Cognitive domains accompany content-based competencies and draw on the levels of Bloom's taxonomy. These are:

- Understanding, including:
    - identifying and retrieving relevant information from texts
    - summarising key details and providing an overview of content
    - recognising the meaning of key vocabulary
    - recognising relevant grammatical and spelling rules.
- Applying, including:
    - using information in documents for a purpose such as predicting future events or interpreting information in texts such as timetables, calendars and charts
    - selecting appropriate words or phrases to complete a sentence grammatically and/or correcting grammar and spelling errors.

Table 55: Cognitive Domains English Reading

Cognitive Domain	Percentages
	Grade 4
Knowledge & Understanding	55%

Application	45%

Table 56: Cognitive Domains Urdu and Sindhi Reading

Cognitive Domain	Percentages
	Grade 4
Knowledge & Understanding	70%
Application	30%

Table 57: Table of Specification for Grade 4 English

Languages Skills/ Competencies	Sub-content areas	Cognitive Skills		
		Understanding 55%	Application 45%	Total
Reading and Critical Thinking Skills 55%	Reading for literary experiences 25%	11 items	2 items	13 items
	Reading for information and Reading to perform a task 15%	7 items	0 items	7 items
	Reading to perform a task 15%	0 items	7 items	7 items
Lexical and formal aspect of language 45%	Vocabulary 19%	2 items	7 items	9 items
	Grammar 26%	6 items	6 items	12 items
Total		26 items	22 items	48 items

Table 58: Table of Specification for Grade 4 Urdu and Sindhi

Languages Skills/ Competencies	Sub-content areas	Cognitive Skills		
		Understanding 70%	**Application 30%**	**Total**
Reading and Critical Thinking Skills 78%	Reading for literary experiences 31%	11 items	5 items	16 items
	Reading for information and Reading to perform a task 31%	11 items	5 items	16 items

	Reading to perform a task 16%	6 items	2 items	8 items
Lexical and formal aspect of language 22%	Vocabulary 11%	4 items	2 items	6 items
	Grammar 11%	4 items	2 items	6 items
Total		36 items	16 items	52 items

Appendix 2: Data collection procedures

Selection and training of field teams

Once sampling had been completed, NAW shared the list of sampled schools with the relevant province and area assessment centres who then nominated test administrators for each school. Test administrators were working teachers up to Grade 16 while lead master trainers were between Grades 17 and 19 with prior experience of conducting test administrator training. A total of 1,450 test administrators and 18 lead master trainers were trained. There were 1,300 sampled schools and each school required one test administrator. An additional 150 (approximately 10%) were trained for contingency. To avoid bias and increase the fairness of test delivery the test administrators were not teachers at the sampled school.

The role of test administrator included taking the assessment tools, administering the achievement tests and returning the materials to NAW-PIE.

Each test administrator received seven hours of training in one of the 18 training centres across the country. This was delivered face-to-face with the use of PowerPoint slides and the Test Administrator Manual. Participants also engaged in role plays of different scenarios, reflecting on the best course of action in various situations and completing random number tables to select students in target grades.

Data management

Test administrators were responsible for returning test materials to NAW-PIE in approved bags. The test was sat on 18 and 19 May 2023. The bags from the sampled schools in Islamabad were delivered by hand by the test administrators on 19 May. In all, 1,299 bags of test materials were received. The final bags arrived on 26 July.

NAW-PIE and the Federal Board of Intermediate and Secondary Education (FBISE) worked together to process the data from the NAT. Standard Operating Procedures were developed to help the two parties collaborate on the data entry for NAT 2023.

NAW-PIE were responsible for designating a data entry coordinator to oversee the data entry process, providing FBISE with the booklets, ensuring data integrity and security during transfer of booklets and to receive and store the booklets once the data entry process was complete. NAW-PIE maintained a record of the booklets sent to FBISE which included the date and time of transfer of materials, while FBISE likewise recorded the booklets they received from NAW-PIE.

FBISE also designated a data entry coordinator. The coordinator oversaw the data entry process and were responsible for the data entry operators entering the data from the NAT booklets into the FBISE database accurately and promptly. Both NAW-PIE and FBISE undertook to work together to resolve any discrepancies or issues.

FBISE developers build the database which was used to collect the data from the NAT booklets. NAW-PIE gradually transferred the NAT booklets to the FBISE in batches. Only when data from one batch was completely entered into the database was the next batch sent for, this mitigated the risk of mixing up data which had been entered with data which had not.

There were 140 data entry operators at the Federal Board who worked on data entry. They worked in two shifts (morning and evening) and NAW experts worked alongside them to resolve technical queries related to the assessment booklets and the background questionnaires. Throughout the data entry

process there was regular monitoring and audits to ensure adherence to the Standard Operating Procedures.

Codebooks, which gave each question a unique ID and identified how to input responses, were produced by Cambridge. FBISE developed the data entry database from these codebooks and NAW experts supported them in this process. The database was continuously improved throughout the data entry process as gaps and errors arose. The data entry operators also used the codebooks to ensure they were inputting the correct code for each question.

The complete and final dataset was sent to Cambridge on 1 September 2023 for analysis.

Appendix 3: Psychometric analysis

1. Methodology

The methodology for the psychometric analyses of the 2023 NAT followed the template provided by AIR's report on the 2019 NAT, which was itself based on guidelines from the Standards for Educational and Psychological Testing (AERA, APA & NCME, 2014[42]). However, in contrast to the previous report:

- Each NAT subject was assessed using a single test version. As well, as meaning there were fewer items to report on in total, this also meant there was no need to perform equating to ascertain equivalencies between scores on different test booklets within the same subject.
- At the time of writing, no policy linking workshops have taken place for the NAT 2023 assessments[43] and so it is not generally possible to report upon the proficiency levels of students.
- Scale scores have also not been calculated at the time of writing. This is because unless combined with the results of policy linking, it would add little extra useful information that could not be seen from analysis of raw scores. Furthermore, the use of scale scores could mask some important features of results – particularly the proportion of students with performances no better than would be expected from random guessing. Note that the way in which scale scores were calculated changed between NAT 2016 and NAT 2019 in any case.

A list of the analyses undertaken for each test is below. These analyses were undertaken using data from students only (i.e., teachers were excluded).

Test reliability statistics. Following the approach used by AIR for the 2019 NAT, Cronbach's alpha was used as the measure of test reliability for each assessment. Test reliability is a measure of internal consistency that estimates the extent to which items measure the same construct. Cronbach's alpha is measured on a scaled of 0 to 1 with values closer to one indicating greater internal consistency. As stated in AIR's report, acceptable values for Cronbach's alpha range are at least 0.80 for achievement tests.

Test Standard Error of Measurement (SEM). In the context of this report, the SEM captures the extent to which students' scores vary due to the exact items that have been included in the test out of the many possible options. In general terms, it estimates the extent to which each student's observed score varies from the average score they would achieve across a large number of test versions of equal difficulty and length to the one they took[44]. The standard error of measurement (SEM) estimates this potential variation in scores using the following formula:

$$SEM = SD\sqrt{1-\alpha}$$

Where SD represents the standard deviation of raw test scores and α represents the value of Cronbach's alpha for the test.

Item-level statistics using classical test theory (CTT). Following AIR's report on the 2019 NAT, the difficulty and point-biserial discrimination of each item has been calculated. In the relevant tables, item difficulty refers to the percentage of students that correctly answered each item on a test. In other places this is sometimes referred to as the *facility* of the item. Following the guidelines from AIR's report,

[42] AERA, APA and NCME (Eds.). (2014). Standards for educational and psychological testing. American Educational Research Association.
[43] These have previously been completed for the NAT 2019 assessments only.
[44] Specifically, the root mean square difference.

the aim is for item difficulties to be between 20% and 80%. Items that were correctly answered by fewer than 20% of students might be considered too hard and items answered correctly by more than 80% of students might be considered too easy. However, this interpretation depends upon the purpose of the test. In particular, since the purpose of the Foundational Literacy test was to identify students that did not have a very basic level of literacy, it is no surprise to see item facility values higher than 80%.

The point-biserial discrimination measures how well an item differentiates between high and low performing students within a test. It is calculated as the Pearson correlation between students' scores on the item and their scores on the rest of the test. As such, point-biserial discriminations can take any value between 0 and 1 with higher values indicating more discrimination. AIR's report on the 2019 NAT suggested that an item's point-biserial discrimination should be higher than 0.20 and, for consistency, this report has continued to report on numbers of items below this threshold.

The percentage of students that omitted or did not reach each item has also been calculated. Note that these percentages are based on the response options in the data exactly as they were provided. In particular, "Omit" has not been replaced with "NR" (that is, not reached) even if it is the last question in the test booklet.

Differential Item Functioning (DIF) by gender. Following the approach used by AIR for the report on the 2019 NAT, the Mantel-Haenszel DIF procedure was used to examine whether each item favoured boys over girls or vice versa. The procedure produces a measure of the statistical significance of DIF in each item using a chi-square statistic. This can be converted into a p-value and items where this is below 0.05 are flagged as having statistically significant DIF. The size of DIF (as opposed to significance) is captured using log-odds ratios. These represent estimates of the extent to which the log of the odds of answering an item correctly for any given overall performance level changes due to being a boy[45]. For items with a statistically significant DIF, log odds ratios greater than 0 mean that boys are advantaged, whereas log odds ratios less than 0 means the girls are advantaged. These were also converted into effect sizes (sometimes labelled "D") typically used in DIF analysis. Based on Zwick (2012)[46], a negligible effect means the absolute size of the Mantel-Haenszel DIF ("D") is less than 1, moderate means the size of the Mantel-Haenszel DIF is greater than 1 but less than 1.5, and large means the size of the Mantel-Haenszel DIF is greater than 1.5.

Note that the DIF analysis was based upon gender as recorded in the assessment data files themselves (as opposed to the student surveys). Within the assessment data files, no students were identified as transgender and so DIF analysis focusses exclusively on the difference between boys and girls.

Item difficulty estimation using IRT. For consistency with AIR's analysis of the 2019 NAT, the difficulty of each item based upon the Rasch model has been estimated. This is mathematically equivalent to a one parameter item response theory (IRT) model. In this report, this model is defined so that:

$$P(Y_{ij} = 1) = \frac{\exp\left(1.7(\theta_i - b_j)\right)}{1 + \exp\left(1.7(\theta_i - b_j)\right)}$$

[45] This is a slightly different definition of log-odds ratios used in AIR's report. In the previous report, these showed the ratio of the log odds value to its sampling variance.

[46] Zwick, R. (2012). A review of ETS differential item functioning assessment procedures: Flagging rules, minimum sample size requirements, and criterion refinement (ETS RR-12-08). Educational Testing Service.

Where $P(Y_{ij} = 1)$ is the probability that of the ith candidate answering the jth item correctly, θ_i is the ability of the ith candidate, and b_j is the difficulty of the jth item. For the purposes of these (unanchored) analyses the ability scale was defined so that mean student ability was equal to 0.

Note that the "1.7" multiplier is included in analysis for consistency with the previous report by AIR. This approach is not universal and so care will be needed if Rasch difficulties in this report are compared with published values for other assessments.

The IRT difficulty parameters differ in their interpretation from those of CTT such that item difficulty indicates the ability level at which examinees are expected to have a 50% chance of answering an item correctly and item discrimination is the rate at which the probability of answering that item correctly increases/decreases relative to an examinee's ability. In the case of the Rasch 1PL IRT model, item discrimination remains constant. Item difficulty parameter values typically range from -3.00 (very easy) to 3.00 (very difficult) and lie on the same scale as the estimates for ability.

Since the Rasch difficulties also may form the basis for test equating, it was felt important also to evaluate the fit of the model. This was done using INFIT and OUTFIT statistics (see Wright and Linacre, 1994)[47]. These fit statistics are calculated by first estimating the IRT ability of each candidate (using the maximum likelihood method) and then determining whether the extent to which the actual item scores differ from expected item scores given ability are as expected under the Rasch model. According to Wright and Linacre (1994) both the INFIT and OUTFIT statistics should be between 0.8 and 1.2 for high-stakes multiple-choice questions.

Means, standard deviations and distributions of raw scores. Cambridge calculated means and standard deviations on the equated raw score scale for each test.

Equating test scores to 2019. Analysis was undertaken to equate scores on the Grade 4 English test to scores in the same subject in the 2019 NAT. Specifically, for every raw score on the 2023 NAT, the aim was to identify an equivalent raw score, representing an equivalent level of achievement in the 2019 NAT. Equating was completed based upon the Rasch model. The precise approach, known as Rasch true score equating, was chosen as being the only viable option given the available data at the time of writing. In order to complete equating, the Rasch difficulties of 5 anchor items that were included in both the 2019 and 2023 NAT were fixed at the values published in AIR's 2019 report[48] and the remainder of the Rasch difficulties were estimated with these values held constant. On the basis of these (revised) Rasch difficulties of all items, the IRT ability associated with each raw score on the 2023 test could be calculated. That is, the ability required so that a student's expected total score on the test was equal to a particular raw score. Finally, the raw scores on the 2019 NAT (specifically, booklet C of the 2019 NAT[49]) associated with the relevant abilities were identified.

Note that equating was only completed for Grade 4 English. It was not possible to complete the same analysis for Grade 8 Maths because:

[47] Wright, B.D., & Linacre, J.M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8:3 p370. https://www.rasch.org/rmt/rmt83b.htm.

[48] Note that this is a separate Rasch model from the unanchored one used simply to evaluate Rasch difficulties escribed earlier.

[49] This booklet was chosen as it was used as the reference booklet in AIR's report.

- Published Rasch difficulties in AIR's report on the 2019 NAT for Grade 8 Maths were inconsistent with the classical difficulties published for the same items in the same report. This made it difficult to be confident in how these values should be interpreted.
- It was not clear how the Rasch difficulties published as part of the report on the 2019 NAT for Grade 8 Maths allowed for equating between booklets. Specifically, whilst Rasch difficulties were only published for the multiple-choice questions, equating tables for the 2019 booklets also seemed to also incorporate scores from constructed-response questions. How this extra step was achieved was not documented.

No other subjects tested as part of the 2023 NAT were also included in the 2019 NAT. As such, they are also not included in the equating analysis.

2. Results

Test reliabilities and standard error of the mean (SEMs)

- The internal consistency reliability for all booklets within each subject test had acceptable values as estimated by Cronbach's alpha. All reliabilities were greater than 0.80, with a range of between 0.88 and 0.94.

The reliability estimates offer a measure of the internal consistency of each subject and booklet, while the SEMs quantify the assessment's measurement error of the student's unknown true score.

The table below indicates that the internal consistency reliability for every subject test and booklet was acceptable, i.e., greater than 0.80 as measured by Cronbach's alpha. The lowest reliability was 0.88 on all the Grade 8 Maths assessment. The highest reliability was for Grade 4 Urdu and Sindhi at 0.94.

The table also shows the SEM of each test. Since the SEM is on the same scale as raw scores, it is unsurprising that this is smallest for the test with the smallest maximum available score (i.e., Grade 4 FL). For the longer tests, all of the SEMs are close to 3. Very roughly this indicates that we would expect students' observed scores to be within 6 score points of their true scores 95% of the time.

Table 59: Test reliabilities and SEMs

Test	N	Max	Mean	SD	Reliability	SEM
Grade 4 English	10591	48	26.91	11.97	0.94	2.96
Grade 4 Maths	10578	48	23.70	10.54	0.92	3.01
Grade 4 Urdu and Sindhi	10662	52	35.43	12.11	0.94	2.86
Grade 4 FL	9354	15	13.24	3.42	0.93	0.91
Grade 8 Maths	11882	52	21.61	9.26	0.88	3.19
Grade 8 Science	11799	52	26.68	10.12	0.90	3.19

Classical test theory

- Based on the item difficulty and item discrimination, the quality of test items was generally acceptable. Across all subjects, only 4 items in total were found to be clearly too hard for the students. Outside of Grade 4 FL, which was specifically designed to be easy for most students, only 6 items were found to be very easy. In terms of discrimination, two items had negative discriminations (both in Grade 4 Maths). A further 11 items (or 5% of the 219 used in total), of which 7 were in Grade 8 Maths, had discriminations that might be considered unacceptable (i.e., positive discrimination, but less than 0.20).

The item-level difficulty and discrimination for each grade and subject was calculated. Note, difficulty simply represents the proportion of candidates answering a question correctly. Thus, higher values of "difficulty" actually indicate that an item was easier. The complete item level statistics are shown later. Following AIR's approach, item difficulty was judged as unacceptable if fewer than 20% of students got the item correct (i.e., too difficult). They were classified as very easy if more than 80% of students responded to the item correctly. Item discrimination was deemed unacceptable if less than 0.20.

Table 60 shows the number of items in categories of item difficulty across each test. As can be seen, very few items were classified as too difficult (i.e., x≤0.2). All items in Grade 4 FL were classified as very easy. This is not surprising given the purpose of this test. Five items from Grade 4 Urdu and Sindhi were also classified as very easy as was 1 item from Grade 4 English.

Table 60:The number of items in each category of difficulty in each test

Test	0≤x<0.2	0.2≤x<0.5	0.5≤x<0.8	x≥0.8	Total
Grade 4 English	0	10	37	1	48
Grade 4 Maths	1	22	25	0	48
Grade 4 Urdu and Sindhi	0	0	47	5	52
Grade 4 FL	0	0	0	15	15
Grade 8 Maths	2	39	11	0	52
Grade 8 Science	1	21	30	0	52

Table 61 summarises the numbers of items in each category of classical discrimination. Two items, in Grade 4 Maths actually had negative discriminations. This may possibly indicate a common misunderstanding in how to handle these particular questions. Eleven other items (mostly in Grade 8 Maths) also had discriminations that might be considered unacceptably low.

Table 61: The number of items in each category of classical discrimination in each test

Test	x<0	0≤x<0.2	0.2≤x<0.5	0.5≤x<0.8	x≥0.8	Total
Grade 4 English	0	1	24	23	0	48
Grade 4 Maths	2	1	33	12	0	48
Grade 4 Urdu and Sindhi	0	0	27	25	0	52
Grade 4 FL	0	0	0	15	0	15
Grade 8 Maths	0	7	43	2	0	52
Grade 8 Science	0	2	50	0	0	52

Differential item functioning

A large proportion of items display statistically significant levels of differential item functioning (DIF) between boys and girls. Indeed, the proportion is noticeably larger than for the 2019 NAT. The increase in the proportion of items with statistically significant DIF is most likely due to the greatly increased sample size since 2019. This allows smaller differences between girls and boys to be detected as statistically significant. The results generally show a balance between items that favour boys and those that favour girls. Looking at the size (as opposed to the significance) of DIF revealed that only 5 items in total displayed DIF of non-negligible size.

Following the approach of AIR, to evaluate the differential item functioning (DIF) for boys and girls, the Mantel-Haenszel test of significance was used. Mantel-Haenszel provides a breakdown of items for each test in terms of the number of items that advantaged boys or girls based on log odds. The significant items for boys or girls are classified into negligible, moderate, and large effects. Full results for individual items are shown later.

Table 62 shows the total number of items in each test and then the number displaying significant ($p<0.05$) DIF between boys and girls split by the size of the DIF. As can be seen, quite a high proportion of items (over half in several tests) displayed significant DIF. However, the size of the DIF was generally classified as negligible. This indicates that the increase in the number of items displaying significant DIF since NAT 2019 is likely due to the larger sample size making the detection of small DIF effects more powerful.

Table 62: The number of items in each test displaying DIF between boys and girls

Subject	Total items	Number of items with significant DIF where size of DIF is…		
		Negligible	Moderate	Large
Grade 4 English	48	27	1	0
Grade 4 Maths	48	32	2	0
Grade 4 Urdu and Sindhi	52	25	1	0
Grade 4 FL	15	2	0	0
Grade 8 Maths	52	31	2	0
Grade 8 Science	52	37	0	0

Table 63 shows the number of items that advantaged boys and girls, where this difference was statistically significant. Items are included in this table regardless of whether the size of DIF was negligible, moderate, or large. In general, the number of items favouring boys was very similar to the number favouring girls.

Table 63: The number of items that advantaged boys and girls

Subject	Total items with significant DIF	Number favouring boys	Number favouring girls
Grade 4 English	28	13	15
Grade 4 Maths	34	15	19
Grade 4 Urdu and Sindhi	26	12	14
Grade 4 FL	2	1	1
Grade 8 Maths	33	20	13
Grade 8 Science	37	17	20

Item response theory

Since only a single test booklet for each subject was used, Rasch difficulty parameters simply provide a different way of looking at item difficulty. They may also be useful for equating the scores from the 2023 NAT to previous or future years. However, this application relies on us being convinced that the Rasch model fits the data.

Table 64 summarises the Rasch difficulties of the items within each test using the same categories as AIR used in their analysis of the 2019 NAT. The results essentially reveal the same patterns described earlier using classical difficulty statistics.

Low Rasch difficulties (e.g. less than -1) indicate that an item is very easy. Nearly all such items occur in Grade 4 FL. Very few hard items are visible in the data (e.g. difficulty greater than 1). The few hard items that are visible occur in Maths (Grade 4 and 8) or in Grade 8 Science.

Table 64: A summary of the Rasch difficulties of the items within each test

Test	$x<-2$	$-2\leq x<-1$	$-1\leq x<0$	$0\leq x<1$	$1\leq x<2$	$x\geq 2$	Total
Grade 4 English	0	1	38	9	0	0	48
Grade 4 Maths	0	0	27	20	1	0	48
Grade 4 Urdu and Sindhi	0	8	44	0	0	0	52
Grade 4 FL	13	2	0	0	0	0	15
Grade 8 Maths	0	0	12	38	2	0	52
Grade 8 Science	0	0	30	21	1	0	52

Table 65 and Table 66 summarise the fit statistics for all items across the various tests. Ideally, the values of the two fit statistics should be between 0.8 and 1.2. Using either the INFIT or the OUTFIT criterion it can be seen that the majority of items in each test have a good fit to the Rasch model. However, by no means all items show an ideal level of fit. As such, it may be worth exploring the use of other equating techniques rather than solely relying on the Rasch model for test equating.

At the time of writing, in the context of equating to the 2019 NAT, there is insufficient information to attempt any other approaches to equating. In order to undertake further investigations, the raw item-level data from the 2019 NAT would be required.

Table 65: The INFIT statistics for all items across each test

Test	INFIT<0.8	0.8≤ INFIT ≤1.2	INFIT>1.2	Total
Grade 4 English	2	41	5	48
Grade 4 Maths	0	44	4	48
Grade 4 Urdu and Sindhi	0	48	4	52
Grade 4 FL	13	2	0	15
Grade 8 Maths	0	52	0	52
Grade 8 Science	0	51	1	52

Table 66: The OUTFIT statistics for all items across each test

Test	OUTFIT<0.8	0.8≤ OUTFIT ≤1.2	OUTFIT>1.2	Total
Grade 4 English	7	35	6	48
Grade 4 Maths	1	39	8	48
Grade 4 Urdu and Sindhi	10	35	7	52
Grade 4 FL	15	0	0	15
Grade 8 Maths	0	46	6	52
Grade 8 Science	0	48	4	52

Equating to the 2019 NAT

As stated earlier, the equating analysis was only performed linking performance in Grade 4 English in the 2023 NAT to performance in the 2019 NAT. This was done based upon Rasch analysis using the common items.

To begin with the Rasch difficulties of every item in the 2023 Grade 4 English test were re-estimated with the difficulties of anchor items (questions 28 to 32) fixed at the values reported for the same items (questions 1, 2, 4, 5 and 7 respectively from 2019 booklet B) in AIR's 2019 report.

The Rasch difficulties of all Grade 4 English items in this analysis are shown in the table below. Note that these difficulties may differ somewhat from the unanchored Rasch difficulties presented later in this report. Anchored items had their difficulties fixed at the values identified in AIR's report on the 2019 NAT.

As a check on the anchoring process, Figure 29 explores the fit of each anchor item to the Rasch model. In particular it helps us check whether the anchored difficulty values fit the performance of students at different levels of ability on each item. To create each plot, the ability (θ) of each student was estimated. Students were then grouped by ability. The points in each plot display the proportion of students getting each item correct ($P(\theta)$) in each group against the mean ability of the group. The blue lines represent the expected relationship between ability and the proportion of students answering an item correctly based upon the Rasch model and the item difficulty reported for the 2019 NAT. If either all points are above the line, or all points are below the line, that would indicate that the anchored difficulty value was not appropriate for these items within the 2023 NAT.

The plot reveals that the Rasch difficulties from the 2019 NAT are broadly appropriate. However, as discussed earlier, the Rasch model itself does not provide a perfect fit to the data (i.e., the blue lines do not perfectly coincide with the points). The impact of this issue on the accuracy of equating could be a matter for further research.

Table 67: Rasch difficulties from anchored analysis of Grade 4 English.

