



Volume 23 (2026), pp. 26-54
*American Journal of STEM Education:
Issues and Perspectives*
eISSN 30.3-1190| Print ISSN: 3069-0072
<https://doi.org/10.32674/k72y8m35>

Factors Predicting Students' Science Performance: Insights from a Nationwide Study from Pakistan

Sadia Muzaffar Bhutta
*Aga Khan University,
Institute for Educational Development, Pakistan*
<https://orcid.org/0000-0003-0343-8932>

Nusrat Fatima Rizvi
*Aga Khan University,
Institute for Educational Development, Pakistan*

Sohail Ahmad
*Faculty of Education,
University of Cambridge, England*
<https://orcid.org/0000-0001-7344-0566>

Anil Khamis
*Aga Khan University,
Institute for Educational Development East Africa & Institute for Human
Development, Pakistan*

Aisha Naz Ansari
*School of Education,
Durham University, England*
<https://orcid.org/0000-0003-0938-8980>

ABSTRACT

This large-scale study provides evidence on pedagogical quality contributing to students' science performance when examined alongside teacher, classroom, student, family, school, and regional factors. The sample included 286 science

teachers and 15,287 students from 153 public and private schools across six regions of Pakistan. Using validated classroom observation scales and achievement tests for Grades 5, 6, and 8, the study found that pedagogical quality is a significant predictor of science achievement. Girls, younger students, those in Punjab, private-school students, and students with more educated parents performed better. Professional qualifications showed a negative association, while consistent language use across textbooks, instruction, and assessment predicted higher outcomes. Implications for improving science teaching in the Global South are discussed.

Keywords: Elementary school, large-scale study, predicting factors, science, students' learning outcomes.

© Author(s), 2026. Published by Star Scholars Press.

This article is distributed under the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. <https://creativecommons.org/licenses/by/4.0/>

INTRODUCTION

Students' science learning is shaped by a constellation of influences spanning the classroom, school, and family environment. Within the domain of school-controlled factors, teachers' pedagogical practices consistently emerge as among the strongest predictors of student learning (Hamilton & Hattie, 2022). The instructional approaches teachers use determine the kinds of learning opportunities students encounter, making pedagogical quality central to efforts to improve science achievement. Given this centrality, researchers and policymakers increasingly call for rigorous, context-specific evidence on how pedagogical quality contributes to learning, particularly in systems where such evidence remains limited.

In Pakistan, recent large-scale assessments have documented persistently low levels of science achievement (NEAS, 2016; TIMSS, 2020). While these studies provide important descriptive insights into student performance, they do not examine the contribution of observed pedagogical quality to students' science learning. This constitutes a critical gap because, without understanding how teaching quality interacts with other influential factors, such as student background, classroom composition, school characteristics, and regional disparities, policy responses may remain fragmented and weakly targeted.

Moreover, the relationship between pedagogical quality and learning is unlikely to be linear or isolated. Although pedagogical quality may explain a meaningful share of variance in achievement, other factor baskets, including teacher characteristics, school type, parental education, and regional context, also shape learning outcomes. A comprehensive model is, therefore, needed to estimate the unique contribution of pedagogical quality while accounting for these additional influences.

This nationwide study addresses this gap by examining whether elementary science teachers' pedagogical practices predict students' science achievement after controlling for relevant school-, teacher-, classroom-, and student-level characteristics. Particularly, the study seeks to determine the extent to which teachers' pedagogical practices are associated with students' science achievement in Pakistan after controlling for selected school-level (e.g., school system and region), teacher-level (e.g., qualifications, age, and gender), classroom-level (e.g., grade level and classroom gender composition), and student-level (e.g., gender and parental education) variables.