Item	Rasch difficulty	2019 anchor
Q1	0.25	
Q2	-0.52	
Q3	0.19	
Q4	-0.01	
Q5	0.26	
Q6	-0.23	
Q7	-0.05	
Q8	0.07	
Q9	0.06	
Q10	0.40	
Q11	0.30	
Q12	0.27	
Q13	0.63	
Q14	0.09	
Q15	-0.90	
Q16	-0.38	
Q17	-0.29	
Q18	-0.18	
Q19	-0.28	
Q20	0.20	
Q21	0.26	
Q22	0.19	
Q23	0.08	
Q24	-0.18	
Q25	0.32	
Q26	-0.16	
Q27	0.05	
Q28	-0.24	Booklet B Q1
Q29	-0.11	Booklet B Q2
Q30	0.00	Booklet B Q4
Q31	0.13	Booklet B Q5
Q32	0.10	Booklet B Q7
Q33	-0.34	
Q34	-0.35	
Q35	0.15	
Q36	0.16	
Q37	0.13	
Q38	0.02	
Q39	-0.08	
Q40	-0.18	
Q41	-0.09	
Q42	0.22	
Q43	-0.02	
Q44	0.09	
Q45	-0.07	

Item	Rasch difficulty	2019 anchor
Q46	0.14	
Q47	-0.12	
Q48	-0.03	

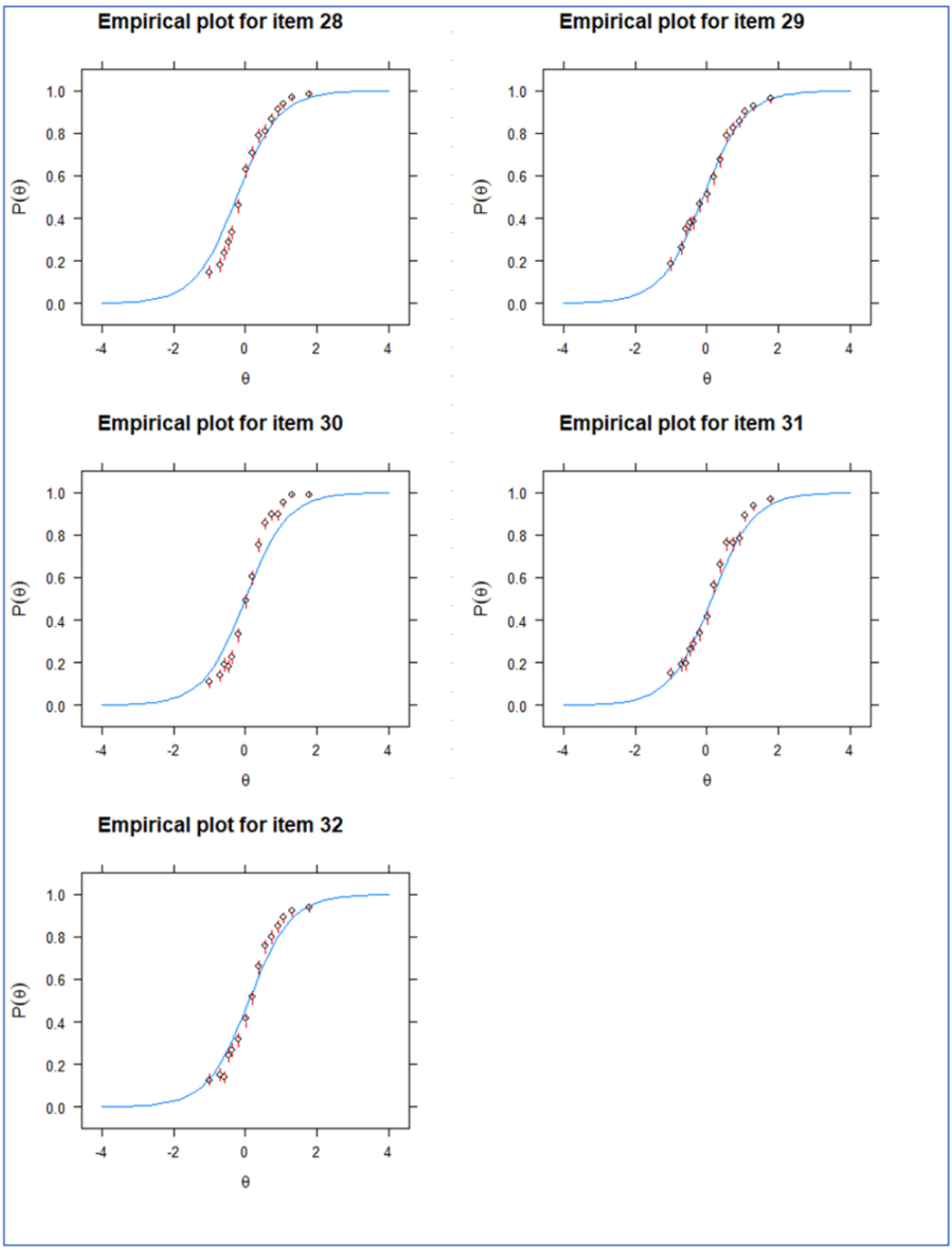


Figure 29: Visual check of item fit to Rasch model for Grade 4 English anchor items.

Finally, using the Rasch difficulties above, the equivalent scores on booklet C of the 2019 NAT (i.e., the reference booklet from 2019[50]) to each whole-number raw score on the 2023 test can be identified. A table of score equivalencies is provided in Table 68. The final column indicates the benchmark scores for global minimum proficiency (GMP) on the 2019 test (page 33 of AIR report). For example, a score of 11 in Grade 4 English in 2023, is equivalent to a score of 10.48 on the 2019 test which, in turn, makes it the first score equivalent a score higher than the benchmark identified for partially meeting (Global Minimum Proficiency) GMP on the 2019 NAT.

Throughout the main report on the 2023 NAT, it is noted that scores representing 25% or less of the maximum test score could easily be achieved by guessing without any knowledge of the subject being tested at all. In the case of Grade 4 English, this means that any scores of 12 or below cannot be taken to indicate any proficiency at all. As such, it is more than a little concerning that a score of 11 on the Grade 4 English test in 2023 equates to a score deemed to partially meet the GMP in the 2019 NAT. Further analysis might explore whether alternative approaches to equating might yield a more acceptable mapping of 2023 scores to GMP benchmarks. Whatever issue effects equating at the lower end of the score scale need not necessarily indicate an issue with equating at the benchmarks for Meets and Exceeds GMP. However, it is difficult to evaluate this without further analysis of raw data from the 2019 NAT.

Table 68: Score equivalencies between Grade 4 English scores on the 2019 and 2023 NAT according to Rasch equating

Score on 2023 Grade 4 English	Equivalent score on 2019 Grade 4 English booklet C	Global minimum proficiency (GMP) cut-off in 2019 booklet C
0	0.00	
1	1.07	
2	2.11	
3	3.12	
4	4.12	
5	5.09	
6	6.03	
7	6.96	
8	7.87	
9	8.75	
10	9.62	
11	10.48	10 (Partially meets GMP)
12	11.31	
13	12.13	
14	12.94	
15	13.74	
16	14.52	
17	15.29	
18	16.05	
19	16.79	
20	17.53	
21	18.26	

[50] This could be used even though the common items came from booklet B since the Rasch analysis of the 2019 NAT placed all Rasch difficulties from all booklets on a common scale.

Score on 2023 Grade 4 English	Equivalent score on 2019 Grade 4 English booklet C	Global minimum proficiency (GMP) cut-off in 2019 booklet C
22	18.98	
23	19.69	
24	20.40	
25	21.09	
26	21.78	
27	22.47	
28	23.14	
29	23.81	
30	24.48	
31	25.14	25 (Meets GMP)
32	25.80	
33	26.45	
34	27.10	
35	27.75	
36	28.39	
37	29.04	
38	29.67	
39	30.31	
40	30.95	
41	31.58	
42	32.21	
43	32.84	
44	33.47	
45	34.10	
46	34.74	
47	35.37	35 (Exceeds GMP)
48	36.00	

Notwithstanding the very serious concerns about equating raised above, Figure 30 shows the percentage of candidates that would be at each proficiency level if these results were relied upon. The results for NAT 2019 are lifted directly from AIR's report. As can be seen, if the equating approach above is trusted, it would indicate a major increase in the proportion of candidates that meet or exceed the GMP.

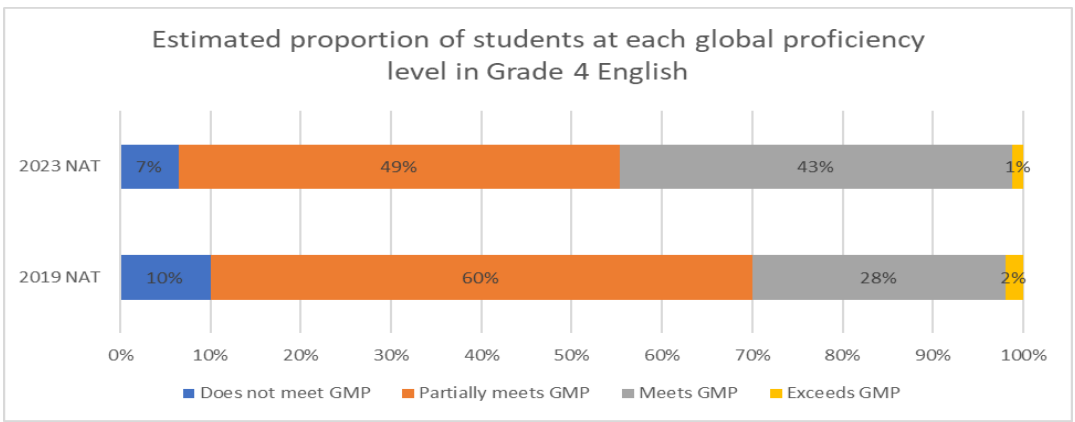


Figure 30: Estimated proportion of students at each global proficiency level according to results of Rasch equating of Grade 4 English test to 2019

The results above reflect the increased proportion of students answering each of the anchor items correctly shown earlier. Descriptive analysis of changes between 2019 and 2023 both on the anchor items (Table 11) and on the tests as a whole (Table 10) were shown within the main report. These showed that, in terms of the proportion of items the students were expected to answer correctly, it is likely that the 2023 Grade 4 English test is of fairly similar difficulty to the ones used in the 2019 NAT (once we average over the 3 NAT 2019 booklets rather than focus purely on booklet C). As such, in general terms, the increased proportion of items answered correctly (shown earlier within the main report) is likely to relate to an improvement in performance.

Although Rasch equating has not been completed for Grade 8 Maths, analysis of performance on the anchor items included in both NAT 2019 and NAT 2023 is also included within the main report.

3. Detailed results of classical item analyses

The tables below show the percentage of students answering each item correctly ("Difficulty") along with the discrimination and the percentage of students omitting or not reaching each item. Where applicable, the final column of each table shows the Student Level Outcome (SLO) targeted by each item.

Table 69: Grade 4 English

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
1	10591	48%	0.328	1%	0%	Scan a simple text for specific information.
2	10591	72%	0.437	1%	0%	Guess meaning of difficult words from context.
3	10591	50%	0.558	1%	0%	Distinguish fact from opinion
4	10591	57%	0.408	1%	0%	Scan a simple text for specific information.
5	10591	48%	0.484	2%	0%	Identify a paragraph as a larger meaningful unit…
6	10591	63%	0.425	1%	0%	Recognise… capitalisation to the initial letter of proper nouns etc
7	10591	58%	0.501	1%	0%	Demonstrate [understanding of] the use of *and, or,* and *but*.
8	10591	54%	0.499	1%	0%	Apply spelling change in plural form…
9	10591	54%	0.455	1%	0%	Use summary skills to identify imp.points
10	10591	44%	0.248	2%	0%	Guess meaning of difficult words from context.
11	10591	47%	0.327	1%	0%	Provide the missing information in a gapped summary.
12	10591	48%	0.419	1%	0%	Recognise elements of a story to: describe the characters in a story
13	10591	37%	0.115	2%	0%	Recognise and use…action verbs…
14	10591	53%	0.384	1%	0%	Classify… nouns as common and proper nouns.
15	10591	81%	0.393	1%	0%	Locate specific information in… a class timetable.
16	10591	68%	0.461	1%	0%	Locate specific information in… a class timetable.

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
17	10591	65%	0.560	1%	0%	Locate specific information in…m , a class timetable.
18	10591	62%	0.380	1%	0%	Locate specific information in… a class timetable.
19	10591	65%	0.547	1%	0%	Locate specific information in… a class timetable.
20	10591	50%	0.448	3%	0%	Locate specific information in… a class timetable.
21	10591	48%	0.323	1%	0%	Locate specific information in… a class timetable.
22	10591	50%	0.512	2%	0%	Locate/ scan specific information to answer short questions.
23	10591	54%	0.418	2%	0%	Recognise elements of a story to: describe the characters in a story
24	10591	62%	0.528	1%	0%	Recognise elements of a story to: describe the characters in a story
25	10591	46%	0.496	2%	0%	Recognise elements of a story to: describe the characters in a story
26	10591	61%	0.508	2%	0%	Recognise elements of a story to: tell where and when a story is set
27	10591	55%	0.578	1%	0%	Locate/ scan specific information to answer short questions.
28	10591	61%	0.581	1%	0%	Scan a simple text for specific information.
29	10591	60%	0.482	1%	0%	Scan a simple text for specific information.
30	10591	57%	0.649	2%	0%	Use summary skills to identify important points
31	10591	54%	0.540	2%	0%	Scan a simple text for specific information.
32	10591	53%	0.568	2%	0%	Scan a simple text for specific information.
33	10591	67%	0.545	2%	0%	Locate, provide and use words similar and opposite in meaning.
34	10591	67%	0.517	2%	0%	Provide missing letters in simple two/three syllable words.
35	10591	51%	0.577	3%	0%	Understand the] use [of] words such as *first*, *second*, *next* and *then* to show a sequence.
36	10591	51%	0.546	3%	0%	Demonstrate the use of subject-verb agreement according to person and number
37	10591	52%	0.637	3%	0%	Locate, provide and use words similar and opposite in meaning.
38	10591	55%	0.583	3%	0%	Locate, identify, differentiate between, and use some simple pairs of words, including homophones
39	10591	59%	0.598	3%	0%	Use some words showing position, time and movement.
40	10591	62%	0.603	2%	0%	Identify countable and uncountable nouns
41	10591	59%	0.565	3%	0%	Recall… rules for the use of *a*, *an,* and *the.*
42	10591	49%	0.367	3%	0%	Identify in text, and change part of speech in a given word.
43	10591	57%	0.557	3%	0%	Identify… simple sentences that show instructions, commands etc.
44	10591	53%	0.465	3%	0%	Classify and change the gender (masculine, feminine, neuter) of… nouns.

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
45	10591	58%	0.462	3%	1%	Provide missing letters in simple two/three syllable words.
46	10591	52%	0.400	3%	1%	Locate, identify, differentiate between, and use some simple pairs of words, including homophones.
47	10591	60%	0.500	2%	1%	Demonstrate [understanding of] the use of and, or, and but.
48	10591	57%	0.463	3%	1%	Distinguish between and use the pronouns as subject, object and for possession.

Table 70: Grade 4 Maths

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
1	10578	79%	0.313	3%	0%	Recognize and identify parallel and non-parallel lines.
2	10578	64%	0.345	2%	0%	Read and write time from digital and analog clocks in 12-hour and 24-hour format.
3	10578	59%	0.451	5%	0%	Identify divisibility rules for 2, 3, 5 and 10 and use them up to 4-digit numbers
4	10578	53%	0.446	2%	0%	Convert units of length from larger to smaller units (Kilometre, metre, centimetre and millimetre).
5	10578	67%	0.437	4%	0%	Solve real-world word problems (including multi step) involving addition and subtraction.
6	10578	64%	0.462	2%	0%	Recognise the place value of each digit in 5-digit numbers.
7	10578	57%	0.369	4%	0%	Describe the outcome of a simple probability experiment (spinner and dice), using mathematical language (i.e. impossible, less likely, more likely, equally likely, unlikely and certain).
8	10578	33%	0.488	3%	0%	Convert units of capacity from larger to smaller units (litre and millilitre).
9	10578	62%	0.420	4%	0%	Compare two numbers up to 5 - digit numbers using symbols “<”, “>”, or “=”
10	10578	31%	0.384	3%	0%	Draw, read and interpret horizontal and vertical single and double bar graphs (including real life problems).
11	10578	55%	0.448	4%	0%	Express decimal numbers up to three decimal places as fractions.
12	10578	50%	0.468	4%	0%	Recognize and identify acute, right, and obtuse angles.
13	10578	55%	0.412	4%	0%	Multiply and divide proper, improper fractions and mixed numbers by a whole number.
14	10578	47%	0.503	4%	0%	Reinforce/recall round off numbers to the nearest tens, hundreds, thousands.
15	10578	52%	0.429	4%	0%	Identify and write expressions or number sentences to represent problems that may involve unknowns.

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
16	10578	62%	0.528	4%	0%	Identify and differentiate between multiples and factors and find: all factor pairs of a number common factors of two numbers
17	10578	50%	0.486	4%	0%	Identify and differentiate between 2-digit prime and composite numbers up to 50.
18	10578	52%	0.505	3%	0%	Convert units of mass from larger to smaller units (kilogram and gram).
19	10578	57%	0.455	4%	0%	Complete the given increasing and decreasing number patterns.
20	10578	64%	0.505	3%	0%	Solve real-world word problems (including multi step) involving addition and subtraction.
21	10578	69%	0.471	4%	0%	Add and subtract up to 5-digit numbers mentally and in written form (with and without regrouping)
22	10578	38%	0.494	4%	0%	Differentiate among: proper fractions improper fractions mixed numbers.
23	10578	47%	0.510	6%	0%	Multiply and divide proper, improper fractions and mixed numbers by a whole number.
24	10578	42%	0.413	3%	0%	Draw, read and interpret horizontal and vertical single and double bar graphs (including real life problems).
25	10578	38%	0.522	4%	0%	Solve real-world word problems involving multiplication.
26	10578	55%	0.425	3%	0%	Describe the radius, diameter, and circumference of a circle.
27	10578	20%	-0.035	4%	0%	Reinforce/recall round off numbers to the nearest tens, hundreds, thousands.
28	10578	57%	0.532	4%	0%	Multiply and divide a 2-digit number with one decimal place by a 1-digit number or a 2-digit number.
29	10578	43%	0.350	4%	0%	Calculate duration of different events using start time and end time.
30	10578	46%	0.251	4%	0%	Complete the given increasing and decreasing number patterns.
31	10578	61%	0.464	4%	0%	Identify and differentiate between multiples and factors and find: all factor pairs of a number common factors of two numbers
32	10578	24%	0.132	3%	0%	Recognise, read, write decimal numbers and identify the place value of decimal numbers with up to three decimal places.
33	10578	24%	0.204	5%	0%	Convert, add and subtract mass to solve real-world word problems.
34	10578	43%	0.483	5%	0%	Multiply and divide a 2-digit number with one decimal place by 10 and 100.
35	10578	33%	0.250	4%	0%	Recognize and identify acute, right, and obtuse angles.
36	10578	16%	-0.035	5%	0%	Describe the outcome of a simple probability experiment (spinner and dice), using mathematical language (i.e.

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
						impossible, less likely, more likely, equally likely, unlikely and certain).
37	10578	55%	0.522	5%	0%	Differentiate among: proper fractions improper fractions mixed numbers.
38	10578	44%	0.269	6%	0%	Apply formulas to find the perimeter and area of squares, rectangles, and rectilinear shapes.
39	10578	54%	0.508	4%	0%	Convert larger units to smaller units of time (hours, minutes, seconds, years, months, weeks and days).
40	10578	53%	0.315	4%	0%	Compare and order unlike fractions.
41	10578	48%	0.463	4%	0%	Recognise the place value of each digit in 5-digit numbers.
42	10578	53%	0.560	4%	1%	Draw, read and interpret horizontal and vertical single and double bar graphs (including real life problems).
43	10578	60%	0.481	4%	1%	Solve real-world word problems (including multi step) involving addition and subtraction.
44	10578	47%	0.458	4%	1%	Identify and differentiate between multiples and factors and find: all factor pairs of a number common factors of two numbers
45	10578	39%	0.392	5%	1%	Recognize and draw lines of symmetry in 2-D shapes and complete symmetrical figures with respect to a given line of symmetry.
46	10578	41%	0.528	4%	1%	Draw, read and interpret horizontal and vertical single and double bar graphs (including real life problems).
47	10578	47%	0.497	6%	1%	Express decimal numbers up to three decimal places as fractions.
48	10578	60%	0.539	5%	1%	Draw, read and interpret horizontal and vertical single and double bar graphs (including real life problems).

Table 71: Grade 4 Urdu and Sindhi

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
1	10662	80%	0.455	1%	0%	Locate/ scan specific information to answer short questions.
2	10662	77%	0.508	1%	0%	Locate/ scan specific information to answer short questions.
3	10662	81%	0.504	1%	0%	Identify facts in the text (as indicated through these words; day, date, place, etc.).
4	10662	67%	0.487	1%	0%	Identify facts in the text (as indicated through these words; day, date, place, etc.).

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
5	10662	79%	0.488	1%	0%	Predict what follows in the text using context and prior knowledge.
6	10662	75%	0.361	1%	0%	Predict what follows in the text using context and prior knowledge.
7	10662	52%	0.383	2%	0%	Guess meaning of difficult words from context.
8	10662	59%	0.361	1%	0%	Locate/ scan specific information to answer short questions.
9	10662	86%	0.412	2%	0%	Locate/ scan specific information to answer short questions.
10	10662	62%	0.505	1%	0%	Scan a simple text for specific information.
11	10662	75%	0.562	1%	0%	Locate/ scan specific information to answer short questions.
12	10662	72%	0.523	1%	0%	Scan a simple text for specific information.
13	10662	65%	0.480	1%	0%	Tenses
14	10662	82%	0.469	1%	0%	Locate/ scan specific information to answer short questions.
15	10662	72%	0.460	1%	0%	Scan a simple text for specific information.
16	10662	67%	0.472	1%	0%	Predict what follows in the text using context and prior knowledge.
17	10662	58%	0.355	2%	0%	Locate, provide and use words similar and opposite in meaning.
18	10662	84%	0.484	1%	0%	Identify facts in the text (as indicated through these words; day, date, place, etc.).
19	10662	53%	0.430	1%	0%	Locate, provide and use words similar and opposite in meaning.
20	10662	58%	0.603	2%	0%	Locate, identify, differentiate between, and use some simple pairs of words, including homophones.
21	10662	81%	0.514	1%	0%	Locate/ scan specific information to answer short questions.
22	10662	71%	0.601	1%	0%	Identify facts in the text (as indicated through these words; day, date, place, etc.).
23	10662	68%	0.572	1%	0%	Identify facts in the text (as indicated through these words; day, date, place, etc.).
24	10662	66%	0.539	1%	0%	Use summary skills to identify important points
25	10662	66%	0.536	1%	0%	Use summary skills to identify important points
26	10662	73%	0.543	1%	0%	Locate, provide and use words similar and opposite in meaning.
27	10662	53%	0.444	2%	0%	Read tables and charts in textbooks.
28	10662	51%	0.496	2%	0%	Locate, provide and use words similar and opposite in meaning.
29	10662	76%	0.526	1%	0%	Locate/ scan specific information to answer short questions.
30	10662	78%	0.467	1%	0%	Use summary skills to identify important points

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
31	10662	69%	0.599	1%	0%	Use summary skills to identify important points
32	10662	60%	0.512	1%	0%	Classify… nouns as common and proper nouns.
33	10662	54%	0.523	1%	0%	Apply spelling change in plural form of regular and irregular nouns and regular verb forms.
34	10662	69%	0.499	1%	0%	Guess meaning of difficult words from context.
35	10662	74%	0.539	1%	0%	Identify facts in the text (as indicated through these words; day, date, place, etc.).
36	10662	70%	0.516	1%	0%	Locate/ scan specific information to answer short questions.
37	10662	62%	0.523	1%	0%	Use summary skills to identify important points
38	10662	59%	0.235	1%	0%	Recognise meaning of common adjectives in relation to each other e.g. huge versus big.
39	10662	57%	0.490	2%	0%	Recognise meaning of common adjectives in relation to each other e.g. huge versus big.
40	10662	75%	0.593	1%	0%	Classify… nouns as common and proper nouns.
41	10662	75%	0.492	1%	0%	Locate/ scan specific information to answer short questions.
42	10662	73%	0.567	1%	0%	Predict what follows in the text using context and prior knowledge.
43	10662	71%	0.359	1%	0%	Predict what follows in the text using context and prior knowledge.
44	10662	73%	0.444	2%	0%	Guess meaning of difficult words from context.
45	10662	70%	0.503	1%	0%	Tenses
46	10662	59%	0.412	1%	0%	Apply spelling change in plural form of regular and irregular nouns and regular verb forms.
47	10662	77%	0.540	2%	0%	Use textual aids such as table of content and glossary for greater comprehension of texts.
48	10662	64%	0.520	2%	0%	Use textual aids such as table of content and glossary for greater comprehension of texts.
49	10662	72%	0.494	2%	0%	Use textual aids such as table of content and glossary for greater comprehension of texts.
50	10662	68%	0.545	2%	0%	Use textual aids such as table of content and glossary for greater comprehension of texts.
51	10662	55%	0.331	2%	1%	Use textual aids such as table of content and glossary for greater comprehension of texts.
52	10662	50%	0.408	2%	1%	Read tables and charts in textbooks.

Table 72: Grade 4 Foundational Literacy

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached
1	9354	90%	0.571	1%	0%
2	9354	89%	0.730	2%	0%
3	9354	90%	0.680	2%	0%
4	9354	90%	0.693	2%	0%
5	9354	85%	0.553	2%	0%
6	9354	87%	0.659	2%	0%
7	9354	83%	0.704	3%	0%
8	9354	90%	0.699	2%	0%
9	9354	89%	0.733	2%	0%
10	9354	89%	0.550	2%	0%
11	9354	89%	0.714	1%	0%
12	9354	87%	0.623	2%	0%
13	9354	88%	0.629	2%	0%
14	9354	90%	0.696	2%	0%
15	9354	89%	0.685	2%	0%

Table 73: Grade 8 Maths

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
1	11882	73%	0.182	1%	0%	Construct different types of quadrilaterals (square, rectangle, parallelogram, trapezium, rhombus and kite).
2	11882	74%	0.309	1%	0%	Calculate direct and inverse and compound proportion and solve real-world word problems related to direct, inverse and compound proportion.
3	11882	59%	0.351	1%	0%	Solve real world word problems involving profit %, loss %, discount, profit, mark-up, insurance, partnership and inheritance.
4	11882	71%	0.331	2%	0%	Convert Pakistani currency to well-known international currencies and vice versa.
5	11882	70%	0.394	1%	0%	Describe sets using language (tabular, descriptive and set-builder notation) and Venn diagrams
6	11882	50%	0.294	5%	0%	Solve real life problems involving number sequences and patterns.

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
7	11882	18%	0.284	1%	0%	Calculate range, variance and standard deviation for ungrouped data and solve related real-world problems.
8	11882	45%	0.109	1%	0%	Identify and differentiate between decimal numbers as terminating (non-recurring) and non-terminating (recurring).
9	11882	52%	0.307	3%	0%	Enlarge a figure (with the given scale factor) and find the centre and scale factor of enlargement.
10	11882	58%	0.283	1%	0%	Differentiate between rational and irrational numbers.
11	11882	47%	0.278	2%	0%	Demonstrate the following properties: closure property associative property existence of identify element existence of inverses commutative property distributive property
12	11882	56%	0.321	1%	0%	Apply sets in real-life word problems.
13	11882	68%	0.381	1%	0%	Describe sets using language (tabular, descriptive and set-builder notation) and Venn diagrams
14	11882	43%	0.407	2%	0%	Solve real-world word problems involving two simultaneous linear equations in two variables.
15	11882	40%	0.264	4%	0%	Perform probability experiments (for example tossing a coin, rolling a die, spinning a spinner etc. for certain number of times) to estimate probability of a simple event
16	11882	54%	0.368	1%	0%	Calculate range, variance and standard deviation for ungrouped data and solve related real-world problems.
17	11882	58%	0.370	2%	0%	State the Pythagoras theorem and use it to solve right angled triangles.
18	11882	30%	0.139	3%	0%	Calculate range, variance and standard deviation for ungrouped data and solve related real-world problems.
19	11882	47%	0.411	2%	0%	Round off numbers up to 5 significant figures.
20	11882	42%	0.272	2%	0%	Recognise the following algebraic identities and use them to expand expressions:

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
						(a + b)2 = a2 +2ab + b2 (a - b)2 = a2 - 2ab + b2 a2 - b2 = (a - b)(a + b)
21	11882	40%	0.254	3%	0%	Recall the multiplication of polynomials.
22	11882	38%	0.300	2%	0%	Solve real life problems involving number sequences and patterns.
23	11882	44%	0.419	2%	0%	Find terms of an arithmetic sequence using: term to term rule position to term rule
24	11882	42%	0.389	2%	0%	Represent real numbers on a number line and Recognise the absolute value of a real number.
25	11882	32%	0.157	3%	0%	Find the square root of natural numbers, common fractions and decimal numbers (up to 6-digit).
26	11882	39%	0.333	2%	0%	Simplify algebraic expressions involving addition, subtraction, multiplication and division.
27	11882	37%	0.408	3%	0%	Solve real life word problems using Pythagoras theorem.
28	11882	39%	0.441	3%	0%	Calculate the surface area and volume of the pyramid, sphere, hemisphere and cone.
29	11882	41%	0.156	2%	0%	Explain and calculate profit percentage, loss percentage and discount.
30	11882	32%	0.238	3%	0%	Solve real-world word problems involving two simultaneous linear equations in two variables.
31	11882	45%	0.403	2%	0%	Perform probability experiments (for example tossing a coin, rolling a die, spinning a spinner etc. for certain number of times) to estimate probability of a simple event
32	11882	39%	0.450	2%	0%	Factorise expressions of the following types: ka + kb + kc ac + ad + bck + bd a2 +- 2ab + b2 a2 – b2 a2 +- 2ab + b2 – c2
33	11882	30%	0.364	3%	0%	State the Pythagoras theorem and use it to solve right angled triangles.
34	11882	42%	0.379	2%	0%	Construct different types of quadrilaterals (square, rectangle, parallelogram,

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
						trapezium, rhombus and kite).
35	11882	43%	0.413	2%	0%	Construct different types of quadrilaterals (square, rectangle, parallelogram, trapezium, rhombus and kite).
36	11882	32%	0.469	2%	0%	Solve real-world word problems involving cubes and cube roots.
37	11882	44%	0.461	2%	0%	Deduce and apply the following laws of Exponents/ Indices: Product Law Quotient Law Power Law
38	11882	37%	0.393	2%	0%	Solve simple linear inequalities
39	11882	31%	0.384	3%	0%	Explain and calculate profit/mark-up, principal amount and mark-up rate.
40	11882	41%	0.478	2%	0%	Deduce and apply the following laws of Exponents/ Indices: Product Law Quotient Law Power Law
41	11882	35%	0.154	3%	0%	Calculate range, variance and standard deviation for ungrouped data and solve related real-world problems.
42	11882	30%	0.339	3%	0%	Find terms of an arithmetic sequence using: term to term rule position to term rule
43	11882	36%	0.427	3%	0%	State the Pythagoras theorem and use it to solve right angled triangles.
44	11882	26%	0.221	3%	0%	Solve real-world word problems involving approximation
45	11882	23%	0.231	2%	0%	Construct frequency distribution tables, histograms (of equal widths) and frequency polygons and solve related real-world problems.
46	11882	26%	0.532	3%	0%	Solve simple linear inequalities
47	11882	31%	0.420	2%	0%	Perform probability experiments (for example tossing a coin, rolling a die, spinning a spinner etc. for certain number of times) to estimate probability of a simple event

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
48	11882	22%	0.265	4%	0%	Perform probability experiments (for example tossing a coin, rolling a die, spinning a spinner etc. for certain number of times) to estimate probability of a simple event
49	11882	30%	0.296	3%	0%	Compare experimental and theoretical probability in simple events.
50	11882	26%	0.354	2%	0%	Explain and calculate profit percentage, loss percentage and discount.
51	11882	34%	0.501	3%	0%	Solve simultaneous linear equations in two variables using: elimination method substitution method graphical method division and factorisation method
52	11882	17%	0.118	3%	1%	Rotate an object and find the centre of rotation by rotation.

Table 74: Grade 8 Science

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
1	11799	76%	0.262	1%	0%	Describe the structure and functions of the nervous system
2	11799	79%	0.376	1%	0%	Identify the characteristics that can be transferred from parents to off springs.
3	11799	62%	0.412	2%	0%	Explain that how do astronauts survive and research in space
4	11799	17%	0.356	1%	0%	Interconvert smaller units and bigger units
5	11799	79%	0.343	1%	0%	Draw and label human excretory system
6	11799	62%	0.289	1%	0%	Interconvert smaller units and bigger units
7	11799	31%	0.385	2%	0%	Describe development of tools and technologies used in space exploration
8	11799	68%	0.333	1%	0%	Investigate the means used by scientist and engineers to overcome the problems of expansion and

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
						contraction in everyday life
9	11799	59%	0.461	1%	0%	Describe the role of kidney in excretion of waste
10	11799	34%	0.236	1%	0%	Explain the balancing of a chemical reaction
11	11799	63%	0.429	1%	0%	Describe the image formation using a lens by ray diagram
12	11799	24%	0.201	2%	0%	Identify new technologies used on earth that have developed as a result of the development of space technology
13	11799	73%	0.387	2%	0%	Interpret SI units in the daily life
14	11799	58%	0.429	2%	0%	Describe the causes and effects of ozone depletion
15	11799	67%	0.286	1%	0%	Describe the term atmospheric pressure
16	11799	64%	0.418	1%	0%	List some biotechnological products used in daily life.
17	11799	67%	0.356	1%	0%	Select and use measuring instruments
18	11799	48%	0.432	2%	0%	Define the terms acid, alkali and salt
19	11799	28%	0.242	1%	0%	Define the term Pressure
20	11799	66%	0.427	1%	0%	List some biotechnological products used in daily life.
21	11799	52%	0.495	2%	0%	Carry out research to explain global warming and its likely effects on life on earth.
22	11799	34%	0.258	2%	0%	Describe the term atmospheric pressure
23	11799	48%	0.331	1%	0%	Explain the types of chemical reactions with examples
24	11799	43%	0.194	2%	0%	Suggest the ways to solve the problems that have resulted from space exploration
25	11799	53%	0.306	1%	0%	Describe the structure and functions of the nervous system
26	11799	45%	0.194	1%	0%	Explain that how do astronauts survive and research in space

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
27	11799	53%	0.467	1%	0%	Compare and contrast the working of a human eye with the lens camera
28	11799	61%	0.319	1%	0%	Differentiate between mitosis and meiosis
29	11799	32%	0.341	2%	0%	Describe basic component of an electronic system
30	11799	35%	0.381	2%	0%	Explain the types of chemical reactions with examples
31	11799	50%	0.372	1%	0%	Explain the Greenhouse effect
32	11799	62%	0.349	1%	0%	Design a spacecraft and explain the key features of design to show its suitability as a spacecraft
33	11799	34%	0.448	1%	0%	Describe basic component of an electronic system
34	11799	54%	0.476	1%	0%	Use indicators to identify acids, alkalis and neutral substances
35	11799	37%	0.320	1%	0%	Explain that how do astronauts survive and research in space
36	11799	51%	0.340	1%	0%	Describe the uses of expansion and contraction of liquids
37	11799	67%	0.440	1%	0%	Explain the working of the model generator
38	11799	45%	0.392	1%	0%	Identify the characteristics that can be transferred from parents to off springs.
39	11799	39%	0.318	1%	0%	Explain the uses of acid, alkali and salt in daily life
40	11799	41%	0.450	1%	0%	Suggest techniques to cure problems of kidneys
41	11799	48%	0.382	1%	0%	Identify the technological tools used in space exploration
42	11799	54%	0.409	1%	0%	Identify the simple devices that generate electricity in daily life
43	11799	43%	0.398	1%	0%	Define the term Pressure
44	11799	48%	0.446	1%	0%	Identify DNA and chromosomes in the cell diagram

Item	Number of students	Difficulty	Point-biserial discrimination	Omitted	Not reached	SLO
45	11799	61%	0.448	1%	0%	Define chemical reactions and give examples
46	11799	56%	0.291	1%	0%	Define chemical reactions and give examples
47	11799	42%	0.396	1%	0%	Differentiate between mitoses and meiosis
48	11799	51%	0.442	1%	0%	Explain the sources, properties and harmful effects of air pollutants
49	11799	59%	0.375	1%	0%	Differentiate between voluntary and involuntary actions they have experienced
50	11799	55%	0.425	1%	0%	Explain the energy changes in chemical reactions
51	11799	61%	0.382	1%	0%	List general applications of biotechnology in various fields.
52	11799	30%	0.363	1%	0%	Describe the image formation using a lens by ray diagram

4. DIF by gender

The Mantel-Haenszel (MH) Chi-square test of significance was used to determine whether each item favoured boys over girls. The test produces a chi-square statistic (column "MH Chi Square"), p-value (column "P-Value"), and log odds ratio (column "Log Odds Ratio") for each item. The log odds ratio column represents estimates of the extent to which the log of the odds of answering each item correctly for any given overall performance level changes due to being a boy[51]. Log odds ratios greater than 0 indicate that boys are advantaged, while log odds ratios less than 0 indicate that girls are advantaged. More specifically, the ETS Delta Scale was applied to the results, which produces the calculated effect size according to the scale (column "MH D-DIF") and an interpretation (column "Effect Size"). If the calculated effect size is less than 1 it is interpreted as "Negligible," if the calculated effect size is greater than 1 but less than 1.5 it is interpreted as "Moderate," and if the calculated effect size is greater than 1.5 it is interpreted as "Large."

Table 75: Grade 4 English

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q1	11.71	0.00	-0.15	0.34	Negligible
Q2	2.84	0.09	-0.08	0.20	Negligible

[51] This is a slightly different definition of log-odds ratios used in AIR's report. In the previous report, these showed the ratio of the log odds value to its sampling variance.

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q3	47.57	0.00	0.34	-0.80	Negligible
Q4	2.54	0.11	0.07	-0.17	Negligible
Q5	10.67	0.00	-0.15	0.35	Negligible
Q6	2.72	0.10	-0.08	0.18	Negligible
Q7	7.07	0.01	0.13	-0.30	Negligible
Q8	48.83	0.00	0.33	-0.77	Negligible
Q9	42.47	0.00	-0.29	0.69	Negligible
Q10	42.61	0.00	-0.27	0.63	Negligible
Q11	31.44	0.00	-0.24	0.56	Negligible
Q12	0.08	0.77	-0.01	0.03	Negligible
Q13	8.71	0.00	0.12	-0.29	Negligible
Q14	0.84	0.36	0.04	-0.10	Negligible
Q15	4.93	0.03	-0.13	0.30	Negligible
Q16	19.19	0.00	0.21	-0.50	Negligible
Q17	0.58	0.45	0.04	-0.09	Negligible
Q18	1.37	0.24	0.05	-0.12	Negligible
Q19	1.08	0.30	0.05	-0.13	Negligible
Q20	17.30	0.00	0.19	-0.44	Negligible
Q21	108.94	0.00	0.44	-1.04	Moderate
Q22	2.30	0.13	-0.07	0.17	Negligible
Q23	0.04	0.85	-0.01	0.02	Negligible
Q24	2.52	0.11	0.08	-0.19	Negligible
Q25	4.93	0.03	-0.10	0.25	Negligible
Q26	6.20	0.01	0.12	-0.28	Negligible
Q27	32.56	0.00	-0.28	0.66	Negligible
Q28	18.11	0.00	-0.22	0.51	Negligible
Q29	46.32	0.00	-0.32	0.75	Negligible
Q30	0.29	0.59	-0.03	0.07	Negligible
Q31	0.01	0.91	-0.01	0.02	Negligible
Q32	0.90	0.34	0.05	-0.11	Negligible
Q33	3.56	0.06	0.10	-0.23	Negligible
Q34	0.46	0.50	-0.04	0.08	Negligible
Q35	13.43	0.00	-0.18	0.43	Negligible
Q36	11.98	0.00	0.17	-0.40	Negligible
Q37	7.80	0.01	0.15	-0.35	Negligible
Q38	0.05	0.82	0.01	-0.03	Negligible
Q39	1.20	0.27	0.06	-0.13	Negligible
Q40	4.67	0.03	-0.11	0.27	Negligible
Q41	15.00	0.00	0.19	-0.46	Negligible

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q42	1.40	0.24	0.05	-0.12	Negligible
Q43	30.24	0.00	-0.27	0.63	Negligible
Q44	27.34	0.00	0.24	-0.56	Negligible
Q45	14.38	0.00	0.18	-0.41	Negligible
Q46	38.21	0.00	-0.27	0.64	Negligible
Q47	0.81	0.37	0.04	-0.10	Negligible
Q48	7.43	0.01	-0.13	0.29	Negligible

Table 76: Grade 4 Maths

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q1	48.20	0.00	-0.37	0.86	Negligible
Q2	10.81	0.00	0.15	-0.34	Negligible
Q3	14.48	0.00	-0.17	0.41	Negligible
Q4	61.48	0.00	0.35	-0.83	Negligible
Q5	0.61	0.44	-0.04	0.09	Negligible
Q6	4.13	0.04	-0.10	0.23	Negligible
Q7	7.30	0.01	0.12	-0.28	Negligible
Q8	70.49	0.00	0.41	-0.97	Negligible
Q9	9.99	0.00	-0.14	0.34	Negligible
Q10	1.13	0.29	0.05	-0.12	Negligible
Q11	15.15	0.00	-0.18	0.41	Negligible
Q12	100.63	0.00	-0.46	1.08	Moderate
Q13	2.79	0.09	0.07	-0.18	Negligible
Q14	6.88	0.01	-0.12	0.29	Negligible
Q15	1.67	0.20	0.06	-0.14	Negligible
Q16	1.07	0.30	-0.05	0.12	Negligible
Q17	9.18	0.00	-0.14	0.33	Negligible
Q18	43.98	0.00	0.31	-0.73	Negligible
Q19	1.30	0.25	-0.05	0.12	Negligible
Q20	14.97	0.00	0.19	-0.45	Negligible
Q21	2.93	0.09	-0.09	0.20	Negligible
Q22	19.52	0.00	-0.21	0.50	Negligible
Q23	1.58	0.21	0.06	-0.14	Negligible
Q24	41.33	0.00	0.28	-0.67	Negligible
Q25	36.36	0.00	-0.29	0.69	Negligible
Q26	24.75	0.00	-0.22	0.52	Negligible
Q27	4.45	0.03	0.11	-0.25	Negligible
Q28	8.83	0.00	0.14	-0.34	Negligible

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q29	23.29	0.00	0.21	-0.49	Negligible
Q30	1.03	0.31	-0.04	0.10	Negligible
Q31	5.08	0.02	-0.11	0.25	Negligible
Q32	17.93	0.00	0.20	-0.47	Negligible
Q33	0.95	0.33	-0.05	0.11	Negligible
Q34	25.68	0.00	0.24	-0.56	Negligible
Q35	4.55	0.03	-0.09	0.22	Negligible
Q36	28.82	0.00	-0.29	0.68	Negligible
Q37	41.41	0.00	-0.30	0.72	Negligible
Q38	6.37	0.01	-0.11	0.25	Negligible
Q39	182.05	0.00	0.64	-1.51	Large
Q40	0.24	0.63	0.02	-0.05	Negligible
Q41	29.14	0.00	0.25	-0.58	Negligible
Q42	14.10	0.00	0.18	-0.43	Negligible
Q43	0.86	0.35	-0.04	0.10	Negligible
Q44	0.45	0.50	-0.03	0.07	Negligible
Q45	12.73	0.00	-0.16	0.37	Negligible
Q46	0.81	0.37	0.04	-0.10	Negligible
Q47	25.13	0.00	-0.23	0.55	Negligible
Q48	5.02	0.03	-0.11	0.26	Negligible

Table 77: Grade 4 Urdu and Sindhi

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q1	14.10	0.00	-0.21	0.50	Negligible
Q2	23.59	0.00	-0.28	0.65	Negligible
Q3	18.81	0.00	-0.26	0.62	Negligible
Q4	52.40	0.00	-0.35	0.83	Negligible
Q5	4.53	0.03	-0.12	0.29	Negligible
Q6	1.14	0.29	-0.05	0.13	Negligible
Q7	34.86	0.00	0.26	-0.62	Negligible
Q8	2.93	0.09	-0.08	0.18	Negligible
Q9	0.28	0.60	-0.04	0.09	Negligible
Q10	0.28	0.60	0.03	-0.06	Negligible
Q11	15.29	0.00	-0.23	0.53	Negligible
Q12	2.21	0.14	-0.08	0.19	Negligible
Q13	5.29	0.02	0.11	-0.26	Negligible
Q14	0.13	0.72	0.02	-0.05	Negligible

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q15	5.00	0.03	-0.11	0.27	Negligible
Q16	14.41	0.00	0.19	-0.43	Negligible
Q17	50.93	0.00	0.31	-0.74	Negligible
Q18	5.62	0.02	-0.15	0.36	Negligible
Q19	16.05	0.00	0.19	-0.44	Negligible
Q20	0.17	0.68	-0.02	0.05	Negligible
Q21	3.70	0.05	0.12	-0.28	Negligible
Q22	0.00	0.97	0.00	-0.01	Negligible
Q23	17.49	0.00	-0.22	0.52	Negligible
Q24	64.64	0.00	-0.40	0.95	Negligible
Q25	0.65	0.42	0.04	-0.10	Negligible
Q26	0.57	0.45	0.04	-0.10	Negligible
Q27	16.72	0.00	-0.19	0.43	Negligible
Q28	98.38	0.00	0.48	-1.13	Moderate
Q29	0.00	0.99	0.00	-0.01	Negligible
Q30	0.27	0.60	0.03	-0.07	Negligible
Q31	2.97	0.08	0.10	-0.23	Negligible
Q32	3.09	0.08	0.09	-0.20	Negligible
Q33	39.37	0.00	0.31	-0.72	Negligible
Q34	0.00	0.96	0.00	0.01	Negligible
Q35	2.90	0.09	-0.09	0.22	Negligible
Q36	3.68	0.05	0.10	-0.23	Negligible
Q37	3.34	0.07	0.09	-0.21	Negligible
Q38	16.45	0.00	0.17	-0.40	Negligible
Q39	13.85	0.00	0.18	-0.42	Negligible
Q40	7.23	0.01	0.16	-0.38	Negligible
Q41	0.22	0.64	0.03	-0.06	Negligible
Q42	14.39	0.00	-0.21	0.50	Negligible
Q43	0.00	0.99	0.00	0.00	Negligible
Q44	10.52	0.00	-0.17	0.39	Negligible
Q45	0.64	0.42	-0.04	0.10	Negligible
Q46	0.43	0.51	-0.03	0.07	Negligible
Q47	0.01	0.93	0.01	-0.02	Negligible
Q48	0.95	0.33	0.05	-0.12	Negligible
Q49	4.13	0.04	0.11	-0.25	Negligible
Q50	5.65	0.02	0.12	-0.29	Negligible
Q51	2.21	0.14	-0.06	0.15	Negligible
Q52	74.96	0.00	-0.38	0.90	Negligible

Table 78: Grade 4 FL

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q1	1.25	0.26	-0.11	0.25	Negligible
Q2	0.86	0.35	0.11	-0.25	Negligible
Q3	0.26	0.61	0.06	-0.13	Negligible
Q4	0.02	0.89	0.02	-0.04	Negligible
Q5	3.95	0.05	-0.16	0.37	Negligible
Q6	0.17	0.68	-0.04	0.09	Negligible
Q7	4.04	0.04	0.18	-0.43	Negligible
Q8	1.37	0.24	0.13	-0.30	Negligible
Q9	1.69	0.19	0.14	-0.34	Negligible
Q10	0.88	0.35	-0.08	0.19	Negligible
Q11	0.02	0.90	0.02	-0.04	Negligible
Q12	0.00	0.96	-0.01	0.02	Negligible
Q13	2.59	0.11	-0.15	0.34	Negligible
Q14	0.15	0.70	0.05	-0.11	Negligible
Q15	0.25	0.62	0.05	-0.13	Negligible

Table 79: Grade 8 Maths

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q1	45.84	0.00	-0.29	0.69	Negligible
Q2	85.03	0.00	0.42	-0.99	Negligible
Q3	10.68	0.00	0.14	-0.32	Negligible
Q4	8.06	0.00	0.13	-0.30	Negligible
Q5	85.60	0.00	-0.43	1.01	Moderate
Q6	15.24	0.00	0.15	-0.36	Negligible
Q7	4.66	0.03	0.11	-0.27	Negligible
Q8	5.16	0.02	0.09	-0.20	Negligible
Q9	1.52	0.22	-0.05	0.12	Negligible
Q10	0.26	0.61	-0.02	0.05	Negligible
Q11	1.93	0.16	0.05	-0.13	Negligible
Q12	47.13	0.00	-0.28	0.65	Negligible
Q13	89.79	0.00	-0.43	1.01	Moderate
Q14	12.36	0.00	0.15	-0.35	Negligible
Q15	2.49	0.11	-0.06	0.15	Negligible
Q16	7.47	0.01	-0.11	0.26	Negligible
Q17	8.61	0.00	0.12	-0.29	Negligible
Q18	0.20	0.65	-0.02	0.04	Negligible

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q19	32.16	0.00	-0.24	0.55	Negligible
Q20	0.09	0.77	-0.01	0.03	Negligible
Q21	1.34	0.25	0.05	-0.11	Negligible
Q22	19.05	0.00	-0.18	0.42	Negligible
Q23	42.31	0.00	0.27	-0.64	Negligible
Q24	0.51	0.48	-0.03	0.07	Negligible
Q25	0.11	0.74	-0.01	0.03	Negligible
Q26	10.65	0.00	0.13	-0.31	Negligible
Q27	24.81	0.00	-0.21	0.50	Negligible
Q28	85.96	0.00	-0.40	0.94	Negligible
Q29	38.84	0.00	0.24	-0.56	Negligible
Q30	0.33	0.57	0.02	-0.06	Negligible
Q31	32.21	0.00	0.24	-0.56	Negligible
Q32	5.92	0.01	-0.11	0.25	Negligible
Q33	0.01	0.92	0.01	-0.01	Negligible
Q34	26.25	0.00	0.21	-0.50	Negligible
Q35	13.80	0.00	0.16	-0.37	Negligible
Q36	3.42	0.06	-0.09	0.20	Negligible
Q37	1.04	0.31	0.04	-0.10	Negligible
Q38	0.17	0.68	-0.02	0.04	Negligible
Q39	20.70	0.00	0.20	-0.48	Negligible
Q40	0.02	0.89	0.01	-0.02	Negligible
Q41	0.53	0.47	-0.03	0.07	Negligible
Q42	9.03	0.00	-0.13	0.31	Negligible
Q43	4.25	0.04	0.09	-0.22	Negligible
Q44	4.17	0.04	0.09	-0.21	Negligible
Q45	10.19	0.00	0.15	-0.34	Negligible
Q46	19.19	0.00	-0.23	0.55	Negligible
Q47	12.71	0.00	-0.16	0.38	Negligible
Q48	0.00	0.97	0.00	-0.01	Negligible
Q49	3.64	0.06	-0.08	0.19	Negligible
Q50	7.92	0.00	0.13	-0.31	Negligible
Q51	3.65	0.06	-0.09	0.21	Negligible
Q52	10.57	0.00	0.16	-0.38	Negligible

Table 80: Grade 8 Science

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q1	13.77	0.00	-0.17	0.40	Negligible

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q2	18.29	0.00	-0.22	0.52	Negligible
Q3	66.13	0.00	0.36	-0.84	Negligible
Q4	25.07	0.00	0.29	-0.67	Negligible
Q5	17.06	0.00	0.21	-0.49	Negligible
Q6	14.63	0.00	0.16	-0.37	Negligible
Q7	14.44	0.00	-0.17	0.41	Negligible
Q8	9.22	0.00	0.13	-0.31	Negligible
Q9	0.97	0.32	-0.04	0.10	Negligible
Q10	3.59	0.06	-0.08	0.19	Negligible
Q11	39.40	0.00	0.28	-0.66	Negligible
Q12	33.63	0.00	0.26	-0.62	Negligible
Q13	0.67	0.41	-0.04	0.09	Negligible
Q14	6.07	0.01	0.11	-0.25	Negligible
Q15	2.54	0.11	0.07	-0.16	Negligible
Q16	5.54	0.02	-0.10	0.24	Negligible
Q17	0.16	0.69	0.02	-0.04	Negligible
Q18	80.03	0.00	-0.38	0.89	Negligible
Q19	82.30	0.00	0.40	-0.94	Negligible
Q20	31.35	0.00	-0.25	0.59	Negligible
Q21	22.19	0.00	-0.21	0.49	Negligible
Q22	31.55	0.00	0.24	-0.55	Negligible
Q23	11.71	0.00	-0.14	0.33	Negligible
Q24	0.02	0.88	0.01	-0.02	Negligible
Q25	0.01	0.93	0.00	0.01	Negligible
Q26	9.74	0.00	0.12	-0.29	Negligible
Q27	1.18	0.28	-0.05	0.11	Negligible
Q28	12.37	0.00	-0.14	0.34	Negligible
Q29	0.60	0.44	-0.04	0.08	Negligible
Q30	0.27	0.60	0.02	-0.06	Negligible
Q31	7.37	0.01	0.11	-0.26	Negligible
Q32	43.12	0.00	0.28	-0.65	Negligible
Q33	0.43	0.51	0.03	-0.07	Negligible
Q34	11.58	0.00	-0.15	0.35	Negligible
Q35	42.60	0.00	0.28	-0.66	Negligible
Q36	12.86	0.00	-0.15	0.34	Negligible
Q37	17.64	0.00	0.19	-0.45	Negligible
Q38	60.89	0.00	-0.32	0.76	Negligible
Q39	8.57	0.00	-0.12	0.29	Negligible
Q40	7.53	0.01	-0.12	0.28	Negligible

Item	MH Chi Square	P-Value	Log Odds Ratio	MH D-DIF	Effect Size
Q41	2.49	0.11	0.07	-0.15	Negligible
Q42	77.27	0.00	0.37	-0.87	Negligible
Q43	2.08	0.15	0.06	-0.14	Negligible
Q44	90.86	0.00	-0.40	0.94	Negligible
Q45	5.23	0.02	-0.10	0.24	Negligible
Q46	0.00	0.95	0.00	-0.01	Negligible
Q47	16.29	0.00	-0.17	0.40	Negligible
Q48	1.62	0.20	-0.05	0.13	Negligible
Q49	16.60	0.00	-0.17	0.40	Negligible
Q50	18.62	0.00	-0.18	0.43	Negligible
Q51	7.33	0.01	-0.11	0.27	Negligible
Q52	18.18	0.00	0.20	-0.46	Negligible

5. Rasch difficulties

Rasch difficulties from unanchored analyses of each test are below. The ability scale used in these analyses was defined to have a mean of zero. The fit of each item to the Rasch model in terms of the INFIT and OUTFIT indices is also shown (see Wright and Masters, 1990[52]). According to Wright and Linacre (1994)[53] these values should ideally be between 0.8 and 1.2 for high stakes[54] MCQ tests.

Table 81: Grade 4 English

Item	N	Difficulty	INFIT	OUTFIT
1	10591	0.04	1.23	1.30
2	10591	-0.72	0.98	0.95
3	10591	-0.02	0.91	0.87
4	10591	-0.22	1.10	1.09
5	10591	0.06	1.01	1.00
6	10591	-0.44	1.05	1.03
7	10591	-0.25	0.97	0.94
8	10591	-0.13	0.98	1.00
9	10591	-0.15	1.04	1.03
10	10591	0.17	1.35	1.46

[52] Wright, B.D., and Masters, G.N. (1990). Computation of OUTFIT and INFIT Statistics. Rasch Measurement Transactions, 3:4 p84-5. https://www.rasch.org/rmt/rmt34e.htm

[53] Wright, B.D., Linacre, J.M. (1994). Reasonable mean-square fit values. Rasch Measurement Transactions, 8:3 p370. https://www.rasch.org/rmt/rmt83b.htm

[54] Although the NAT is not a high stakes test for students, we use the recommendation of Wright and Linacre as a guideline to inform interpretation.

Item	N	Difficulty	INFIT	OUTFIT
11	10591	0.09	1.22	1.26
12	10591	0.06	1.09	1.10
13	10591	0.40	1.50	1.73
14	10591	-0.12	1.14	1.21
15	10591	-1.11	0.95	0.89
16	10591	-0.59	0.97	1.00
17	10591	-0.49	0.86	0.79
18	10591	-0.38	1.12	1.16
19	10591	-0.49	0.88	0.82
20	10591	-0.01	1.06	1.10
21	10591	0.05	1.24	1.29
22	10591	-0.02	0.97	0.94
23	10591	-0.13	1.09	1.10
24	10591	-0.38	0.92	0.88
25	10591	0.11	0.99	0.98
26	10591	-0.37	0.95	0.90
27	10591	-0.16	0.88	0.83
28	10591	-0.36	0.85	0.79
29	10591	-0.33	0.99	0.98
30	10591	-0.23	0.78	0.71
31	10591	-0.13	0.93	0.90
32	10591	-0.09	0.89	0.88
33	10591	-0.55	0.88	0.78
34	10591	-0.55	0.91	0.84
35	10591	-0.06	0.88	0.83
36	10591	-0.04	0.92	0.87
37	10591	-0.07	0.80	0.76
38	10591	-0.18	0.87	0.81
39	10591	-0.28	0.84	0.79
40	10591	-0.38	0.82	0.75
41	10591	-0.29	0.88	0.86
42	10591	0.01	1.17	1.19
43	10591	-0.22	0.90	0.85
44	10591	-0.12	1.03	0.99
45	10591	-0.27	1.02	0.99
46	10591	-0.06	1.12	1.13
47	10591	-0.32	0.97	0.93
48	10591	-0.23	1.02	1.01

Table 82: Grade 4 Maths

Item	N	Difficulty	INFIT	OUTFIT
1	10578	-0.97	1.00	1.03
2	10578	-0.43	1.05	1.10
3	10578	-0.27	0.96	0.94
4	10578	-0.10	0.98	0.98
5	10578	-0.52	0.94	0.91
6	10578	-0.43	0.92	0.87
7	10578	-0.23	1.05	1.08
8	10578	0.50	0.92	0.89
9	10578	-0.35	0.98	0.95
10	10578	0.57	1.03	1.08
11	10578	-0.15	0.97	0.95
12	10578	0.00	0.96	0.94
13	10578	-0.16	1.01	1.02
14	10578	0.07	0.92	0.89
15	10578	-0.06	1.00	0.99
16	10578	-0.36	0.86	0.79
17	10578	-0.01	0.94	0.91
18	10578	-0.08	0.91	0.88
19	10578	-0.21	0.96	0.91
20	10578	-0.44	0.88	0.82
21	10578	-0.60	0.89	0.83
22	10578	0.34	0.92	0.89
23	10578	0.07	0.91	0.87
24	10578	0.22	1.02	1.02
25	10578	0.34	0.89	0.88
26	10578	-0.15	1.00	0.99
27	10578	1.00	1.39	2.09
28	10578	-0.21	0.87	0.82
29	10578	0.19	1.10	1.16
30	10578	0.10	1.22	1.27
31	10578	-0.34	0.94	0.88
32	10578	0.82	1.25	1.58
33	10578	0.85	1.20	1.40
34	10578	0.20	0.94	0.92
35	10578	0.50	1.19	1.29
36	10578	1.20	1.37	2.18
37	10578	-0.15	0.89	0.85
38	10578	0.18	1.19	1.22

Item	N	Difficulty	INFIT	OUTFIT
39	10578	-0.14	0.91	0.85
40	10578	-0.11	1.13	1.21
41	10578	0.06	0.96	0.96
42	10578	-0.10	0.85	0.81
43	10578	-0.29	0.92	0.88
44	10578	0.08	0.97	0.96
45	10578	0.31	1.04	1.07
46	10578	0.25	0.88	0.86
47	10578	0.08	0.92	0.90
48	10578	-0.29	0.85	0.82

Table 83: Grade 4 Urdu and Sindhi

Item	N	Difficulty	INFIT	OUTFIT
1	10662	-1.09	0.96	0.99
2	10662	-0.99	0.91	0.90
3	10662	-1.14	0.89	0.81
4	10662	-0.59	1.01	1.02
5	10662	-1.07	0.93	0.85
6	10662	-0.88	1.13	1.19
7	10662	-0.07	1.16	1.22
8	10662	-0.31	1.23	1.48
9	10662	-1.43	0.94	0.92
10	10662	-0.43	1.00	0.93
11	10662	-0.91	0.86	0.77
12	10662	-0.79	0.93	0.87
13	10662	-0.53	1.02	0.95
14	10662	-1.18	0.92	0.90
15	10662	-0.78	1.02	1.11
16	10662	-0.58	1.03	1.13
17	10662	-0.29	1.22	1.25
18	10662	-1.30	0.89	0.73
19	10662	-0.10	1.09	1.12
20	10662	-0.28	0.85	0.78
21	10662	-1.14	0.88	0.76
22	10662	-0.74	0.83	0.71
23	10662	-0.63	0.89	0.81
24	10662	-0.55	0.94	0.91
25	10662	-0.54	0.94	0.95

Item	N	Difficulty	INFIT	OUTFIT
26	10662	-0.82	0.90	0.75
27	10662	-0.10	1.09	1.09
28	10662	-0.04	0.99	0.97
29	10662	-0.95	0.90	0.83
30	10662	-1.01	0.96	1.09
31	10662	-0.65	0.85	0.75
32	10662	-0.33	0.99	0.90
33	10662	-0.13	0.95	0.90
34	10662	-0.66	0.98	0.88
35	10662	-0.84	0.90	0.81
36	10662	-0.71	0.96	0.83
37	10662	-0.41	0.98	0.93
38	10662	-0.31	1.40	1.72
39	10662	-0.24	1.01	0.95
40	10662	-0.88	0.82	0.73
41	10662	-0.88	0.96	0.86
42	10662	-0.82	0.87	0.77
43	10662	-0.72	1.17	1.19
44	10662	-0.80	1.03	1.21
45	10662	-0.70	0.97	0.86
46	10662	-0.31	1.12	1.10
47	10662	-0.96	0.88	0.80
48	10662	-0.47	0.97	0.90
49	10662	-0.78	0.97	0.90
50	10662	-0.64	0.92	0.86
51	10662	-0.18	1.26	1.37
52	10662	-0.03	1.15	1.21

Table 84: Grade 4 FL

Item	N	Difficulty	INFIT	OUTFIT
1	9354	-2.52	0.80	0.63
2	9354	-2.41	0.54	0.35
3	9354	-2.46	0.63	0.38
4	9354	-2.46	0.60	0.38
5	9354	-2.00	0.75	0.59
6	9354	-2.22	0.64	0.47
7	9354	-1.85	0.52	0.39
8	9354	-2.51	0.60	0.34

Item	N	Difficulty	INFIT	OUTFIT
9	9354	-2.38	0.54	0.30
10	9354	-2.36	0.83	0.65
11	9354	-2.37	0.57	0.36
12	9354	-2.18	0.69	0.52
13	9354	-2.28	0.69	0.50
14	9354	-2.51	0.60	0.37
15	9354	-2.43	0.61	0.42

Table 85: Grade 8 Maths

Item	N	Difficulty	INFIT	OUTFIT
1	11882	-0.68	1.03	1.16
2	11882	-0.70	0.94	0.89
3	11882	-0.27	0.95	0.93
4	11882	-0.62	0.93	0.89
5	11882	-0.57	0.88	0.81
6	11882	-0.01	1.03	1.03
7	11882	1.03	1.01	1.11
8	11882	0.13	1.20	1.31
9	11882	-0.07	1.01	1.01
10	11882	-0.24	1.01	1.04
11	11882	0.08	1.05	1.06
12	11882	-0.19	0.98	0.98
13	11882	-0.52	0.90	0.84
14	11882	0.18	0.94	0.93
15	11882	0.25	1.07	1.09
16	11882	-0.12	0.95	0.93
17	11882	-0.24	0.95	0.90
18	11882	0.54	1.19	1.24
19	11882	0.08	0.93	0.91
20	11882	0.21	1.06	1.06
21	11882	0.26	1.08	1.09
22	11882	0.31	1.04	1.04
23	11882	0.15	0.93	0.91
24	11882	0.20	0.96	0.95
25	11882	0.50	1.17	1.21
26	11882	0.28	1.01	1.01
27	11882	0.33	0.94	0.94
28	11882	0.29	0.91	0.90
29	11882	0.22	1.16	1.23

Item	N	Difficulty	INFIT	OUTFIT
30	11882	0.49	1.10	1.12
31	11882	0.12	0.94	0.94
32	11882	0.28	0.91	0.89
33	11882	0.56	0.98	0.99
34	11882	0.20	0.97	0.95
35	11882	0.16	0.94	0.91
36	11882	0.50	0.88	0.89
37	11882	0.16	0.89	0.86
38	11882	0.36	0.96	0.96
39	11882	0.52	0.96	0.98
40	11882	0.23	0.88	0.86
41	11882	0.39	1.18	1.24
42	11882	0.56	1.00	1.03
43	11882	0.37	0.92	0.93
44	11882	0.68	1.11	1.16
45	11882	0.79	1.09	1.14
46	11882	0.71	0.81	0.81
47	11882	0.52	0.93	0.92
48	11882	0.84	1.05	1.11
49	11882	0.57	1.05	1.05
50	11882	0.71	0.98	1.02
51	11882	0.43	0.86	0.84
52	11882	1.07	1.16	1.33

Table 86: Grade 8 Science

Item	N	Difficulty	INFIT	OUTFIT
1	11799	-0.81	1.01	1.12
2	11799	-0.91	0.90	0.85
3	11799	-0.37	0.94	0.92
4	11799	1.12	0.93	1.12
5	11799	-0.91	0.93	0.91
6	11799	-0.35	1.06	1.09
7	11799	0.56	0.99	1.00
8	11799	-0.52	0.99	0.99
9	11799	-0.28	0.90	0.85
10	11799	0.45	1.18	1.21
11	11799	-0.39	0.92	0.87
12	11799	0.81	1.18	1.28

Item	N	Difficulty	INFIT	OUTFIT
13	11799	-0.69	0.92	0.90
14	11799	-0.25	0.94	0.89
15	11799	-0.49	1.04	1.07
16	11799	-0.41	0.93	0.88
17	11799	-0.49	0.98	0.96
18	11799	0.04	0.95	0.91
19	11799	0.66	1.14	1.20
20	11799	-0.49	0.91	0.84
21	11799	-0.07	0.88	0.85
22	11799	0.45	1.12	1.19
23	11799	0.03	1.05	1.02
24	11799	0.17	1.21	1.26
25	11799	-0.10	1.07	1.09
26	11799	0.13	1.20	1.22
27	11799	-0.09	0.91	0.86
28	11799	-0.33	1.03	1.08
29	11799	0.53	1.02	1.11
30	11799	0.43	0.99	1.03
31	11799	-0.03	1.01	1.01
32	11799	-0.35	1.00	1.00
33	11799	0.45	0.92	0.94
34	11799	-0.14	0.90	0.85
35	11799	0.36	1.05	1.11
36	11799	-0.03	1.04	1.02
37	11799	-0.49	0.90	0.84
38	11799	0.12	0.99	1.03
39	11799	0.30	1.06	1.08
40	11799	0.24	0.93	0.91
41	11799	0.04	1.00	0.99
42	11799	-0.13	0.97	0.93
43	11799	0.18	0.99	0.98
44	11799	0.03	0.94	0.91
45	11799	-0.33	0.91	0.85
46	11799	-0.18	1.07	1.09
47	11799	0.22	0.99	0.99
48	11799	-0.04	0.94	0.91
49	11799	-0.26	0.98	1.00
50	11799	-0.16	0.95	0.92
51	11799	-0.32	0.97	0.94
52	11799	0.59	0.99	1.07

Appendix 4: Background questionnaire responses

The following tables show the percentage of survey respondents giving each available answer to each question in the student, parent, teacher and headteacher questionnaires. Each table also provides the total number of respondents answering each question excluding those giving invalid responses (e.g. "Not Attempted" or "Not applicable") and those for whom the question was clearly not relevant. For example, teachers of subjects other than Maths are not included in the totals for the questionnaire section relating to the teaching of Maths. Similarly, as another example, students who have said that they are not physically punished in school are not included in the totals for the questions regarding the reasons they are punished.

Results for questions that allowed free responses are not included in the following tables. This includes questions asking respondents "how many" of certain items they have access to. At the time of writing, considerable data cleaning is needed to make the data from these questions usable and so it is not possible to include this data in this early report.

Headteacher survey

Table 1: Distribution of headteachers by gender

	Grade 4	Grade 8
	%	%
Male	53.6	45.8
Female	46.4	54.2
Total number of respondents	481	517
Percentage of respondents providing a valid response	100%	100%

Table 2: Total teaching and managerial experience *of headteacher*

	Grade 4	Grade 8
	%	%
< 2 years	2.9	1.4
2-5 years	0.8	2.3
6-10 years	8.5	4.4
11-15 years	10.0	8.9
more than 15 years	77.8	83.0
Total number of respondents	481	517
Percentage of respondents providing a valid response	100%	100%

Table 3: Total teaching and managerial experience of headteacher in current school

	Grade 4	Grade 8
	%	%
<2 years	18.1	26.3
2-5 years	18.7	21.3
6-10 years	20.8	17.4
11-15 years	12.9	12.4
more than 15 years	29.5	22.6
Total number of respondents	481	517
Percentage of respondents providing a valid response	100%	100%

Table 4: Headteacher academic qualification

	Grade 4	Grade 8
	%	%
Secondary School Certificate/Metric	1.5	1.4
Intermediate	3.5	0.0
Diploma	1.7	0.4
Bachelor Degree (B.A, B.Sc, AD)	25.8	10.4
Master's degree (M.A, M.Sc.) B.S. Honors	56.3	68.9
M. Phil or Ph.D.	7.5	18.8
Not attempted	3.7	0.2
Total number of respondents	481	517
Percentage of respondents providing a valid response	100%	100%

Table 5: Headteacher professional qualification

	Grade 4	Grade 8
	%	%
Primary Teacher Certificate (P.T.C)	25.4	1.7
Certificate in Teaching (C.T)	6.4	2.5
Diploma in Education	0.2	1.5
Bachelor in Education (B.Ed.) or Bachelor in Science Education (B.S.Ed.)	23.9	24.2
Master in Education (M.Ed.) or Master in Science Education (M.S.Ed.)	37.0	57.1
M.Phil/ M.S.	3.3	8.7
Ph.D.	1.5	3.7
Not attempted	2.3	0.6
Total number of respondents	481	517
Percentage of respondents providing a valid response	100%	100%

Table 6: Headteacher continuous professional development

	Grade 4		Grade 8	
	yes	no	yes	no
	%		%	
Have you attended any kind of Continuous Professional Development workshop or training during the previous two years?	48.2	51.8	35.6	64.4
Total number of respondents	481		517	
Percentage of respondents providing a valid response	100%		100%	

Table 7: Headteacher opinion on teachers' experience and parents' interest

		Grade 4	Grade 8
		%	%
The teachers are well-aware about the objectives of the National Curriculum.	the least	5.2	8.5
	less	4.8	7.0
	average	33.7	40.0
	more	40.1	34.6
	the most	16.2	9.9
The teachers have proper command over content and teaching methodologies.	the least	4.8	5.4
	less	4.2	7.4
	average	23.9	26.7
	more	43.2	45.3
	the most	23.9	15.3
The teachers use AV aids during classroom teachings.	the least	9.4	12.2
	less	12.1	18.0
	average	33.1	37.9
	more	33.7	24.4
	the most	11.9	7.5
Parents extend their cooperation with teachers to bring improvement in teaching learning process.	the least	27.0	24.6
	less	26.4	33.8
	average	27.0	28.0
	more	11.4	8.7
	the most	8.1	4.8
Parents take interest in teaching learning activities of the school.	the least	30.4	26.7
	less	28.9	29.2
	average	21.8	32.7
	more	13.9	8.5
	the most	5.0	2.9
The Parents Teacher Association facilitates and cooperates with school to improve the standard of learning	the least	10.2	9.1
	less	11.6	11.2
	average	38.0	40.4
	more	26.2	31.1
	the most	13.9	8.1
The Parents Teacher Association help school to solve problems.	the least	13.3	8.9
	less	17.0	14.9
	average	30.6	36.2
	more	23.1	31.5
	the most	16.0	8.5

Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 8: Headteacher opinion on parental invitation to participate in activities

	Grade 4		Grade 8	
	%		%	
Do you invite parents in the school to participate in the following activities?	yes	no	yes	no
Co-curricular activities e.g. debates and sports etc.	77.5	22.5	79.1	20.9
For the financial aid of the school.	33.7	66.3	40.6	59.4
For voluntarily services of the school.	70.1	29.9	72.0	28.0
To inform about the punitive activities of their child / children.	89.2	20.8	89.6	10.4
To inform about the academic progress of the students.	90.9	9.1	92.3	7.7
To meet the teachers on open day or at result declaration day.	86.5	13.5	89.4	10.6
To inform them and for the solution of physical, psychological, mental and social issues of the student	89.8	10.2	91.3	8.7
Total number of respondents	481		517	
Percentage of respondents providing a valid response	97.5%		98.3%	

Table 9: Headteacher opinion on personal development of the students

		Grade 4	Grade 8
Do you give proper importance to the following indicators for the personal development of the students?		%	%
Motivational activities in the Morning Meeting/Assembly.	never	4.6	6.2
	seldom	9.4	7.5
	often	38.3	45.6
	always	47.8	40.6
To conduct the programs for the development of the students.	never	3.3	5.0
	seldom	19.3	14.5
	often	44.5	46.4
	always	32.8	34.0
To diagnose mistakes of students	never	4.8	5.8
	seldom	5.2	3.5
	often	22.2	33.1
	always	67.8	57.6
To correct the remedial issues of the academically weaker students.	never	4.0	5.4
	seldom	5.0	8.5
	often	30.1	34.6
	always	60.9	51.5

		Grade 4	Grade 8
To provide equal opportunities of co-curricular activities to all students.	never	3.3	3.1
	seldom	12.7	13.5
	often	41.0	41.8
	always	43.0	41.6
To invite the influential and learned dignitaries of the locality in the school.	never	11.2	7.0
	seldom	41.6	40.0
	often	23.7	31.1
	always	23.5	21.9
To take steps to avoid the corporeal punishment.	never	4.0	4.6
	seldom	10.6	10.1
	often	17.7	20.1
	always	67.8	65.2
To arrange excursions and field trips for students.	never	43.9	28.6
	seldom	30.6	40.0
	often	13.9	15.7
	always	11.6	15.7
To inform the students about the current affairs, issues and incidents in the society.	never	4.6	5.4
	seldom	15.8	13.0
	often	34.5	44.1
	always	45.1	37.5
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 10: Headteacher opinion on regularity of activities

		Grade 4	Grade 8
How many times do you do the following activities?		%	%
Assessment of the students.	never	2.1	4.3
	annually	4.8	4.8
	monthly	43.9	45.1
	weekly	29.7	28.2
	daily	19.5	17.6
Provision of the basic necessities of the students.	never	8.1	5.6
	annually	7.9	6.0
	monthly	14.1	10.4
	weekly	12.9	13.9
	daily	57.0	64.0
Consultation for Lesson Planning.	never	5.6	6.0
	annually	2.9	5.2
	monthly	20.4	24.4
	weekly	36.2	36.8
	daily	34.9	27.7
Inspection of Classrooms.	never	3.7	2.3
	annually	2.9	4.8
	monthly	5.2	7.5
	weekly	18.1	18.8
	daily	70.1	66.5
Inspection of teacher's progress.	never	4.0	4.6

		Grade 4	Grade 8
	annually	4.6	5.0
	monthly	18.7	20.5
	weekly	32.2	28.6
	daily	40.5	41.2
Suggestions for the guidance of the teachers and students.	never	6.0	6.0
	annually	5.2	4.3
	monthly	21.8	25.9
	weekly	28.5	27.5
	daily	38.5	35.4
Meeting with parents for the betterment and advancement of the institution.	never	5.6	8.9
	annually	12.3	10.3
	monthly	56.3	54.5
	weekly	14.1	9.9
	daily	11.6	16.4
Addresses and Announcements about different issues in the Morning Assembly to make aware the student	never	4.2	3.9
	annually	4.4	5.0
	monthly	13.7	11.6
	weekly	34.7	37.5
	daily	43.0	42.0
Cleanliness of the school.	never	4.4	3.5
	annually	3.3	5.2
	monthly	3.1	3.1
	weekly	5.6	5.2
	daily	83.6	83.0
To make sure the provision of the hygienic edibles in the school.	never	15.6	8.9
	annually	7.5	4.6
	monthly	2.5	4.6
	weekly	6.9	4.6
	daily	67.6	77.2
To arrange the co-curricular activities in the school.	never	6.4	5.2
	annually	15.0	20.5
	monthly	35.8	33.8
	weekly	20.0	21.3
	daily	22.9	19.1
To make sure the provision of the AV aids in the classrooms.	never	8.1	6.8
	annually	11.0	5.4
	monthly	12.7	15.1
	weekly	19.5	19.3
	daily	48.6	53.4
To make sure the use of the AV aids in the classrooms to facilitate the students.	never	8.7	6.8
	annually	5.6	5.4
	monthly	9.8	15.1
	weekly	16.2	19.3
	daily	59.7	53.4
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 11: Headteacher opinion on district educational authorities contact

		Grade 4	Grade 8
		%	%
Do the District Educational Authorities contact you for the solution of the issues of your institution?	never	7.9	8.1
	monthly	70.1	63.6
	after 3 months	10.0	11.4
	after 6 months	5.6	8.7
	annually	6.4	8.1
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 12: Headteacher opinion on district educational authorities visit

		Grade 4	Grade 8
		%	%
How many times the District Educational Authorities visit your institution for the solution of the issues of your institution?	never	4.8	8.5
	once a year	17.7	12.8
	after 6 months	9.1	13.2
	after 3 months	68.4	65.6
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 13: Headteacher opinion on district Educational Authorities activities during their visit of the school

		Grade 4	Grade 8
Do the District Educational Authorities do the following activities during their visit of the school?		%	%
To check the daily attendance of the teachers and students.	never	5.0	4.6
	seldom	5.8	13.9
	often	21.4	30.6
	always	67.8	50.9
To check the record of the school.	never	4.0	6.8
	seldom	9.4	16.4
	often	28.5	31.9
	always	58.2	44.9
To evaluate the teaching learning activities of the school.	never	4.8	4.8
	seldom	13.5	15.7
	often	34.1	34.6
	always	47.6	44.9
Overall inspection of the school.	never	6.4	6.2
	seldom	10.6	10.3
	often	22.2	28.6
	always	60.7	54.9
To accept the suggestions of the teachers for the betterment of the school.	never	10.4	13.0
	seldom	26.4	29.4
	often	29.3	35.5

		Grade 4	Grade 8
	always	33.9	25.1
To give instructions for the betterment of the school.	never	5.6	4.4
	seldom	11.9	14.5
	often	20.8	33.8
	always	61.7	47.2
To check the record of co-curricular activities of the school.	never	9.4	10.3
	seldom	20.4	23.4
	often	30.8	34.8
	always	39.5	31.5
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 14: Headteacher opinion on higher authorities giving importance to their opinions

		Grade 4	Grade 8
		%	%
Do the higher authorities give proper importance to your given opinions?	never	15.4	17.4
	seldom	31.8	28.6
	often	32.6	33.3
	always	20.2	20.7
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 15: Headteacher opinion on challenges faced by headteachers

		Grade 4	Grade 8
The challenges you have to face from the following are:		%	%
Shortage of teachers.	never	20.2	18.0
	seldom	23.5	29.8
	often	27.7	28.2
	always	28.7	24.0
Shortage of supporting or non-teaching staff.	never	64.0	55.5
	seldom	28.5	31.7
	often	6.0	9.3
	always	1.5	3.5
Continuous absenteeism of teachers.	never	64.2	47.6
	seldom	27.7	38.5
	often	4.6	10.3
	always	3.5	3.7
Lack of interest of teachers in teaching.	never	45.7	29.0
	seldom	34.7	44.9
	often	12.9	22.1
	always	6.7	4.1
Un-scheduled transfers of teachers.	never	56.5	40.8
	seldom	31.8	48.2
	often	6.7	9.1
	always	5.0	1.9

		Grade 4	Grade 8
To take leaves of teachers for long period.	never	57.4	40.4
	seldom	25.4	43.5
	often	13.3	14.5
	always	3.5	1.5
Lack of command over content material from teachers end.	never	57.8	41.8
	seldom	25.4	39.7
	often	13.3	16.1
	always	3.5	2.5
Lack of professional ethics and approach from teachers end.	never	39.5	26.5
	seldom	30.8	34.0
	often	20.6	11.4
	always	9.1	3.7
Induction of teachers without professional qualification and training.	never	58.2	48.4
	seldom	23.3	36.6
	often	12.9	11.4
	always	5.6	3.7
Lack of in-service or Continuous Professional Development courses or trainings.	never	36.0	28.2
	seldom	41.0	47.8
	often	17.3	18.0
	always	5.7	6.0
Absenteeism attitude of the students.	never	26.2	19.1
	seldom	37.0	46.2
	often	28.1	29.6
	always	8.7	5.0
Lack of interest of the students toward learning.	never	34.5	26.7
	seldom	43.7	50.7
	often	15.6	18.2
	always	6.2	4.4
Lack of discipline in students.	never	31.2	29.0
	seldom	36.4	41.6
	often	24.5	21.9
	always	7.9	7.5
Lack of AV aids in school.	never	27.9	17.8
	seldom	26.8	33.8
	often	27.4	26.4
	always	17.9	12.0
Non-cooperation attitude of the parents.	never	42.6	33.5
	seldom	33.7	44.1
	often	13.5	16.6
	always	10.2	5.8
Non-cooperation attitude of the higher authorities.	never	59.7	43.9
	seldom	21.2	34.2
	often	10.0	19.5
	always	9.1	5.0
Un-due pressure of influential and political personalities.	never	24.5	21.7
	seldom	22.0	26.3
	often	26.6	34.2
	always	26.8	17.8
Lack of financial resources.	never	35.1	28.6
	seldom	24.5	37.3
	often	22.5	22.8

		Grade 4	Grade 8
	always	17.9	11.2
Lack of basic facilities.	never	53.2	41.2
	seldom	24.9	34.2
	often	13.1	19.5
	always	8.7	5.0
Un-interesting teaching learning material.	never	66.5	60.3
	seldom	20.0	22.4
	often	6.7	11.4
	always	6.9	5.8
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 16: Headteacher opinion on student absenteeism

		Grade 4	Grade 8
		%	%
What is the daily absentee ratio/ average of the student in the school?	< 5%	53.6	38.1
	11% - 20%	30.8	45.3
	>20%	15.6	16.6
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 17: Headteacher opinion on condition of the school building

		Grade 4	Grade 8
		%	%
The condition of the building of the school:	satisfactory	74.8	81.2
	unsatisfactory	25.2	18.2
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 18: Headteacher opinion on ownership of the school building

		Grade 4	Grade 8
		%	%
The nature of the school building:	ownership of school	87.5	95.9
	in other's school building	5.0	1.7
	rental building	7.5	2.3
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 19: Headteacher opinion on type of school

		Grade 4	Grade 8
		%	%
The registration of the School:	for boys	42.8	45.8
	in for girls	45.7	48.9
	for co-education	11.4	5.2
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 20: Headteacher opinion on washroom facilities

	Grade 4		Grade 8	
	yes	no	yes	no
Activities	%		%	
Is there the facility of washroom in your school?	92.5	7.5	94.2	5.8
Washroom facilities for boys	75.1	24.9	80.5	19.5
Washroom facilities for girls	51.2	48.8	69.5	30.5
Washroom facilities for teachers	63.4	36.6	76.6	23.4
Washroom facilities for other staff	27.4	72.6	52.8	47.2
Total number of respondents	445		487	
Percentage of respondents providing a valid response	100%		100%	

Table 21: Headteacher opinion on provision of facilities

		Grade 4	Grade 8
What is the provision of the following facilities?		%	%
electricity	not available	24.3	8.5
	available but not in proper/usable condition	4.0	5.2
	available but insufficient	5.4	15.1
	available and in proper condition	66.3	71.2
water	not available	17.9	9.3
	available but not in proper/usable condition	8.1	8.5
	available but insufficient	6.9	10.8
	available and in proper condition	67.2	71.4

		Grade 4	Grade 8
boundary wall	not available	17.9	7.4
	available but not in proper/usable condition	5.4	5.2
	available but insufficient	8.3	13.2
	available and in proper condition	68.4	74.3
playground	not available	43.2	34.6
	available but not in proper/usable condition	7.3	7.9
	available but insufficient	17.5	15.9
	available and in proper condition	32.0	41.6
science laboratory	not available	77.3	16.8
	available but not in proper/usable condition	7.3	11.2
	available but insufficient	4.2	18.0
	available and in proper condition	11.2	54.0
fans	not available	23.5	33.7
	available but not in proper/usable condition	8.3	10.4
	available but insufficient	20.0	18.6
	available and in proper condition	48.2	37.3
fire pot or heater	not available	78.2	10.1
	available but not in proper/usable condition	5.2	10.6
	available but insufficient	5.8	19.0
	available and in proper condition	10.8	60.3
dispensary or medical facility	not available	66.3	66.9
	available but not in proper/usable condition	3.7	5.6
	available but insufficient	17.3	11.2
	available and in proper condition	12.7	16.2
teaching and science kit	not available	61.1	40.0
	available but not in proper/usable condition	7.3	8.9
	available but insufficient	21.0	28.4
	available and in proper condition	10.6	22.6
teacher guides	not available	42.6	37.3
	available but not in proper/usable condition	5.8	8.3
	available but insufficient	19.1	21.9
	available and in proper condition	32.4	32.5
charts, models and globes	not available	30.6	16.8
	available but not in proper/usable condition	8.5	5.6
	available but insufficient	27.9	33.7

		Grade 4	Grade 8
	available and in proper condition	33.1	43.9
curriculum booklets	not available	38.3	34.4
	available but not in proper/usable condition	4.0	4.8
	available but insufficient	22.2	25.9
	available and in proper condition	35.6	34.8
Is there the facility of computer laboratory in your school?	yes	21.2	66.7
	no	78.8	33.3
Is there the facility of internet in your school?	yes	26.8	61.1
	no	73.2	38.9
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 22: Headteacher opinion on library facilities

		Grade 4	Grade 8
		%	%
library	not available	74.6	33.7
	available but not in proper/usable condition	5.4	10.4
	available but insufficient	8.3	18.6
	available and in proper condition	11.6	37.3
number of books in library	<20	29.1	17.0
	20 - 50	8.5	6.8
	51 - 100	5.2	6.6
	>100	18.7	46.2
	not present	38.5	23.4
number of library periods of each class in a week	one	28.3	43.3
	two	7.1	11.2
	three	2.3	2.7
	more than three	4.2	5.4
	no periods	58.2	37.3
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 23: Headteacher opinion on co-curricular activities during the academic year

		Grade 4	Grade 8
Co-Curricular activities during the academic year		%	%
Hamd-o-naat and singing competitions.	not at all	10.0	9.3
	weekly	20.8	13.7
	monthly	22.0	25.7

		Grade 4	Grade 8
	quarterly	18.1	18.4
	annually	29.1	38.9
painting and drawing competitions	not at all	36.4	23.8
	weekly	6.2	5.0
	monthly	15.0	13.2
	quarterly	10.0	19.1
	annually	32.4	38.9
sports	not at all	20.4	13.0
	weekly	14.8	14.3
	monthly	9.6	10.1
	quarterly	8.9	14.9
	annually	46.4	47.8
stage shows / drama	not at all	63.8	51.8
	weekly	4.6	3.3
	monthly	4.8	6.8
	quarterly	3.7	10.1
	annually	23.1	28.0
field trips and excursions	not at all	60.5	48.7
	weekly	3.3	4.4
	monthly	8.3	6.6
	quarterly	6.4	9.7
	annually	21.4	30.6
plantation campaigns	not at all	22.7	15.9
	weekly	6.4	7.7
	monthly	8.5	7.4
	quarterly	20.6	28.0
	annually	41.8	41.0
speeches and debates	not at all	21.4	10.6
	weekly	10.4	11.6
	monthly	15.4	11.2
	quarterly	13.1	10.4
	annually	39.7	35.4
scouting and girl guides	not at all	60.7	31.3
	weekly	7.5	11.6
	monthly	6.0	11.2
	quarterly	6.7	10.4
	annually	19.1	35.4
quiz competitions	not at all	31.2	15.3
	weekly	11.6	10.8
	monthly	15.0	18.0
	quarterly	12.7	21.1

		Grade 4	Grade 8
	annually	29.5	34.8
cleanliness competitions	not at all	16.2	13.7
	weekly	35.3	27.3
	monthly	24.9	26.5
	quarterly	9.6	16.1
	annually	13.9	16.4
national and local festivals	not at all	21.8	14.7
	weekly	13.1	10.3
	monthly	16.0	21.9
	quarterly	17.5	15.3
	annually	31.6	37.9
any other	not at all	61.3	59.0
	weekly	6.4	8.3
	monthly	4.6	7.5
	quarterly	8.7	5.4
	annually	18.9	19.7
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 24: Headteacher opinion on source of income of the parents of the students

		Grade 4	Grade 8
Source of income of the parents of the students		%	%
technical workers e.g. carpenters, plumbers, mechanics etc.	0 – 10%	34.9	30.8
	11 – 25%	26.0	36.2
	26 – 50%	21.2	19.9
	>50%	7.9	6.8
	other	10.0	6.4
daily wagers, vendors	0 – 10%	24.9	31.5
	11 – 25%	21.8	28.8
	26 – 50%	20.2	19.9
	>50%	24.3	15.7
	other	8.7	4.1
shopkeepers / traders	0 – 10%	62.6	56.7
	11 – 25%	18.7	28.5
	26 – 50%	8.3	10.3
	>50%	2.2	2.1
	other	7.9	3.5
government employees	0 – 10%	71.5	63.6
	11 – 25%	13.7	19.9

		Grade 4	Grade 8
	26 – 50%	5.6	9.7
	>50%	1.2	1.9
	other	7.9	4.8
private employees	0 – 10%	56.1	51.6
	11 – 25%	24.3	26.5
	26 – 50%	5.8	10.4
	>50%	5.4	5.2
	other	8.3	6.2
housewives	0 – 10%	40.3	46.0
	11 – 25%	11.2	13.5
	26 – 50%	7.5	9.5
	>50%	28.5	24.6
	other	12.5	6.4
unemployed	0 – 10%	35.1	44.7
	11 – 25%	18.9	23.6
	26 – 50%	13.3	12.0
	>50%	15.8	14.1
	other	16.8	5.6
any other profession	0 – 10%	33.9	45.3
	11 – 25%	11.6	16.4
	26 – 50%	5.4	4.3
	>50%	6.0	3.5
	other	43.0	30.6
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Table 25: Headteacher opinion on staff or students facing violence

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
Have your staff or students faced violence during the previous academic year?	11.6	88.4	16.8	83.2
Total number of respondents	481		517	
Percentage of respondents providing a valid response	100%		100%	

Table 26: Headteacher opinion on wheelchair access

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
Is there a ramp for wheelchair in your school?	40.5	59.5	58.4	41.6
Total number of respondents	481		517	
Percentage of respondents providing a valid response	100%		100%	

Table 27: Headteacher opinion on environmental initiatives

		Grade 4	Grade 8
		%	%
Do you sensitize your students about the importance of trees?	yes	26.8	94.8
	no	73.2	5.2
If you sensitize your students about the importance of trees, then how many times do you celebrate.	once	41.4	26.7
	twice	37.3	47.6
	three times	21.3	25.7
How many plants/ trees were planted this year in your school?	1 – 20	62.6	41.2
	21 - 50	16.2	20.7
	>50	21.2	38.1
Do you sensitize your students about the bad effects of the use of polyethylene/ plastic bags?	yes	95.4	93.2
	no	4.6	6.8
Do you sensitize your students about the proper use of water?	yes	95.6	95.4
	no	4.4	4.6
Do you think that the students reflect some change because of your sensitization?	yes	94.8	94.8
	no	5.2	5.2
Total number of respondents		481	517
Percentage of respondents providing a valid response		100%	100%

Teacher survey

Table 1: Distribution of teacher by age

	Grade 4	Grade 8
What is your age?	%	%
>25 years	2.9	2.4
26 - 30 years	16.8	18.0
31 - 35 years	26.5	23.6

	Grade 4	Grade 8
36 - 40 years	17.3	18.2
41 – 45 years	17.2	21.7
>50 years	19.2	16.1
Total number of respondents	452	423
Percentage of respondents providing a valid response	99.3%	99.3%

Table 2: Teacher academic qualification

	Grade 4	Grade 8
Qualification	%	%
Matriculation (Grade 10)	3.5	5.5
Intermediate (Grade 12)	6.0	1.0
Diploma	1.8	1.2
Bachelor Degree (B.A, B.Sc, AD)	26.2	14.0
Master's degree (M. A, M.Sc.) B.S. Honors	52.1	66.7
M. Phil or Ph.D.	10.4	11.6
Total number of respondents	451	421
Percentage of respondents providing a valid response	99.1%	98.8%

Table 3: Teacher professional qualification

	Grade 4	Grade 8
Qualification	%	%
Primary Teacher Certificate (P.T.C)	19.7	4.6
Certificate in Teaching (C.T)	5.3	7.2
Diploma in Education	1.8	0.5
Bachelor in Education (B.Ed.) or Bachelor in Science Education (B.S.Ed.)	41.6	32.3
Master in Education (M.Ed.) or Master in Science Education (M.S.Ed.)	26.8	48.7
Other	4.8	6.7
Total number of respondents	437	415
Percentage of respondents providing a valid response	96.0%	97.4%

Table 4: Teacher nature of contract

		Grade 4	Grade 8
		%	%
What is the nature of your job?	permanent	92.6	97.4
	contract	3.6	1.4
	other	3.8	1.2
Total number of respondents		444	424
Percentage of respondents providing a valid response		97.6%	99.5%

Table 5: Total teaching experience of teacher

	Grade 4	Grade 8
Experience	%	%
<2 years	10.5	9.4
2-5 years	15.7	19.7
6-10 years	26.0	24.6
11-15 years	16.8	13.3
more than 15 years	30.9	32.7
Total number of respondents	446	407
Percentage of respondents providing a valid response	98.0%	95.5%

Table 6: Total teaching experience of teacher in current school

	Grade 4	Grade 8
Experience	%	%
<2 years	27.2	22.5
2-5 years	25.4	39.7
6-10 years	19.6	22.7
11-15 years	11.2	5.0
more than 15 years	16.5	10.2
Total number of respondents	448	426
Percentage of respondents providing a valid response	98.5%	100%

Table 7: Continuous professional development of teacher

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
Have you received any training related to the teaching methodology during the last two years?	60.0	40.0	28.6	71.4
Total number of respondents	455		426	
Percentage of respondents providing a valid response	100%		100%	

Table 8: Subjects taught by teacher

	Grade 4		Grade 8	
	%		%	
Which of the following subjects do you teach to your class?	yes	no	yes	no
Urdu	65.5	34.5	12.4	87.6
English	59.6	40.4	22.3	77.7
Maths	75.6	24.4	61.7	38.3
Science	52.5	47.5	29.8	70.2
Sindhi	26.4	73.6	11.3	88.7
Social studies	49.9	50.1	-	-
History	-	-	10.1	89.9
Geography	-	-	10.8	89.2
Total number of respondents	455		426	
Percentage of respondents providing a valid response	100%		100%	

Table 9: Teacher opinion on teaching guides available

	Grade 4		Grade 8	
	%		%	
Do you have the following helping material in your school to guide the students?: Teaching Guide	yes	no	yes	no
Urdu	71.8	28.2	49.0	51.0
English	64.2	35.8	47.3	52.7
Maths	50.2	49.8	20.9	79.1
Science	56.3	43.7	12.9	87.1
Sindhi	61.3	38.7	17.8	82.2
Social studies	45.4	54.6	-	-
History	-	-	26.3	73.7
Geography	-	-	14.6	85.4
Total number of respondents	435		412	
Percentage of respondents providing a valid response	94.7%		90.5%	

Table 10: Teacher opinion on teaching aids available for students

	Grade 4		Grade 8	
	%		%	
Do you have the following helping material in your school to aid the students? Teaching Aids	yes	no	yes	no
Urdu	61.3	38.7	38.0	62.0

	Grade 4		Grade 8	
English	58.6	41.4	40.4	59.6
Maths	48.1	51.9	15.2	84.8
Science	57.0	43.0	20.2	79.8
Sindhi	55.0	45.0	16.3	83.7
Social studies	48.5	51.5	-	-
History	-	-	28.9	71.1
Geography	-	-	24.4	75.6
Total number of respondents	435		412	
Percentage of respondents providing a valid response	93.3%		90.0%	

Table 11: Teacher opinion on effectiveness of teaching guides

		Grade 4	Grade 8
		%	%
How effective is the above teaching material?: Teaching guide	not at all	20.7	34.5
	to some extent	20.2	20.9
	to a reasonable extent	32.5	27.2
	much more	26.6	17.4
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 12: Teacher opinion on effectiveness of teaching aids for students

		Grade 4	Grade 8
		%	%
How effective is the above teaching material?: Teaching aids	not at all	21.5	27.2
	to some extent	16.9	16.2
	to a reasonable extent	29.5	38.3
	much more	32.1	18.3
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 13: Teacher opinion on lesson planning

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
Do you plan lessons in advance?	93.0	7.0	90.1	9.9
Total number of respondents	423		426	
Percentage of respondents providing a valid response	93.3%		100%	

Table 14: Teacher opinion on resources used for lesson planning

		Grade 4	Grade 8
If you plan lessons in advance, to what extent do you use the following sources/methods?		%	%
internet	never	19.6	12.5
	occasionally	30.0	31.8
	quire often	36.2	40.6
	much more	14.2	15.1
colleagues	never	26.0	25.3
	occasionally	35.0	35.4
	quire often	27.0	34.1
	much more	12.0	5.2
library book	never	22.2	21.9
	occasionally	19.6	23.7
	quire often	27.7	26.3
	much more	30.5	28.1
other	never	41.6	49.2
	occasionally	29.1	26.8
	quire often	17.5	17.7
	much more	11.8	6.3
Total number of respondents		423	384
Percentage of respondents providing a valid response		100%	100%

Table 15: Teacher opinion on teaching two or more classes at the same time

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
Do you have to teach two or more sections of a group?	33.8	62.2	-	-
Do you have to teach different classes in a period?	44.9	55.1	36.3	63.7
Total number of respondents	455		426	
Percentage of respondents providing a valid response	99.8%		100%	

Table 16: Teacher opinion on regularity of teaching two or more classes

		Grade 4	Grade 8
		%	%
How often do you teach two or more classes in the same period?	daily	55.3	31.5
	once a week	17.6	29.1
	every three weeks	9.5	10.8
	monthly	13.4	15.8
	not attempted	4.2	12.8
Total number of respondents		262	203
Percentage of respondents providing a valid response		100%	100%

Table 17: Teacher opinion on problems arising from teaching two or more classes together

	Grade 4		Grade 8	
	%		%	
What problems do you think arise if you teach two or more classes together and this affects the teaching process?	yes	no	yes	no
discipline problems	64.1	35.9	59.1	40.9
lack of individual attention	71.4	28.6	62.1	37.9
the feedback problem	45.8	54.2	42.9	57.1
disruption to teaching	59.5	40.5	39.4	60.6
psychological stress	38.2	61.8	21.2	78.8
fatigue in students and teachers	45.8	54.2	34.5	65.5
unmanageable number of children	45.4	54.6	41.9	58.1
lack of teaching aids	35.5	64.5	26.1	73.9
inappropriate teaching environment	39.7	60.3	38.9	61.1
any other problem	17.2	82.8	11.8	88.2
Total number of respondents	262		203	
Percentage of respondents providing a valid response	95.8%		89.2%	

Table 18: Teacher opinion on parents

		Grade 4	Grade 8
		%	%
Parents send children neat and clean in school.	never	6.8	5.4
	occasionally	33.4	29.6
	quire often	42.4	46.2
	always	17.4	18.8
Contribute to ensure children's school attendance	never	12.5	20.2
	occasionally	29.5	33.8
	quire often	36.7	35.7
	always	21.3	10.3
Try to meet the educational needs of children.	never	9.0	12.9
	occasionally	30.8	29.6
	quire often	42.6	38.3
	always	17.6	19.2
Make children complete their school homework regularly.	never	18.0	24.2
	occasionally	37.6	43.0
	quire often	29.9	21.4
	always	14.5	11.5
Extend cooperation with teachers regarding academic performance.	never	22.0	30.3
	occasionally	45.1	44.8
	quire often	21.5	13.6
	always	11.4	11.3
Help school management for the betterment of children.	never	35.2	39.0
	occasionally	32.5	35.7
	quire often	18.5	13.1
	always	13.8	12.2
	never	11.9	11.0

		Grade 4	Grade 8
Allow children to participate in curricular and co-curricular activities.	occasionally	24.0	36.6
	quire often	35.4	25.8
	always	28.8	26.5
Help in providing basic facilities for children in school.	never	40.0	39.0
	occasionally	29.2	34.0
	quire often	15.6	16.2
	always	15.2	10.8
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 19: Teacher opinion on process that affect learning

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
absence of children	84.0	16.0	83.8	16.2
lack of educational requirements	53.8	46.2	50.2	49.8
children not completing homework	62.0	38.0	68.3	31.7
lack of moral training of children	38.9	61.1	44.1	55.9
emotional and psychological problems of children	46.8	53.2	45.8	54.2
more interest of children in extra-curricular activities	37.1	62.9	35.2	64.8
excessive use of mobile/internet	49.7	50.3	66.0	34.0
overprotective parents	50.5	49.5	45.5	54.5
Total number of respondents	455		426	
Percentage of respondents providing a valid response	97.8%		99.8%	

Table 20: Teacher opinion on course completion

		Grade 4	Grade 8
		%	%
My opinion is valued in teaching subjects.	never	8.6	10.6
	occasionally	19.8	25.8
	quite often	39.8	33.8
	much more	31.9	29.8
The course ends in time.	never	6.8	6.1
	occasionally	11.6	12.7
	quite often	27.7	33.8
	much more	53.8	47.4
After completing the course, time is saved for repeating the lesson.	never	8.4	14.3
	occasionally	20.0	24.2
	quite often	31.0	29.3
	much more	40.7	32.2
You are given subjects for teaching according to your interest.	never	13.8	8.9
	occasionally	18.7	13.6
	quite often	32.3	38.5
	much more	35.2	39.0
	never	12.7	14.1

		Grade 4	Grade 8
Students approach me for the solution of their personal problems.	occasionally	29.0	31.5
	quite often	30.5	28.4
	much more	27.7	26.1
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 21: Teacher opinion on frequency of reporting to parents

		Grade 4	Grade 8
		%	%
Sending report of academic performance of students to parents	never	5.9	12.9
	annually	22.4	23.0
	nine months	4.6	7.5
	quarterly	20.9	27.0
	monthly	46.2	29.6
Student Attendance Report	never	9.2	14.6
	annually	8.1	13.1
	nine months	2.2	3.1
	quarterly	12.1	10.3
	monthly	68.4	58.9
Performance of students with respect to school discipline	never	12.7	13.6
	annually	10.5	19.7
	nine months	3.5	2.6
	quarterly	14.3	18.8
	monthly	58.9	45.3
Informing parents about children's classroom behaviour	never	8.8	12.4
	annually	6.2	11.0
	nine months	2.0	4.2
	quarterly	12.3	20.9
	monthly	70.8	51.4
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 22: Teacher opinion on evaluation of the performance of students

		Grade 4	Grade 8
		%	%
How often do you evaluate student performance using oral questions and answers?	never	7.0	12.0
	occasionally	13.0	13.4
	quite often	38.7	40.8
	always	41.3	33.8
How often do you evaluate student performance using the class list?	never	19.8	8.5
	occasionally	27.9	8.9
	quite often	31.6	42.0
	always	20.7	40.6
How often do you evaluate student performance using homework?	never	9.5	8.7
	occasionally	8.1	16.2
	quite often	27.3	27.9

		Grade 4	Grade 8
How often do you evaluate student performance using participation in class activities?	always	55.2	47.2
	never	8.8	12.4
	occasionally	16.3	27.0
	quite often	35.6	31.0
	always	39.3	29.6
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 23: Teacher opinion on regularity of informing students about mistakes

		Grade 4	Grade 8
		%	%
How often do you aware your students about their mistakes after their formative assessment?	never	3.1	13.8
	monthly	5.9	21.8
	after 15 days	3.7	9.4
	weekly	20.9	17.8
	daily	66.4	37.1
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 24: Teacher opinion on school rules and staff meetings

		Grade 4	Grade 8
		%	%
School rules and regulations are strictly enforced.	never	5.7	17.8
	monthly	6.2	9.4
	after 15 days	0.9	7.0
	weekly	7.9	11.5
	daily	79.3	54.2
The staff is invited in the meeting for the betterment of the students.	never	7.3	24.9
	monthly	38.2	38.7
	after 15 days	13.2	11.0
	weekly	26.2	14.3
	daily	15.2	11.0
Constructive advice is valued in staff meetings.	never	9.9	27.2
	monthly	39.6	35.9
	after 15 days	11.0	10.6
	weekly	15.2	8.2
	daily	24.4	18.1
I take personal interest in improving the educational quality of the school.	never	5.3	21.4
	monthly	5.3	9.2
	after 15 days	3.3	6.6
	weekly	5.5	7.7
	daily	80.7	55.2
Total number of respondents		455	426
Percentage of respondents providing a valid response		100%	100%

Table 25: Urdu teaching activities to Grade 4 students

		Grade 4
		%
practice the handwriting skills	never	6.0
	occasionally	18.8
	quite often	32.9
	always	42.3
dictation	never	7.7
	occasionally	13.1
	quite often	37.2
	always	41.9
use of allied teaching materials	never	13.4
	occasionally	36.9
	quite often	31.2
	always	18.5
creative writing practice	never	14.4
	occasionally	30.9
	quite often	26.8
	always	27.9
question answering	never	4.7
	occasionally	9.4
	quite often	20.1
	always	65.8
group activities	never	10.7
	occasionally	18.5
	quite often	31.5
	always	39.3
using examples from practical life	never	9.1
	occasionally	9.7
	quite often	23.8
	always	57.4
other	never	35.8
	occasionally	25.5
	quite often	17.4
	always	21.1
Total number of respondents		298
Percentage of respondents providing a valid response		100%

Table 26: Teacher opinion on teaching of Urdu skills in Grade 4

		Grade 4
		%
reading skills	very easy	37.9
	easy	35.2
	average	20.5
	difficult	5.0
	very difficult	1.3
writing skills	very easy	27.2
	easy	37.9
	average	23.5

		Grade 4
	difficult	10.1
	very difficult	1.3
speaking skill	very easy	35.9
	easy	36.2
	average	21.8
	difficult	4.7
	very difficult	1.3
use of grammar	very easy	21.5
	easy	33.2
	average	25.5
	difficult	17.4
	very difficult	2.3
debating skills	very easy	20.8
	easy	30.9
	average	29.9
	difficult	15.7
	very difficult	3.4
Total number of respondents		455
Percentage of respondents providing a valid response		100%

Table 27: Teacher opinion on student proficiency in Urdu skills in Grade 4

		Grade 4
		%
reading with fluency	almost all	36.2
	more than half	50.7
	less than half	9.7
	a few	2.0
	no one	1.3
reading comprehension	almost all	26.8
	more than half	47.3
	less than half	19.1
	a few	4.4
	no one	2.3
creative writing	almost all	53.4
	more than half	31.9
	less than half	10.4
	a few	2.7
	no one	1.7
Total number of respondents		298
Percentage of respondents providing a valid response		100%

Table 28: Urdu activities for Grade 4 students

		Grade 4
To what extent do you use the following issues while teaching Urdu to Grade 4		%
Asking questions about the lesson taught	never	3.7
	occasionally	10.4

		Grade 4
	quite often	27.9
	always	58.1
Writing a story using your own ideas	never	4.7
	occasionally	41.6
	quite often	37.2
	always	16.4
Encourage reading of additional material relating to Urdu subject	never	6.0
	occasionally	30.5
	quite often	40.6
	always	22.8
Encourage participation in speech and debate competition	never	5.4
	occasionally	36.2
	quite often	41.3
	always	17.1
Encourage participation in group discussions	never	4.7
	occasionally	27.2
	quite often	40.3
	always	27.9
Poetry competitions	never	21.5
	occasionally	46.0
	quite often	23.5
	always	9.1
Total number of respondents		298
Percentage of respondents providing a valid response		100%

Table 29: Urdu homework activities for Grade 4 students

		Grade 4
To what extent do you use the following activities when giving Urdu homework?		%
Practice questions from lessons taught	never	4.4
	occasionally	8.1
	quite often	21.5
	always	66.1
Use of other additional materials	never	9.4
	occasionally	37.2
	quite often	36.2
	always	17.1
Making charts etc.	never	8.4
	occasionally	44.6
	quite often	35.2
	always	11.7
Creative and descriptive writing	never	13.8
	occasionally	38.9
	quite often	33.2
	always	14.1
To practice of difficult words	never	3.7
	occasionally	9.4
	quite often	26.2
	always	60.7

		Grade 4
Use of dictionary	never	8.1
	occasionally	20.5
	quite often	36.2
	always	35.2
Total number of respondents		298
Percentage of respondents providing a valid response		100%

Table 30: Teacher opinion on training in the teaching of Urdu

	Grade 4	
	%	
	yes	no
Have you received any training related to teaching Urdu since the last two years?	36.2	63.8
Total number of respondents	298	
Percentage of respondents providing a valid response	100%	

Table 31: Teacher opinion on impact of teacher training to teach Urdu

		Grade 4
		%
Has the teacher training to help you teach Urdu enhanced your teaching skills?	absolutely not	7.4
	to some extent	16.7
	to a reasonable extent	30.6
	much more	45.4
To what extent do you use the methods discussed in the training in the classroom?	never	9.3
	occasionally	15.7
	quite often	46.3
	always	28.7
To what extent have you improved your method of teaching after receiving the training?	absolutely not	11.1
	to some extent	11.1
	to a reasonable extent	36.1
	much more	41.7
To what extent have you improved your assessment methods after receiving the training?	absolutely not	9.3
	to some extent	18.5
	to a reasonable extent	35.2
	much more	37.0
To what extent have you improved your classroom discipline after receiving the training?	absolutely not	8.3
	to some extent	13.9
	to a reasonable extent	22.2
	much more	55.6
To what extent have you increased children's interest in the subject after receiving the training?	absolutely not	9.3
	to some extent	10.2
	to a reasonable extent	23.1
	much more	57.4
To what extent have you improved your proper use of teaching AV aids after receiving the training?	absolutely not	10.2
	to some extent	17.6
	to a reasonable extent	31.5
	much more	40.7

		Grade 4
To what extent have you improved your proper use of modern instructional technology in teaching learning process after receiving the training?	absolutely not	14.8
	to some extent	23.1
	to a reasonable extent	25.9
	much more	36.1
Total number of respondents		108
Percentage of respondents providing a valid response		100%

Table 32: Maths activities for Grade 4 and Grade 8 students from the textbooks

		Grade 4	Grade 8
To what extent do you conduct the following activities from the textbook to teach Maths to Grade 4 and Grade 8students?		%	%
Giving examples from practical life	never	4.9	9.5
	occasionally	13.4	12.9
	quite often	32.8	41.8
	always	48.8	35.7
Conducting practical activities	never	8.1	14.8
	occasionally	28.2	30.8
	quite often	33.4	38.0
	always	30.2	16.3
Providing additional examples	never	12.5	17.5
	occasionally	32.3	31.6
	quite often	32.6	37.3
	always	22.7	13.7
Practice in groups	never	5.5	12.5
	occasionally	12.8	18.5
	quite often	31.4	38.0
	always	50.3	31.9
Conducting practice tests	never	5.8	12.9
	occasionally	16.3	18.3
	quite often	33.1	29.7
	always	44.8	39.2
Total number of respondents		344	263
Percentage of respondents providing a valid response		100%	100%

Table 33: Teacher opinion on teaching of mathematical skills in Grade 4 and Grade 8

		Grade 4	Grade 8
		%	%
numbers	very easy	61.3	70.3
	easy	30.5	21.7
	average	6.4	6.5
	difficult	1.2	0.4
	very difficult	0.6	1.1
multiples and factors	very easy	43.6	62.4
	easy	36.9	27.8
	average	16.9	8.4
	difficult	1.7	0.0

		Grade 4	Grade 8
	very difficult	0.9	1.5
fractions	very easy	41.9	58.2
	easy	35.5	27.0
	average	18.3	11.4
	difficult	3.2	1.9
	very difficult	1.2	1.5
fraction decimals	very easy	39.5	57.0
	easy	35.2	25.1
	average	20.1	13.7
	difficult	4.9	2.7
	very difficult	0.3	1.5
measurement	very easy	43.6	45.2
	easy	27.9	33.8
	average	25.0	16.0
	difficult	11.0	3.0
	very difficult	1.2	1.9
geometry	very easy	35.2	39.5
	easy	27.6	31.6
	average	25.0	19.8
	difficult	11.0	6.1
	very difficult	1.2	3.0
information handling	very easy	39.8	44.9
	easy	33.1	24.3
	average	20.3	23.2
	difficult	4.4	3.0
	very difficult	2.3	4.6
Total number of respondents		344	263
Percentage of respondents providing a valid response		100%	100%

Table 34: Teacher opinion on student interest in Grade 4 and Grade 8 Maths

		Grade 4	Grade 8
		%	%
How many students show interest in Maths subject?	no one	5.2	8.4
	a few	6.7	12.2
	less than half	22.4	20.2
	more than half	41.3	46.8
	almost all	24.4	12.5
Total number of respondents		344	263
Percentage of respondents providing a valid response		100%	100%

Table 35: Maths activities for Grade 4 and Grade 8 students

		Grade 4	Grade 8
To what extent do you use the following methods while teaching Maths to Grade 5 and Grade 8 students?		%	%
Use of Maths in everyday life	never	2.6	6.5

		Grade 4	Grade 8
	occasionally	14.0	15.6
	quite often	41.0	49.8
	always	42.4	28.1
To provide opportunities for brainstorming and question-and-answer	never	4.4	8.4
	occasionally	12.8	18.3
	quite often	44.5	43.7
	always	38.4	29.7
To ask questions outside the textbook	never	5.2	9.5
	occasionally	30.2	36.1
	quite often	36.0	32.7
	always	28.5	21.7
To encourage the students to ask questions about the topic	never	3.8	7.6
	occasionally	9.6	15.2
	quite often	32.3	28.9
	always	54.4	48.3
To encourage group activities	never	4.4	6.5
	occasionally	14.8	22.8
	quite often	34.0	30.3
	always	46.8	40.7
Total number of respondents		344	263
Percentage of respondents providing a valid response		100%	100%

Table 36: Maths homework activities for Grade 4 and Grade 8 students

		Grade 4	Grade 8
To what extent do you use the following activities when giving Maths homework?		%	%
Giving real-life examples of the lessons taught	never	2.9	8.7
	occasionally	19.2	25.1
	quite often	41.0	36.1
	always	36.9	30.0
Making charts or models etc. about lessons taught	never	7.8	14.4
	occasionally	43.3	40.7
	quite often	33.4	28.9
	always	15.4	16.0
Practice additional allied material of Maths	never	5.8	11.4
	occasionally	27.6	27.4
	quite often	41.0	30.4
	always	25.6	30.8
Total number of respondents		344	263
Percentage of respondents providing a valid response		100%	100%

Table 37: Teacher opinion on training in the teaching of Maths

	Grade 4		Grade 8	
	%		%	
	yes	no	yes	no
Have you received any training related to teaching Maths since the last two years?	34.9	65.1	16.0	84.0
Total number of respondents	344		263	
Percentage of respondents providing a valid response	100%		100%	

Table 38: Teacher opinion on impact of teacher training to teach Maths

		Grade 4	Grade 8
		%	%
To what extent do you use the methods discussed in the training in the classroom?	never	10.0	35.7
	occasionally	9.2	9.5
	quite often	45.8	35.7
	always	35.0	19.0
To what extent have you improved your method of teaching after receiving the training?	not at all	11.7	38.1
	to some extent	16.7	14.3
	to a reasonable extent	29.2	19.0
	much more	42.5	28.6
To what extent have you improved your assessment methods after receiving the training?	not at all	14.2	28.6
	to some extent	13.3	21.4
	to a reasonable extent	34.2	28.6
	much more	38.3	21.4
To what extent have you improved your classroom discipline after receiving the training?	absolutely not	15.0	38.1
	to some extent	9.2	7.1
	to a reasonable extent	25.0	21.4
	much more	50.8	33.3
To what extent have you improved your pre-lesson preparation after receiving the training?	absolutely not	13.3	33.3
	to some extent	12.5	9.5
	to a reasonable extent	20.8	23.8
	much more	53.3	33.3
To what extent have you increased children's interest in the subject after receiving the training?	absolutely not	12.5	33.3
	to some extent	15.0	4.8
	to a reasonable extent	25.0	38.1
	much more	47.5	23.8
To what extent have you improved your proper use of teaching AV aids after receiving the training?	absolutely not	15.0	38.1
	to some extent	13.3	14.3
	to a reasonable extent	28.3	31.0
	much more	43.3	16.7
To what extent have you improved your proper use of modern instructional technology in teaching learning process after receiving the training?	absolutely not	20.8	42.9
	to some extent	13.3	28.6
	to a reasonable extent	33.3	7.1
	much more	32.5	21.4
Total number of respondents		120	42
Percentage of respondents providing a valid response		100%	100%

Table 39: English activities for Grade 4 students

		Grade 4
To what extent, do you use following tasks to teach English of Grade 4?		%
writing practice	never	12.5
	occasionally	6.3
	quite often	23.6
	always	57.6
spelling practice	never	13.7
	occasionally	9.2
	quite often	25.1
	always	52.0
grammar translation methods	never	15.5
	occasionally	14.8
	quite often	35.4
	always	34.3
paired and group work	never	15.1
	occasionally	19.9
	quite often	32.8
	always	32.1
supplementary reading work	never	19.6
	occasionally	19.9
	quite often	34.3
	always	26.2
innovative and creative writing practice	never	20.7
	occasionally	27.3
	quite often	28.8
	always	23.2
Total number of respondents		271
Percentage of respondents providing a valid response		100%

Table 40: Teacher opinion on students' understanding of English

		Grade 4
How many students of Grade 4 can perform these tasks with ease?		%
fluent in reading	almost all	19.2
	more than half	46.9
	less than half	22.5
	a few	6.3
	no one	5.2
reading with understanding	almost all	17.7
	more than half	33.6
	less than half	31.7
	a few	10.7
	no one	6.3
writing	almost all	40.6
	more than half	33.6
	less than half	15.5
	a few	5.5

		Grade 4
	no one	4.8
Total number of respondents		271
Percentage of respondents providing a valid response		100%

Table 41: English activities for Grade 4 students

		Grade 4
While teaching English, how many times do you instruct students to do the following?		%
Asking questions about the lesson	never	14.8
	sometimes	11.1
	often	25.1
	always	49.1
Innovative and Creative writing	never	17.7
	sometimes	22.1
	often	36.2
	always	24.0
Group discussion	never	15.5
	sometimes	24.0
	often	32.8
	always	27.7
Speech contests	never	21.0
	sometimes	39.5
	often	24.0
	always	15.5
Extra Reading	never	25.5
	sometimes	31.7
	often	31.0
	always	11.8
Watching English Videos	never	46.9
	sometimes	25.5
	often	19.2
	always	8.5
Watching English TV Programs	never	53.1
	sometimes	25.5
	often	12.5
	always	8.9
Listening to English programs	never	46.5
	sometimes	30.6
	often	10.7
	always	12.2
Total number of respondents		271
Percentage of respondents providing a valid response		100%

Table 42: English homework activities for Grade 4 students

		Grade 4
To what extent, do you use following activities in assigning English homework?		%
Only textbook exercises	never	15.5
	occasionally	11.8
	quite often	25.1
	always	47.6
Innovative and Creative Reading	never	19.2
	occasionally	27.3
	quite often	32.8
	always	20.7
Extra Readings	never	25.5
	occasionally	34.7
	quite often	27.3
	always	12.5
Preparing charts/models	never	24.0
	occasionally	38.4
	quite often	26.2
	always	11.4
Learning vocabulary	never	17.0
	occasionally	22.5
	quite often	29.9
	always	30.6
Total number of respondents		
Percentage of respondents providing a valid response		100%

Table 43: Teacher opinion on training in the teaching of English

	Grade 4	
	%	
	yes	no
Have you received any training related to teaching English since the last two years?	47.6	52.4
Total number of respondents	271	
Percentage of respondents providing a valid response	100%	

Table 44: Teacher opinion on impact of teacher training to teach English

		Grade 4
		%
Has the teacher training to help you teach English enhanced your teaching skills?	not at all	16.3
	a little	9.3
	some	19.4
	a lot	55.0
To what extent do you use the methods discussed in the training in the classroom?	never	19.4
	sometimes	7.8
	often	35.7
	always	37.2
To what extent have you improved your method of teaching after receiving the training?	never	16.3
	sometimes	13.2

		Grade 4
	often	26.4
	always	44.2
To what extent have you improved your assessment methods after receiving the training?	never	17.1
	sometimes	12.4
	often	36.4
	always	34.1
To what extent have you improved your classroom discipline after receiving the training?	never	16.3
	sometimes	10.1
	often	24.8
	always	48.8
To what extent have you increased children's interest in the subject after receiving the training?	never	20.2
	sometimes	9.3
	often	32.6
	always	38.0
To what extent have you improved your effective use of teaching aids after receiving the training?	never	20.2
	sometimes	17.1
	often	31.8
	always	31.0
To what extent have you improved your use of technology in teaching after receiving the training?	never	24.8
	sometimes	17.8
	often	29.5
	always	27.9
Total number of respondents		129
Percentage of respondents providing a valid response		100%

Table 45: Science activities for Grade 8 students

		Grade 8
To what extent do you use the following activities with the textbook to teach Science to Grade 8 students?		%
Giving examples from practical life	never	11.8
	occasionally	5.5
	quite often	27.6
	always	55.1
Conducting practical activities	never	12.6
	occasionally	33.9
	quite often	29.9
	always	23.6
Providing additional content	never	15.2
	occasionally	39.4
	quite often	30.7
	always	15.0
Practice in groups	never	11.0
	occasionally	25.2
	quite often	34.6
	always	29.1
Carrying out experiments in the laboratory	never	29.9
	occasionally	29.1
	quite often	25.2
	always	15.7

		Grade 8
Practice a mental test	never	16.5
	occasionally	23.6
	quite often	31.5
	always	28.3
Total number of respondents		127
Percentage of respondents providing a valid response		100%

Table 46: Teacher opinion on teaching of science domains in Grade 8

		Grade 8
		%
Biology	very easy	41.7
	easy	32.3
	average	19.7
	difficult	1.6
	very difficult	4.7
Chemistry	very easy	30.7
	easy	42.5
	average	18.9
	difficult	3.9
	very difficult	3.9
Physics	very easy	29.9
	easy	33.9
	average	24.4
	difficult	4.7
	very difficult	7.1
Earth Science	very easy	35.4
	easy	38.6
	average	18.9
	difficult	4.7
	very difficult	2.4
Total number of respondents		127
Percentage of respondents providing a valid response		100%

Table 47: Teacher opinion on student interest in Grade 8 science

		Grade 8
		%
How many students show interest in Maths subject?	no one	15.0
	a few	10.2
	less than half	11.8
	more than half	44.1
	almost all	18.9
Total number of respondents		127
Percentage of respondents providing a valid response		100%

Table 48: Science activities for Grade 8 students

		Grade 8
To what extent do you use the following teaching activities while teaching science to Grade 8 students?		%
Asking allied questions relating to the lesson taught	never	17.3
	occasionally	10.2
	quite often	22.0
	always	50.4
Ask students to do practical experiments on their own	never	18.9
	occasionally	27.6
	quite often	32.3
	always	21.3
Making charts and projects of different topics	never	17.3
	occasionally	33.1
	quite often	26.0
	always	23.6
Encourage critical thinking about the different topics of the Science	never	18.1
	occasionally	16.5
	quite often	29.9
	always	35.4
Encourage keen observation of Earth and the Solar System	never	20.5
	occasionally	30.7
	quite often	23.6
	always	25.2
Encourage group activities	never	19.7
	occasionally	21.3
	quite often	30.7
	always	28.3
Encourage students to ask questions about the topic	never	18.1
	occasionally	9.4
	quite often	22.0
	always	50.4
Total number of respondents		127
Percentage of respondents providing a valid response		100%

Table 49: Science homework activities for Grade 8 students

		Grade 8
To what extent do you use the following activities when giving science homework?		%
To solving the exercise questions only	never	23.6
	occasionally	15.7
	quite often	32.3
	always	28.3
To instruct for practical experiments	never	22.0
	occasionally	32.3
	quite often	31.5
	always	14.2
Creating a chart or model	never	18.9
	occasionally	40.2
	quite often	17.3

		Grade 8
	always	23.6
To ask for additional study other than the textbook	never	21.3
	occasionally	20.5
	quite often	38.6
	always	19.7
Collecting other material relating to the topic	never	18.9
	occasionally	20.5
	quite often	41.7
	always	18.9
Total number of respondents		127
Percentage of respondents providing a valid response		100%

Table 50: Teacher opinion on training in the teaching of science

	Grade 8	
	%	
	yes	no
Have you received any training related to teaching science since the last two years?	26.0	74.0
Total number of respondents	127	
Percentage of respondents providing a valid response	100%	

Table 51: Teacher opinion on impact of teacher training to teach Science

		Grade 8
		%
Has the teacher training to help you teach science enhanced your teaching skills?	not at all	36.4
	a little	18.2
	some	33.3
	a lot	12.1
To what extent do you use the methods discussed in the training in the classroom?	never	36.4
	sometimes	18.2
	often	33.3
	always	12.1
To what extent have you improved your method of teaching after receiving the training?	never	42.4
	sometimes	15.2
	often	15.2
	always	27.3
To what extent have you improved your assessment methods after receiving the training?	never	33.3
	sometimes	33.3
	often	9.1
	always	24.2
To what extent have you improved your classroom discipline after receiving the training?	never	42.4
	sometimes	3.0
	often	33.3
	always	21.2
To what extent have you increased children's interest in the subject after receiving the training?	never	39.4
	sometimes	3.0
	often	30.3
	always	27.3

		Grade 8
To what extent have you improved your effective use of teaching aids after receiving the training?	never	39.4
	sometimes	18.2
	often	21.2
	always	21.2
To what extent have you improved your use of technology in teaching after receiving the training?	never	51.5
	sometimes	9.1
	often	12.1
	always	27.3
To what extent have you improved your use of pre-lesson preparation after receiving the training?	never	39.4
	sometimes	6.1
	often	12.1
	always	42.4
Total number of respondents		33
Percentage of respondents providing a valid response		100%

Student survey

Table 1: Distribution of student by gender

	Grade 4	Grade 8
	%	%
male	48.2	49.8
female	51.7	50.2
transgender	0.0	0.0
Total number of respondents	10378	11558
Percentage of respondents providing a valid response	100%	100%

Table 2: Home spoken language of student

	Grade 4		Grade 8	
	%		%	
	yes	**no**	**yes**	**no**
Urdu	35.9	64.1	44.6	55.4
Brahawi	4.1	95.9	3.2	96.8
Balochi	4.9	95.1	4.2	95.8
Punjabi	25.2	74.8	24.9	75.1
Pashtu	28.4	71.6	23.6	76.4
Sindhi	10.8	89.2	13.1	86.9
Siraiki	7.4	92.6	7.3	92.7
Kashmiri	2.5	97.5	3.4	96.6
Hindko	4.7	95.3	6.9	93.1

	Grade 4		Grade 8	
Farsi	2.6	97.4	2.9	97.1
Bulti	2.2	97.8	3.4	96.6
Shena	3.9	96.1	4.4	95.6
Any other	5.1	94.9	5.6	94.4
Total number of respondents	10378		11558	
Percentage of respondents providing a valid response	97.9%		98.3	

Table 3: Student opinion on homework help

	Grade 4		Grade 8	
	%		%	
At home, who helps you in doing your homework?	yes	no	yes	no
Father	19.3	80.7	15.4	84.6
Mother	18.2	81.8	10.4	89.6
Both (Father & Mother)	12.4	87.6	10.1	89.9
Brother	20.4	79.6	22.3	77.7
Sister	26.3	73.7	22.3	77.7
Tutor	14.0	86.0	13.7	86.3
Any other	4.5	95.5	6.4	93.6
No one	10.7	89.3	23.6	76.4
Total number of respondents	10378		11558	
Percentage of respondents providing a valid response	97.5%		97.9%	

Table 4: Student opinion on mode of transport to school

	Grade 4		Grade 8	
	%		%	
How do you go to school?	yes	no	yes	no
On foot	84.8	15.2	64.9	35.1
On bicycle	6.4	93.6	8.7	91.3
In car	5.6	94.4	13.3	86.7
In school bus	2.7	97.3	4.4	95.6
Through public transport (bus, Rickshaw, van etc.)	5.8	94.2	15.3	84.7
Total number of respondents	10378		11558	
Percentage of respondents providing a valid response	98.0%		97.7%	

Table 5: Student opinion on time taken to get to school

	Grade 4	Grade 8
How much time, you take to reach your school?	%	
< 30 minutes	87.5	81.3
31 minutes – 1 hour	8.8	13.1
> 1 hour	3.7	5.6
Total number of respondents	10378	11558
Percentage of respondents providing a valid response	98.8%	99%

Table 6: Student opinion on distance of home from school

	Grade 4	Grade 8
How much distance is of your school from your home?	%	
< 1km	67.7	55.4
1 km – 2 km	18.1	26.2
2 km – 3 km	7.2	9.2
4km – 5 km	3.4	8.3
>5 km	3.6	0.9
Total number of respondents	10378	11558
Percentage of respondents providing a valid response	98.5%	98.2

Table 7: Students having breakfast before school

		Grade 4	Grade 8
		%	%
Do you take breakfast before going to school?	never	5.9	6.6
	seldom	13.2	24.3
	often	6.8	6.3
	always	74.1	62.9
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		100%	100%

Table 8: Student opinion on punishment or reprimands

		Grade 4	Grade 8
		%	%
Are you physically punished in your school?	never	56.8	52.8
	seldom	35.5	40.9
	often	3.2	2.5
	always	4.4	3.8
punished for being late arriving at school	yes	29.6	33.1
	no	70.4	66.9
punished for incomplete homework	yes	27.8	26.3
	no	72.2	73.7

		Grade 4	Grade 8
punished for violation of discipline	yes	15.8	17.0
	no	84.2	83.0
punished for not learning the assigned lesson	yes	34.7	29.7
	no	65.3	70.3
Have you to face abusing, scolding and threatening in the school?	never	34.2	35.3
	seldom	56.1	58.0
	often	6.0	3.9
	always	3.7	2.8
abuse, scolding or threatened for being late arriving at school	yes	29.8	31.8
	no	70.2	68.2
abuse, scolding or threatened for incomplete homework	yes	29.9	26.9
	no	70.1	73.1
abuse, scolding or threatened for violation of discipline	yes	21.9	20.4
	no	78.1	79.6
abuse, scolding or threatened for not learning the assigned lesson	yes	34.5	27.6
	no	65.5	72.4
Total number of respondents		6832	7481
Percentage of respondents providing a valid response		94.1%	92.1%

Table 9: Student opinion on resources at home

	Grade 4		Grade 8	
	%		%	
Do you use the following things at your home?	yes	no	yes	no
books	86.4	13.6	86.9	13.1
calculator	23.5	76.5	43.9	56.1
computer	19.2	80.8	26.1	73.9
internet	24.4	75.6	41.8	58.2
dictionary / thesaurus	21.6	78.4	37.5	62.5
television	49.3	50.7	53.0	47.0
mobile phone	52.8	47.2	59.7	40.3
tablet	16.9	83.1	18.7	81.3
telephone	21.4	78.6	23.2	76.8
Total number of respondents	10378		11558	
Percentage of respondents providing a valid response	97.6%		98.2%	

Table 10: Student opinion on time spent on devices

		Grade 4	Grade 8
How much time, do you spend using the following things?		%	%
television	not at all	41.0	43.0
	> 1 hour	36.1	35.1

	1 – 2 hours	15.8	15.6
	2 – 3 hours	4.0	3.3
	>3 hours	3.1	3.0
mobile phone	not at all	53.6	44.1
	> 1 hour	32.8	35.8
	1 – 2 hours	8.8	12.8
	2 – 3 hours	2.8	3.7
	>3 hours	2.0	3.7
computer / laptop	not at all	84.0	77.0
	> 1 hour	10.3	13.8
	1 – 2 hours	3.8	6.0
	2 – 3 hours	1.2	1.7
	>3 hours	0.8	1.5
tablet	not at all	87.5	86.3
	> 1 hour	17.1	9.2
	1 – 2 hours	4.6	2.5
	2 – 3 hours	1.5	0.9
	>3 hours	1.6	1.1
video games	not at all	-	76.9
	> 1 hour	-	15.8
	1 – 2 hours	-	4.3
	2 – 3 hours	-	1.4
	>3 hours	-	1.7
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		100%	100%

Table 11: Student opinion on coverage of textbooks for each subject

		Grade 4	Grade 8
		%	%
Urdu	less than half	11.9	-
	half	8.7	-
	more than half	12.2	-
	complete book	67.2	-
English	less than half	13.9	-
	half	10.3	-
	more than half	17.0	-
	complete book	58.7	-
Maths	less than half	12.9	11.4
	half	11.0	9.4
	more than half	15.8	27.6
	complete book	60.3	51.6
Sindhi	less than half	34.3	-
	half	11.3	-
	more than half	9.5	-
	complete book	44.9	-
Science	less than half	-	9.6
	half	-	8.2
	more than half	-	21.4
	complete book	-	60.8

		Grade 4	Grade 8
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		98.7%	100%

Table 12: Student opinion on homework regularity

		Grade 4	Grade 8
		%	%
How much time do you study at your home?	not at all	7.6	7.0
	> 1 hour	29.1	19.4
	1 – 2 hours	37.2	33.5
	2 – 3 hours	15.7	19.8
	>3 hours	10.4	20.3
How often does your teacher assign Urdu homework?	never	8.4	-
	seldom	14.1	-
	often	8.9	-
	always	68.7	-
How often does your teacher assign English homework?	never	14.8	-
	seldom	13.8	-
	often	9.1	-
	always	62.2	-
How often does your teacher assign Maths homework?	never	15.0	8.8
	seldom	11.4	13.7
	often	8.7	10.3
	always	64.9	67.2
How often does your teacher assign Sindhi homework?	never	71.1	
	seldom	6.4	
	often	3.5	
	always	19.1	
How often does your teacher assign Science homework?	never	-	6.6
	seldom	-	20.3
	often	-	13.1
	always	-	60.0
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		100%	100%

Table 13: Student opinion on homework completion

		Grade 4	Grade 8
		%	%
Do you complete your Urdu homework?	never	8.0	-
	seldom	9.3	-
	often	10.5	-
	always	72.2	-
Do you complete your English homework?	never	15.3	-
	seldom	9.1	-
	often	10.9	-
	always	64.7	-

		Grade 4	Grade 8
Do you complete your Maths homework?	never	16.3	8.1
	seldom	8.3	9.3
	often	10.6	9.5
	always	64.8	73.1
Do you complete your Sindhi homework?	never	69.2	-
	seldom	4.3	-
	often	3.2	-
	always	23.3	-
Do you complete your Science homework?	never	-	5.7
	seldom	-	10.2
	often	-	10.0
	always	-	74.1
Do your teachers check your Urdu homework?	never	6.6	-
	seldom	7.1	-
	often	6.3	-
	always	80.1	-
Do your teachers check your English homework?	never	13.4	-
	seldom	6.3	-
	often	6.6	-
	always	73.8	-
Do your teachers check your Maths homework?	never	14.5	8.8
	seldom	6.0	8.9
	often	6.4	7.6
	always	73.1	74.6
Do your teachers check your Sindhi homework?	never	68.4	-
	seldom	3.8	-
	often	2.5	-
	always	25.3	-
Do your teachers check your Science homework?	never	-	6.4
	seldom	-	9.8
	often	-	7.9
	always	-	75.9
Do your teachers identify and correct your mistakes in your Urdu homework?	never	8.8	-
	seldom	12.5	-
	often	9.8	-
	always	68.9	-
Do your teachers identify and correct your mistakes in your English homework?	never	15.7	-
	seldom	10.7	-
	often	9.5	-
	always	64.0	-
Do your teachers identify and correct your mistakes in your Maths homework?	never	16.6	11.6
	seldom	10.4	16.0
	often	8.7	9.6
	always	64.3	62.8
Do your teachers identify and correct your mistakes in your Sindhi homework?	never	69.0	-
	seldom	5.2	-
	often	3.8	-
	always	22.0	-
Do your teachers identify and correct your mistakes in your Science homework?	never	-	8.8
	seldom	-	17.3
	often	-	10.5

		Grade 4	Grade 8
	always	-	63.4
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		100%	100%

Table 14: Student opinion on activities that affect students' studies at home

	Grade 4		Grade 8	
	%		%	
What affects your studies at home?	yes	no	yes	no
Helping parents with household chores	66.3	33.7	63.0	37.0
Bringing grocery from the shops	42.0	58.0	28.5	71.5
Meeting relatives	34.1	65.9	28.5	71.5
Sports	49.6	50.4	35.2	64.8
Parents not helping in studies	27.4	72.6	18.1	81.9
Excessive use of mobile phones	21.5	78.5	22.7	77.3
Due to load shedding	34.8	65.2	37.0	63.0
Closure of school due to emergencies (corona, flood, sit-in, strike, border firing).	27.2	72.8	29.7	70.3
Helping parents in daily routines	28.7	71.3	27.0	73.0
Due to other reasons	19.9	80.1	20.8	79.2
Total number of respondents	10378		11558	
Percentage of respondents providing a valid response	97.6%		96.6%	

Table 15: Student opinion on teacher language for explanations

		Grade 4	Grade 8
To what extent, your teachers use local or mother language to explain the lesson during teaching learning process?		%	%
Urdu	never	13.7	-
	seldom	19.3	-
	often	12.5	-
	always	54.5	-
English	never	26.2	-
	seldom	21.5	-

		Grade 4	Grade 8
	often	15.3	-
	always	37.0	-
Maths	never	31.9	22.3
	seldom	16.9	25.5
	often	13.3	13.6
	always	37.9	38.6
Science	never	-	18.6
	seldom	-	27.2
	often	-	14.0
	always	-	40.1
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		98.7%	100%

Table 16: Students' opinion on their education

		Grade 4	Grade 8
What is your opinion on the following statements about education?		%	%
In class, teachers allow me to ask questions related to the study	never	9.0	9.9
	seldom	13.8	15.3
	often	8.5	8.2
	always	68.6	66.5
Academic discussions are encouraged in class	never	13.8	15.3
	seldom	16.9	19.1
	often	14.4	14.2
	always	54.9	51.4
Academic assessment (tests, assignments etc.) is conducted periodically in the school	never	10.9	11.5
	seldom	19.6	22.2
	often	18.2	18.1
	always	51.3	48.3
The school provides conducive environment for learning	never	15.9	17.8
	seldom	12.9	13.5
	often	10.4	10.5
	always	60.8	58.2
I face unnecessary reprimands in class	never	51.2	59.9
	seldom	23.6	25.7
	often	5.8	5.0
	always	19.4	12.4
Teachers see my homework and give useful guidance	never	13.2	12.7
	seldom	11.2	15.1
	often	11.0	13.9
	always	64.6	58.3
My teachers use blackboard during teaching	never	13.2	11.0
	seldom	11.2	9.0
	often	11.0	9.1
	always	64.6	70.9
We are provided with free textbooks at the school at the beginning of the session	never	10.9	13.8
	seldom	6.8	7.6
	often	7.5	6.4

		Grade 4	Grade 8
	always	74.9	72.2
During the study, if something is not understood, the teacher explains again and again	never	13.1	9.9
	seldom	6.2	7.3
	often	6.1	7.6
	always	74.6	75.2
My parents ask me about my studies in school	never	10.8	11.2
	seldom	6.6	18.6
	often	8.1	12.7
	always	74.6	57.5
I read other informative books or stories along with textbooks	never	13.8	17.2
	seldom	18.1	37.2
	often	12.9	14.7
	always	55.2	30.9
I like the environment of my school	yes	90.8	90.0
	no	9.2	10.0
My peers and I participate in sports at school	yes	83.6	77.4
	no	16.4	22.6
I can express my ideas in class	yes	82.2	78.3
	no	17.8	21.7
I like my class	yes	87.8	89.1
	no	12.2	10.9
I ask the teacher for help about learning related matters during the class	yes	85.9	85.9
	no	14.1	14.1
I can ask questions in class	yes	85.6	88.3
	no	14.4	11.7
Total number of respondents	yes	10378	11558
Percentage of respondents providing a valid response		100%	98%

Table 17: Student opinion on their absenteeism

		Grade 4	Grade 8
		%	%
How often are you absent from school?	not at all	29.5	32.8
	1 – 2 times a month	57.6	56.9
	5 times a month	8.3	6.9
	> 5 times a month	4.5	3.4
Total number of respondents		10378	11558
Percentage of respondents providing a valid response		100%	100%

Table 18: Student opinion on activities that affect students' studies at home

	Grade 4		Grade 8	
	%		%	
What are the reasons of your absence from school?	yes	no	yes	no
Due to illness	81.2	18.8	81.2	18.8
Meeting relatives	22.6	77.4	22.6	77.4

	Grade 4		Grade 8	
Helping parents in work	22.1	77.9	22.1	77.9
Incomplete homework	13.0	87.0	13.0	87.0
Lack of test preparation	14.3	85.7	14.3	85.7
Emergencies (corona, flood, sit-in, border firing)	22.4	77.6	22.4	77.6
Other	13.5	86.5	13.5	86.5
Total number of respondents	10378		11558	
Percentage of respondents providing a valid response	96.8%		94.7	

Parent survey

Table 19: Parental educational qualification

	Grade 4			Grade 8		
What is your educational qualification?	father	mother	guardian	father	mother	guardian
Illiterate/below primary	23.8	26.3	19.0	16.4	22.8	17.1
Primary	22.3	27.2	15.3	13.3	17.8	10.4
Middle	15.1	14.1	19.6	15.4	15.2	12.2
Matriculation (Grade 10)	21.2	17.1	20.2	26.6	22.2	20.4
Intermediate (Grade 12)	8.0	7.4	10.7	13.4	10.9	17.7
Diploma	1.3	1.1	2.2	2.0	1.5	3.0
BA/BSc/AD	3.9	3.3	4.4	6.9	4.8	7.5
MA/MSc/B.S (Hons)	2.7	2.2	4.2	4.1	3.0	7.0
M. Phil / Ph. D.	1.6	1.2	4.4	2.0	1.8	4.6
Total number of respondents	10373			11558		
Percentage of respondents providing a valid response	100%			94.5%		

Table 20: Parental profession

	Grade 4			Grade 8		
What is your profession?	father	mother	guardian	father	mother	guardian
Farmer	21.5	8.4	20.7	16.6	6.9	18.9
Skilled worker	10.3	6.9	9.9	8.8	4.9	9.2
Daily wager	19.1	3.2	15.2	14.1	2.6	11.0
Shopkeeper	12.9	2.6	8.5	13.3	2.2	9.7
Government employee	11.3	3.7	9.3	19.6	6.0	12.6
Private employee	8.7	4.8	8.1	10.9	4.7	9.7
Housewife	4.3	49.9	9.3	4.3	56.0	8.7

	Grade 4			Grade 8		
Unemployed	7.4	17.2	11.6	6.0	13.0	10.0
Other profession	4.6	3.2	7.5	6.2	3.7	10.2
Total number of respondents	10373			11558		
Percentage of respondents providing a valid response	100%			90.9%		

Table 21: Parental opinion on child's homework

		Grade 4	Grade 8
		%	%
Does your child get homework?	do not know	4.5	4.2
	never	3.1	6.3
	seldom	10.8	12.1
	often	10.2	12.4
	always	71.4	67.7
Does your child do homework?	never	5.2	4.9
	seldom	11.1	9.9
	often	14.4	11.4
	always	69.3	73.8
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 22: Parental opinion on who provides help with homework

	Grade 4		Grade 8	
	%		%	
Who helps your child to do his homework?	yes	no	yes	no
Father	16.4	83.6	15.2	84.8
Mother	17.3	82.7	12.7	87.3
Both father and mother	11.8	88.2	9.2	90.8
Brother	21.0	79.0	22.3	77.7
Sister	29.0	71.0	24.5	75.5
Tutor	14.6	85.4	14.8	85.2
Somebody else	5.6	94.4	7.8	92.2
No one	11.9	88.1	25.2	74.8
Total number of respondents	10373		11558	
Percentage of respondents providing a valid response	97.0%		96.7%	

Table 23: Parental opinion on how long their child spend doing homework

		Grade 4	Grade 8
		%	%
How long time does your child take to complete homework?	does not do	4.7	4.7
	> 1 hour	36.2	28.0
	1 – 2 hours	43.2	41.1
	2 – 3 hours	12.0	16.3
	>3 hours	4.0	9.9
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		97.5%	98.1%

Table 24: Parental opinion on the teachers checking homework

		Grade 4	Grade 8
Do the teachers check the homework of your child?		%	%
Urdu homework	do not know	7.2	-
	never	6.8	-
	seldom	12.9	-
	often	12.2	-
	always	60.9	-
English homework	do not know	12.4	-
	never	9.6	-
	seldom	11.1	-
	often	11.6	-
	always	55.3	-
Maths homework	do not know	13.0	11.3
	never	9.8	5.7
	seldom	10.5	7.9
	often	10.7	9.6
	always	56.0	65.4
Sindhi homework	do not know	38.9	-
	never	33.0	-
	seldom	4.3	-
	often	4.1	-
	always	19.8	-
Science homework	do not know	-	8.3
	never	-	7.2
	seldom	-	16.3
	often	-	12.7
	always	-	55.5
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 25: Parental opinion on the teachers identifying and correcting mistakes in homework

		Grade 4	Grade 8
Do the teachers identify and correct the mistakes in the homework of your child?		%	%
Urdu homework	do not know	7.2	-
	never	6.8	-
	seldom	12.9	-
	often	12.2	-
	always	60.9	-
English homework	do not know	12.4	-
	never	9.6	-
	seldom	11.1	-
	often	11.6	-
	always	55.3	-
Maths homework	do not know	13.0	11.2
	never	9.8	7.6
	seldom	10.5	14.7
	often	10.7	11.6
	always	56.0	55.0
Sindhi homework	do not know	38.9	
	never	33.0	
	seldom	4.3	
	often	4.1	
	always	19.8	
Science homework	do not know	-	8.3
	never	-	7.2
	seldom	-	16.3
	often	-	12.7
	always	-	55.5
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 26: Parental opinion on use of the local language by teachers

		Grade 4	Grade 8
To what extent the teachers use local language to teach the following subjects?		%	%
Urdu	do not know	9.5	-
	never	9.7	-
	seldom	16.6	-
	often	12.7	-
	always	51.5	-
English	do not know	18.2	-
	never	15.6	-
	seldom	17.9	-
	often	14.3	-
	always	34.0	-
Maths	do not know	21.9	15.1

		Grade 4	Grade 8
	never	18.0	14.8
	seldom	14.1	21.7
	often	12.1	13.4
	always	33.9	35.1
Science	do not know		12.4
	never		13.5
	seldom		23.6
	often		14.3
	always		36.3
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 27: Parental opinion on the number of books in the home

		Grade 4	Grade 8
		%	%
How many books are there in your house apart from textbooks?	not even 1	43.2	29.7
	1 - 20	48.5	52.3
	21 - 40	3.5	7.8
	41 - 60	1.0	2.6
	> 60	3.8	7.5
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		97.5%	100%

Table 28: Parental opinion of time spent on devices by child

		Grade 4	Grade 8
How much time, do you spend using the following things?		%	%
television	not at all	42.8	45.0
	> 1 hour	31.3	31.0
	1 – 2 hours	18.9	16.9
	2 – 3 hours	4.8	4.3
	>3 hours	2.2	2.8
computer	not at all	82.8	78.5
	> 1 hour	10.9	12.7
	1 – 2 hours	3.9	5.8
	2 – 3 hours	1.3	1.6
	>3 hours	1.0	1.5
mobile phone	not at all	82.2	73.4
	> 1 hour	10.7	16.0
	1 – 2 hours	4.3	6.6
	2 – 3 hours	1.6	2.0
	>3 hours	1.2	1.9
tablet	not at all	90.2	89.0
	> 1 hour	5.9	7.4
	1 – 2 hours	2.2	2.1
	2 – 3 hours	0.7	0.7

		Grade 4	Grade 8
	>3 hours	0.9	0.8
video games	not at all	82.9	83.4
	> 1 hour	11.2	10.8
	1 – 2 hours	3.5	3.7
	2 – 3 hours	1.3	1.1
	>3 hours	1.1	0.9
ipad / ipod	not at all	92.6	91.8
	> 1 hour	4.4	5.0
	1 – 2 hours	1.5	1.8
	2 – 3 hours	0.6	0.5
	>3 hours	0.8	0.9
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 29: Parental opinion on child screen time

		Grade 4	Grade 8
		%	%
Is screen time (TV, video games, cell phones, etc.) is affecting upon your child's sleep?	never	43.8	48.7
	seldom	32.7	34.5
	often	13.6	8.4
	always	9.8	8.6
Is screen time (TV, video games, cell phones, etc.) is affecting upon your child's academic performance?	never	42.5	49.0
	to some extent	31.2	31.4
	often	13.3	10.0
	too much	13.0	9.6
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 30: Parental opinion on child's participation in sport

		Grade 4	Grade 8
		%	%
Does your child participate in physical games/sports?	not at all	21.9	24.4
	to some extent	34.0	35.0
	often	22.1	20.4
	much more	22.1	20.2
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 31: Parental opinions on the performance of the school

	Grade 4	Grade 8
If you are not satisfied with the performance of the school, what is/are the reasons?	%	%

	Grade 4		Grade 8	
	yes	no	yes	no
Shortage of teachers	40.7	59.3	35.1	64.9
Absence of teachers	20.5	79.5	24.4	75.6
Unsatisfactory teaching methods of teachers	17.8	82.2	23.8	76.2
Lack of basic facilities	38.3	61.7	43.1	56.9
Lack of teaching facilities	25.9	74.1	28.5	71.5
Lack of teaching aids	21.4	78.6	24.1	75.9
Corporal punishment is given to the child	16.9	83.1	22.9	77.1
The child is subjected to unnecessary scolding	18.4	81.6	26.6	73.4
Total number of respondents	4174		4205	
Percentage of respondents providing a valid response	83.0%		84.8%	

Table 32: Parental opinions on important factors that lead to school improvement

	Grade 4		Grade 8	
Which of the following factors do you think is most important to improve the school performance?	%		%	
	yes	no	yes	no
The head of the school should be hard-working, sympathetic and spontaneous decisive	61.5	38.5	64.3	35.7
Teachers should be hardworking and kind	51.4	48.6	63.0	37.0
Textbooks should be available for children on time	44.3	55.7	51.0	49.0
Provision of basic teaching facilities should be ensured	42.3	57.7	42.6	57.4
Teachers should be experts in teaching matters	41.3	58.7	40.1	59.9
The authorities should regularly inspect the school	42.8	57.2	39.2	60.8
Parents should be involved in the school activities	32.9	67.1	42.5	57.5
Apart from this, if there is any other suggestion	61.5	38.5	24.2	75.8
Total number of respondents	10373		11558	
Percentage of respondents providing a valid response	92.9%		94.4%	

Table 33: Parental opinion on their child's education

		Grade 4	Grade 8
What do you think of the following statements about your child's education?		%	%
Teachers provide helpful guidance by checking homework	never	14.3	14.8
	seldom	11.8	15.8
	often	15.2	15.2
	always	58.7	54.2
Free textbooks are provided on time	never	20.2	19.9
	seldom	9.0	11.4
	often	8.5	9.8
	always	62.3	58.9

		Grade 4	Grade 8
You contact the school regarding the child's performance	never	20.7	19.2
	seldom	25.9	28.4
	often	15.0	15.1
	always	38.4	37.3
You are satisfied with school's performance in regarding your child's education	never	17.2	15.0
	seldom	8.1	8.6
	often	15.0	12.8
	always	59.6	63.6
Academic assessment (tests, assignments, etc.) of the child is conducted from time to time	never	17.2	14.9
	seldom	15.0	16.3
	often	17.9	17.8
	always	49.9	51.1
You are satisfied with the educational environment of the school	never	18.1	15.7
	seldom	7.3	7.7
	often	12.2	11.0
	always	62.4	65.6
You attend parent teacher meeting (PTM) regularly	never	30.1	32.8
	seldom	19.7	20.4
	often	14.3	12.1
	always	36.0	34.7
Your opinion is given importance in parent teacher meeting (PTM)	never	29.7	32.3
	seldom	12.9	14.0
	often	14.0	12.9
	always	43.4	40.7
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 34: Parental opinion on disability and problems faced

		Grade 4	Grade 8
		%	%
Is there a person in your family affected by any kind of disability?	no one	80.4	81.2
	one	13.5	13.5
	two	3.3	3.0
	more than two	2.8	2.4
How far is the nearest dispensary or hospital from your home?	1km – 2km	53.2	53.4
	2km – 3km	17.9	18.7
	3km – 5km	11.9	11.8
	5km – 10km	8.7	7.6
	>10km	8.3	8.5
Does your child have any type of disability?	physical	6.6	7.2
	mental	5.0	4.3
	other	2.1	1.7
	no disability	86.2	86.8
To what extent does your child face eyesight problems during study?	not at all	80.5	76.1
	to some extent	10.6	14.6
	often	3.5	4.3
	much more	5.4	5.1
	not at all	87.5	87.2

		Grade 4	Grade 8
To what extent does your child face listening problems during study?	to some extent	6.8	7.7
	often	2.8	2.5
	much more	3.0	2.5
To what extent does your child face difficulty walking?	not at all	88.1	87.6
	to some extent	6.1	6.6
	often	2.5	2.6
	much more	3.4	3.2
To what extent does your child face concentration problems during study?	not at all	73.5	76.1
	to some extent	17.2	15.4
	often	4.6	4.7
	much more	4.7	3.9
To what extent does your child face problems relating to remembering lessons during study?	not at all	66.4	68.8
	to some extent	20.8	20.4
	often	6.4	5.5
	much more	6.4	5.3
Total number of respondents		10373	11558
Percentage of respondents providing a valid response		100%	100%

Table 35: Parental opinions on child's social behaviour

	Grade 4		Grade 8	
What is your opinion about your child's social behaviour?	%		%	
	yes	no	yes	no
Our boy/girl participates in sports	75.6	24.4	68.3	31.7
He/she has friends in school	83.7	16.3	85.5	14.5
He/she has friends in the neighborhood	73.0	27.0	61.9	38.1
Participates in academic activities at school	73.4	26.6	73.8	26.2
Participates in conversation with family members	72.2	27.8	74.6	25.4
Likes solitude/loneliness	27.5	72.5	35.1	64.9
Has self confidence	66.9	33.1	75.2	24.8
Is irritated	28.0	72.0	27.6	72.4
Total number of respondents	10373		11558	
Percentage of respondents providing a valid response	96.6%		98.1%	

Table 36: Activities that affect students' studies at home

	Grade 4		Grade 8	
	%		%	
What are the reasons your child does not complete homework?	yes	no	yes	no
Helping parents with household chores	48.5	51.5	50.3	49.7
Bringing grocery from the shops	25.0	75.0	18.2	81.8
Meeting relatives	17.6	82.4	16.5	83.5

	Grade 4		Grade 8	
Sports	37.1	62.9	24.7	75.3
Parents not helping in studies	13.2	86.8	10.7	89.3
Excessive use of mobile phones	14.3	85.7	17.4	82.6
Due to load shedding	24.4	75.6	30.6	69.4
Closure of school due to emergencies (corona, flood, sit-in, strike, border firing).	17.8	82.2	21.6	78.4
Helping parents in daily routines	15.5	84.5	16.3	83.7
Due to other reasons	13.1	86.9	15.4	84.6
Total number of respondents	10373		11558	
Percentage of respondents providing a valid response	94.7%		94.5%	

Table 37: Parents' opinion on what students like

	Grade 4		Grade 8	
	%		%	
What type of activities does your child like?	yes	no	yes	no
Television	33.9	66.1	27.6	72.4
Computer	12.9	87.1	16.9	83.1
Video games	15.8	84.2	13.1	86.9
Physical exercise	15.3	84.7	19.7	80.3
Sports	36.2	63.8	33.9	66.1
Reading books	48.8	51.2	56.3	43.7
Total number of respondents	10373		11558	
Percentage of respondents providing a valid response	97.6%		96.6%	

Appendix 5: Assessment results for each province

This appendix repeats results provided earlier in the main body of the report. The only difference is that the results relating to each individual province have been grouped together. Due to concerns over data quality, results for Grade 4 FL are not split by gender or by whether schools were in an urban or rural location. Furthermore, items in Grade 4 FL were not split into cognitive and content domains.

Assessment results in Azad Jammu and Kashmir

Table 1: Overall student performance in each assessment in AJK

Subject	N students	N items	Mean raw score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	594	48	23.6	49.2	21	10.8	14.8
Grade 4 FL	542	15	12.6	84.1	15	3.9	4.8
Grade 4 Maths	592	48	20.9	43.5	18	9.8	21.5
Grade 4 Urdu and Sindhi	582	52	33.7	64.8	36	10.8	4.6
Grade 8 Maths	791	52	17.0	32.7	16	5.1	22.4
Grade 8 Science	828	52	22.4	43.0	21	7.0	7.3

Table 2: Performance in each assessment in AJK by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	318	276	24.25 (2.13)	22.85 (1.88)	10.98	10.47	-1.41
Grade 4 Maths	48	320	272	22.84 (1.90)	18.54 (1.79)	9.53	9.55	-4.30
Grade 4 Urdu and Sindhi	52	317	265	34.32 (1.95)	32.92 (1.43)	11.51	9.92	-1.39
Grade 8 Maths	52	398	393	17.53 (0.82)	16.44 (0.41)	5.38	4.83	-1.09
Grade 8 Science	52	425	403	22.65 (1.27)	22.03 (0.88)	7.58	6.36	-0.61

Table 3: Performance in each assessment in AJK by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	366	228	23.8 (1.8)	23.3 (2.3)	10.4	11.3	-0.46
Grade 4 Maths	48	363	229	20.6 (1.7)	21.3 (2.4)	8.9	11.0	0.64
Grade 4 Urdu and Sindhi	52	365	217	34.9 (1.3)	31.6 (2.5)	9.3	12.7	-3.26
Grade 8 Maths	52	204	364	18.6 (1.4)	16.6 (0.5)	5.9	4.3	-2.01
Grade 8 Science	52	206	398	24.2 (2.1)	22.8 (1.0)	8.4	6.8	-1.41

Table 4: Percentage of items answered correctly in AJK in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	47.5 (1.7)	Understanding	47.2 (1.7)
	Reading & CTS	50.5 (1.9)	Applying	51.5 (2.0)
Grade 4 Maths	Algebra, Measurement and Geometry	43.3 (3.0)	Knowing	44.1 (3.1)
	Numbers and Operations	45.3 (2.2)	Applying	43.2 (2.1)
	Statistics and Probability	37.4 (5.8)	Reasoning	42.7 (4.8)
Grade 4 Urdu and Sindh	Grammar	49.8 (3.4)	Understanding	66.1 (2.0)
	Reading for information or task	65.7 (2.0)	Applying	61.9 (3.0)
	Reading for literary experiences	71.5 (2.8)		
	Vocabulary	60.9 (4.9)		
Grade 8 Maths	Algebra	26.9 (2.3)	Knowing	33.7 (4.8)
	Measurement and Geometry	33.7 (6.1)	Applying	33.2 (3.4)
	Numbers and Operations	41.5 (4.6)	Reasoning	30.3 (2.9)
	Statistics and Probability	25.2 (3.1)		
Grade 8 Science	Earth and Space Sciences	37.5 (5.9)	Knowing	46.6 (3.7)
	Life Sciences	48.0 (4.0)	Applying	38.7 (4.3)
	Physical Sciences	41.2 (3.6)	Reasoning	44.4 (4.9)

Assessment results in Balochistan

Table 5: Overall student performance in each assessment in Balochistan

Subject	N students	N items	Mean score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	896	48	24.6	51.2	25	11.0	16.4
Grade 4 FL	823	15	13.0	86.9	15	3.7	4.6
Grade 4 Maths	876	48	19.4	40.4	17	8.8	24.0
Grade 4 Urdu and Sindhi	913	52	33.7	64.8	35	12.0	4.6
Grade 8 Maths	877	52	15.9	30.6	15	5.6	36.3
Grade 8 Science	883	52	25.1	48.4	25	8.2	7.4

Table 6: Performance in each assessment in Balochistan by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	436	460	23.52 (1.67)	25.60 (2.15)	9.99	11.75	2.07
Grade 4 Maths	48	448	428	19.61 (1.34)	19.13 (1.76)	8.52	9.15	-0.48
Grade 4 Urdu and Sindhi	52	444	469	32.01 (1.82)	35.27 (2.07)	11.63	12.06	3.26
Grade 8 Maths	52	379	498	16.11 (0.95)	15.72 (0.76)	6.15	5.18	-0.38
Grade 8 Science	52	352	531	24.78 (1.15)	25.38 (1.48)	7.30	8.78	0.60

Table 7: Performance in each assessment in Balochistan by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	311	409	22.5 (2.3)	25.5 (2.3)	10.7	11.6	3.04
Grade 4 Maths	48	277	425	15.3 (1.4)	21.2 (1.6)	7.2	8.5	5.90
Grade 4 Urdu and Sindhi	52	315	419	30.4 (1.9)	35.8 (2.2)	10.5	12.3	5.41
Grade 8 Maths	52	283	521	15.8 (1.2)	15.8 (0.7)	6.5	5.3	-0.08
Grade 8 Science	52	285	526	26.0 (1.9)	24.5 (1.3)	8.6	8.1	-1.51

Table 8: Percentage of items answered correctly in Balochistan in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	49.9 (1.7)	Understanding	49.7 (1.5)
	Reading & CTS	52.2 (1.8)	Applying	53.0 (2.0)
Grade 4 Maths	Algebra, Measurement and Geometry	39.7 (3.5)	Knowing	43.3 (3.7)
	Numbers and Operations	42.3 (2.5)	Applying	38.8 (2.0)
	Statistics and Probability	34.7 (6.0)	Reasoning	37.9 (4.6)
Grade 4 Urdu and Sindh	Grammar	51.9 (1.6)	Understanding	66.4 (1.6)
	Reading for information or task	65.5 (2.0)	Applying	61.2 (3.3)
	Reading for literary experiences	70.3 (2.8)		
	Vocabulary	62.4 (4.5)		
Grade 8 Maths	Algebra	27.7 (2.1)	Knowing	32.2 (3.9)
	Measurement and Geometry	32.7 (6.0)	Applying	29.0 (2.7)
	Numbers and Operations	36.4 (3.6)	Reasoning	30.8 (3.1)
	Statistics and Probability	22.8 (2.3)		
Grade 8 Science	Earth and Space Sciences	42.7 (6.8)	Knowing	50.7 (3.1)
	Life Sciences	57.7 (3.0)	Applying	45.6 (5.3)
	Physical Sciences	43.6 (4.0)	Reasoning	49.2 (5.2)

Assessment results in Gilgit-Baltistan

Table 9: Overall student performance in each assessment in Gilgit-Baltistan

Subject	N students	N items	Mean score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	514	48	18.8	39.2	16	8.8	24.3
Grade 4 FL	455	15	12.3	81.7	15	4.5	7.3
Grade 4 Maths	489	48	18.5	38.6	15	8.8	26.0
Grade 4 Urdu and Sindhi	505	52	29.5	56.7	30	9.4	3.6
Grade 8 Maths	616	52	19.0	36.6	18	6.4	18.0

Grade 8 Science	618	52	22.3	42.9	20	8.6	11.7

Table 10: Performance in each assessment in Gilgit-Baltistan by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	258	256	17.81 (1.89)	19.80 (1.97)	8.68	8.74	1.99
Grade 4 Maths	48	243	246	18.88 (1.71)	18.18 (2.27)	8.26	9.32	-0.70
Grade 4 Urdu and Sindhi	52	250	255	27.49 (1.66)	31.40 (1.70)	9.72	8.69	3.91
Grade 8 Maths	52	330	286	18.40 (1.18)	19.75 (1.46)	5.98	6.73	1.35
Grade 8 Science	52	329	289	20.88 (1.67)	23.95 (1.99)	7.99	8.87	3.08

Table 11: Performance in each assessment in Gilgit-Baltistan by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	218	170	21.1 (2.5)	16.9 (1.3)	10.1	7.1	-4.15
Grade 4 Maths	48	193	170	20.7 (2.2)	18.1 (2.9)	8.9	10.1	-2.63
Grade 4 Urdu and Sindhi	52	210	170	32.4 (2.1)	28.3 (1.3)	9.5	8.6	-4.03
Grade 8 Maths	52	322	180	20.0 (1.5)	18.0 (1.0)	7.2	4.7	-2.07
Grade 8 Science	52	334	173	23.1 (2.2)	21.6 (1.5)	9.8	6.8	-1.49

Table 12: Percentage of items answered correctly in Gilgit-Baltistan in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	37.6 (1.6)	Understanding	36.2 (1.8)
	Reading & CTS	40.4 (2.4)	Applying	42.7 (2.5)
Grade 4 Maths	Algebra, Measurement and Geometry	38.4 (3.8)	Knowing	40.1 (3.6)
	Numbers and Operations	40.2 (2.3)	Applying	35.9 (2.2)
	Statistics and Probability	33.3 (4.8)	Reasoning	40.8 (3.9)
Grade 4 Urdu and Sindh	Grammar	33.6 (3.9)	Understanding	59.1 (2.7)
	Reading for information or task	58.2 (2.8)	Applying	51.1 (4.0)
	Reading for literary experiences	65.6 (3.6)		
	Vocabulary	53.9 (5.8)		
Grade 8 Maths	Algebra	31.1 (2.7)	Knowing	38.7 (4.8)
	Measurement and Geometry	37.3 (6.7)	Applying	35.1 (3.5)
	Numbers and Operations	45.3 (4.4)	Reasoning	36.0 (3.6)
	Statistics and Probability	29.2 (4.1)		
Grade 8 Science	Earth and Space Sciences	37.2 (4.8)	Knowing	44.6 (3.3)
	Life Sciences	49.5 (3.6)	Applying	40.9 (3.7)
	Physical Sciences	40.1 (2.9)	Reasoning	43.5 (4.0)

Assessment results in Islamabad Capital Territory

Table 13: Overall student performance in each assessment in ICT

Subject	N students	N items	Mean score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	489	48	20.5	42.7	18	9.9	21.7
Grade 4 FL	418	15	14.4	95.8	15	1.4	0.0
Grade 4 Maths	487	48	18.5	38.5	16	8.3	23.6
Grade 4 Urdu and Sindhi	496	52	30.6	58.8	33	10.9	6.9
Grade 8 Maths	427	52	19.3	37.1	19	5.6	13.6
Grade 8 Science	464	52	21.4	41.2	21	6.3	8.2

Table 14: Performance in each assessment in ICT by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	260	229	22.03 (2.39)	18.69 (1.79)	10.83	8.34	-3.34
Grade 4 Maths	48	259	228	19.75 (2.30)	17.04 (1.23)	9.54	6.26	-2.72
Grade 4 Urdu and Sindhi	52	265	231	31.23 (2.43)	29.84 (2.08)	11.42	10.20	-1.39
Grade 8 Maths	52	223	204	18.02 (0.48)	20.71 (1.07)	4.83	5.96	2.69
Grade 8 Science	52	217	247	20.14 (0.62)	22.57 (1.16)	5.52	6.72	2.42

Table 15: Performance in each assessment in ICT by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	261	228	21.1 (2.2)	19.7 (2.2)	10.5	9.1	-1.46
Grade 4 Maths	48	261	226	18.3 (2.0)	18.7 (1.9)	8.6	7.8	0.35
Grade 4 Urdu and Sindhi	52	265	231	30.2 (2.3)	31.0 (2.3)	11.0	10.7	0.79
Grade 8 Maths	52	178	229	18.0 (0.5)	20.0 (1.0)	5.0	5.8	1.98
Grade 8 Science	52	195	249	19.7 (0.9)	22.6 (1.0)	5.5	6.7	2.87

Table 16: Percentage of items answered correctly in ICT in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	42.9 (1.9)	Understanding	39.4 (1.6)
	Reading & CTS	42.4 (2.4)	Applying	46.5 (2.7)
Grade 4 Maths	Algebra, Measurement and Geometry	38.2 (4.4)	Knowing	42.1 (3.9)
	Numbers and Operations	40.4 (2.7)	Applying	36.0 (3.0)

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
	Statistics and Probability	32.5 (5.5)	Reasoning	36.6 (4.8)
Grade 4 Urdu and Sindh	Grammar	36.8 (3.9)	Understanding	60.7 (2.5)
	Reading for information or task	61.7 (2.2)	Applying	54.6 (3.9)
	Reading for literary experiences	66.4 (3.5)		
	Vocabulary	53.0 (6.7)		
Grade 8 Maths	Algebra	31.4 (2.6)	Knowing	39.4 (5.2)
	Measurement and Geometry	36.6 (7.4)	Applying	35.7 (3.6)
	Numbers and Operations	45.7 (5.1)	Reasoning	36.3 (4.6)
	Statistics and Probability	31.6 (4.9)		
Grade 8 Science	Earth and Space Sciences	36.2 (6.0)	Knowing	41.2 (3.4)
	Life Sciences	47.3 (3.8)	Applying	40.2 (4.4)
	Physical Sciences	38.6 (3.4)	Reasoning	43.5 (5.0)

Assessment results in Khyber Pakhtunkhwa and NMD

Table 17: Overall student performance in each assessment in KP & NMD

Subject	N students	N items	Mean score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	2910	48	21.3	44.5	18	10.5	21.2
Grade 4 FL	2514	15	13.9	92.4	15	2.3	0.8
Grade 4 Maths	2934	48	19.0	39.5	17	9.0	25.5
Grade 4 Urdu and Sindhi	2926	52	29.6	56.9	29	12.3	9.0
Grade 8 Maths	3048	52	18.6	35.8	18	6.5	20.8
Grade 8 Science	2982	52	22.8	43.9	21	8.6	11.1

Table 18: Performance in each assessment in KP & NMD by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	1477	1433	19.36 (0.93)	23.37 (1.18)	9.27	11.18	4.02
Grade 4 Maths	48	1510	1424	18.62 (0.93)	19.35 (0.87)	9.35	8.48	0.74
Grade 4 Urdu and Sindhi	52	1499	1427	26.54 (1.16)	32.78 (1.03)	12.29	11.36	6.23
Grade 8 Maths	52	1527	1521	18.15 (0.52)	19.04 (0.64)	6.08	6.79	0.89
Grade 8 Science	52	1479	1503	21.27 (0.75)	24.38 (0.85)	8.18	8.80	3.11

Table 19: Performance in each assessment in KP & NMD by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	1263	1153	23.3 (1.3)	19.3 (1.0)	11.3	9.1	-4.10
Grade 4 Maths	48	1268	1170	20.0 (1.0)	18.1 (1.0)	8.8	8.9	-1.93
Grade 4 Urdu and Sindhi	52	1276	1161	31.4 (1.3)	27.7 (1.2)	12.7	11.6	-3.68
Grade 8 Maths	52	1456	1394	18.3 (0.5)	18.4 (0.6)	6.0	6.5	0.13
Grade 8 Science	52	1431	1362	22.2 (0.8)	22.7 (0.8)	8.5	8.4	0.48

Table 20: Percentage of items answered correctly in KP & NMD in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	43.2 (1.3)	Understanding	42.6 (1.3)
	Reading & CTS	45.4 (1.8)	Applying	46.6 (1.9)
Grade 4 Maths	Algebra, Measurement and Geometry	40.0 (3.3)	Knowing	42.0 (3.2)
	Numbers and Operations	41.0 (2.0)	Applying	37.3 (2.0)
	Statistics and Probability	33.2 (4.6)	Reasoning	39.2 (3.7)
Grade 4 Urdu and Sindh	Grammar	45.1 (2.9)	Understanding	58.0 (1.6)
	Reading for information or task	56.7 (1.7)	Applying	54.4 (2.3)
	Reading for literary experiences	62.9 (2.3)		
	Vocabulary	55.3 (2.4)		
Grade 8 Maths	Algebra	31.3 (1.9)	Knowing	37.3 (4.6)
	Measurement and Geometry	37.2 (6.5)	Applying	35.0 (2.9)
	Numbers and Operations	43.6 (3.8)	Reasoning	34.9 (2.9)
	Statistics and Probability	27.6 (3.5)		
Grade 8 Science	Earth and Space Sciences	37.3 (4.8)	Knowing	44.7 (3.1)
	Life Sciences	49.3 (3.3)	Applying	41.9 (3.4)
	Physical Sciences	42.2 (2.8)	Reasoning	46.5 (4.4)

Assessment results in Punjab

Table 21: Overall student performance in each assessment in Punjab

Subject	N students	N items	Mean score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	3567	48	34.1	71.1	37	10.2	5.1
Grade 4 FL	3145	15	14.0	93.5	15	2.7	2.1
Grade 4 Maths	3548	48	30.6	63.7	33	9.5	4.8
Grade 4 Urdu and Sindhi	3567	52	42.2	81.1	45	9.2	0.9
Grade 8 Maths	3662	52	27.5	52.9	26	10.4	6.5
Grade 8 Science	3587	52	33.5	64.3	34	9.8	1.7

Table 22: Performance in each assessment in Punjab by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	1762	1805	32.37 (1.02)	35.81 (0.87)	10.63	9.40	3.44
Grade 4 Maths	48	1776	1772	29.52 (0.94)	31.64 (0.82)	9.98	8.80	2.12
Grade 4 Urdu and Sindhi	52	1754	1813	41.00 (0.82)	43.32 (0.71)	9.74	8.48	2.32
Grade 8 Maths	52	1818	1844	26.16 (0.94)	28.80 (0.99)	10.28	10.43	2.64
Grade 8 Science	52	1761	1826	31.21 (0.83)	35.62 (0.87)	9.58	9.54	4.41

Table 23: Performance in each assessment in Punjab by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	1818	1387	33.5 (1.0)	34.7 (1.1)	10.2	10.1	1.15
Grade 4 Maths	48	1801	1370	30.5 (0.9)	30.6 (1.0)	9.6	9.3	0.02
Grade 4 Urdu and Sindhi	52	1816	1380	41.5 (0.8)	43.0 (0.8)	9.7	8.5	1.43
Grade 8 Maths	52	1659	1663	28.8 (1.1)	26.0 (1.0)	10.8	10.0	-2.77
Grade 8 Science	52	1627	1615	34.3 (0.9)	32.9 (0.9)	9.6	9.9	-1.45

Table 24: Percentage of items answered correctly in Punjab in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	72.1 (2.1)	Understanding	68.2 (2.0)
	Reading & CTS	70.3 (1.8)	Applying	74.5 (1.5)
Grade 4 Maths	Algebra, Measurement and Geometry	61.5 (3.8)	Knowing	63.2 (5.0)
	Numbers and Operations	67.0 (3.3)	Applying	64.6 (2.9)
	Statistics and Probability	56.8 (8.5)	Reasoning	62.9 (4.9)
Grade 4 Urdu and Sindh	Grammar	73.2 (2.6)	Understanding	82.5 (1.3)
	Reading for information or task	81.6 (1.5)	Applying	78.0 (2.2)
	Reading for literary experiences	85.2 (2.1)		
	Vocabulary	77.5 (3.9)		
Grade 8 Maths	Algebra	51.5 (2.0)	Knowing	55.3 (4.4)
	Measurement and Geometry	55.8 (5.7)	Applying	51.1 (3.6)
	Numbers and Operations	58.5 (4.9)	Reasoning	52.4 (3.7)
	Statistics and Probability	42.3 (4.4)		
Grade 8 Science	Earth and Space Sciences	54.0 (5.4)	Knowing	66.1 (3.0)
	Life Sciences	72.4 (2.9)	Applying	62.2 (4.3)
	Physical Sciences	61.9 (3.3)	Reasoning	65.0 (4.7)

Assessment results in Sindh

Table 25: Overall student performance in each assessment in Sindh

Subject	N students	N items	Mean score	Mean % items answered correctly	Median raw score	SD of scores	% achieving no more than 25%
Grade 4 English	1621	48	28.1	58.5	29	11.8	13.1
Grade 4 FL	1457	15	10.8	72.3	13	4.7	13.3
Grade 4 Maths	1652	48	23.7	49.4	23	9.4	11.8
Grade 4 Urdu and Sindhi	1673	52	36.1	69.4	39	11.6	4.3
Grade 8 Maths	2461	52	21.2	40.8	18	9.0	18.3
Grade 8 Science	2437	52	25.6	49.2	24	9.6	8.7

Table 26: Performance in each assessment in Sindh by gender

Subject	N items	N students		Mean raw score (SE)		SD		Difference
		Male	Female	Male	Female	Male	Female	
Grade 4 English	48	875	746	26.51 (1.48)	29.91 (1.35)	12.30	10.90	3.40
Grade 4 Maths	48	879	773	23.28 (1.06)	24.24 (1.21)	8.89	9.93	0.96
Grade 4 Urdu and Sindhi	52	896	777	35.54 (1.27)	36.75 (1.27)	11.70	11.46	1.21
Grade 8 Maths	52	1225	1236	21.26 (0.99)	21.12 (1.03)	8.95	9.11	-0.14
Grade 8 Science	52	1226	1211	24.25 (1.06)	26.91 (0.98)	9.89	9.05	2.66

Table 27: Performance in each assessment in Sindh by whether urban or rural location (if available)

Subject	N items	N		Mean (SE)		SD		Difference
		Rural	Urban	Rural	Urban	Rural	Urban	
Grade 4 English	48	751	658	26.6 (1.5)	30.2 (1.7)	11.9	11.6	3.64
Grade 4 Maths	48	756	672	23.3 (1.0)	24.4 (1.4)	8.8	9.6	1.08
Grade 4 Urdu and Sindhi	52	759	699	34.9 (1.3)	38.3 (1.3)	12.0	10.8	3.34
Grade 8 Maths	52	907	1205	23.1 (1.2)	18.9 (0.8)	9.4	7.6	-4.27
Grade 8 Science	52	906	1206	27.1 (1.3)	24.0 (0.9)	10.1	8.9	-3.09

Table 28: Percentage of items answered correctly in Sindh in each content and cognitive domain

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 English	Formal & Lexical	59.1 (1.5)	Understanding	57.5 (1.5)
	Reading & CTS	58.0 (1.7)	Applying	59.6 (1.8)
Grade 4 Maths	Algebra, Measurement and Geometry	48.9 (4.1)	Knowing	50.5 (4.3)
	Numbers and Operations	51.5 (2.8)	Applying	49.0 (2.9)
	Statistics and Probability	43.1 (8.3)	Reasoning	48.2 (5.2)

Subject	Content Domain	Percentage of possible marks achieved (indicative standard error)	Cognitive domain	Percentage of possible marks achieved (indicative standard error)
Grade 4 Urdu and Sindh	Grammar	61.9 (1.8)	Understanding	71.5 (1.3)
	Reading for information or task	68.8 (2.3)	Applying	64.7 (3.2)
	Reading for literary experiences	74.1 (2.5)		
	Vocabulary	68.0 (3.4)		
Grade 8 Maths	Algebra	38.9 (1.9)	Knowing	41.4 (3.7)
	Measurement and Geometry	39.7 (3.8)	Applying	40.1 (3.5)
	Numbers and Operations	47.7 (4.2)	Reasoning	40.9 (2.7)
	Statistics and Probability	32.7 (4.6)		
Grade 8 Science	Earth and Space Sciences	43.4 (5.6)	Knowing	51.0 (2.4)
	Life Sciences	55.6 (3.0)	Applying	47.7 (3.8)
	Physical Sciences	46.5 (2.9)	Reasoning	48.6 (5.2)

Technical Assistance Provided by

CAMBRIDGE

cambridge.org/partnership

PIE

PAKISTAN INSTITUTE OF EDUCATION