LITERATURE REVIEW

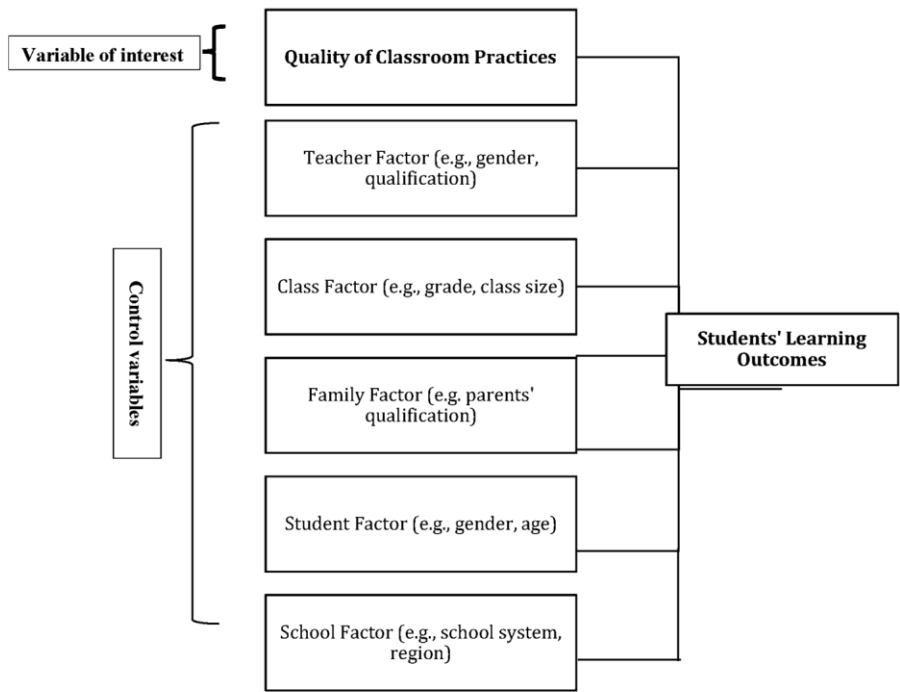
Large-scale assessment research has long attempted to explain variability in student achievement by drawing on school effectiveness frameworks (Creemers, 1994; Kyriakides et al., 2013; Nilsen et al., 2016). These models consistently highlight pedagogical quality as a central driver of learning, alongside student-, teacher-, classroom-, school-, and home-level conditions, as illustrated in Figure 1. Yet, despite this consensus, most evidence, particularly from developing contexts, relies on perceived rather than observed classroom practices. As noted in the literature, "the quality of teaching is largely determined by the pedagogical approaches that teachers use" (Hamilton & Hattie, 2022, p. 140), yet the field still lacks robust, large-scale studies that capture these approaches as they unfold in real classrooms.

This gap is especially pronounced in LMICs, where classroom processes are rarely documented through systematic observation. Existing large-scale studies overwhelmingly depend on teacher self-reports or student perceptions, despite well-established concerns about their validity (Croninger et al., 2012; Ahmad et al., 2022). Observational data, by contrast, provide access to the enacted pedagogical environment, including interactional patterns and instructional sequencing, features central to science learning but often invisible in survey-based studies.

The present study responds directly to this gap by examining the contribution of observed pedagogical quality to students' science performance while controlling for theoretically grounded baskets of variables. The following subsections therefore do not function as isolated summaries; rather, they justify

why each variable group is necessary to isolate the unique effect of pedagogy in the Pakistani context.

Figure 1
Conceptual Framework of the Study



Pedagogical Quality

Pedagogical quality is widely recognised as a central determinant of students’ learning opportunities because it reflects how teachers design, sequence, and enact instruction in real time (Fischer et al., 2020). High-quality science instruction typically involves guided collaborative learning, purposeful questioning, active student engagement, and coherent lesson progression, features consistently associated with improved learning outcomes across contexts (Blömeke et al., 2016; Hamilton & Hattie, 2022). In science specifically, participatory pedagogies enable students to make sense of abstract concepts, build conceptual links, and retain information more effectively (Cairns, 2019; Holzberger et al., 2020).

Despite this consensus, most empirical studies linking pedagogical quality to achievement rely on perceived rather than observed teaching (Fischer et al., 2020; Zhang et al., 2022). This reliance is especially pronounced in LMICs, where

large-scale studies rarely incorporate systematic classroom observation. Yet observational data are essential because they capture the enacted instructional environment, including teacher–student and peer interactions, that cannot be reliably inferred from self-reports (Croninger et al., 2012; van de Grift, 2014). As noted in methodological critiques, independent observation helps “unveil what is happening inside the classroom” beyond the limitations of perceived information (Ahmad et al., 2022; Fraenkel et al., 2015).

This study therefore addresses a critical gap by examining observed pedagogical quality in elementary science classrooms and its association with students’ learning outcomes. To isolate its unique contribution, the subsequent sections outline the theoretically grounded teacher, classroom, student, home, and school factors that must be accounted for when modelling variability in science achievement.

A teacher has a pivotal role in enhancing students’ performance as pedagogical strategies used in the classroom shape learning opportunities for students. The teacher takes decisions about the selection of teaching techniques, designing activities, developing materials, time allocation for each activity, and involving students in discussion through various approaches, which has impact on students’ performance (Fischer et al., 2020). While synthesising systematic reviews on students’ learning in Lower-and-Middle Income Countries (LMICs), Hamilton and Hattie, (2022) reported that “the quality of teaching is largely determined by the pedagogical approaches that teachers use” (p.140) and how it situates in the context and subject. Effective pedagogical quality in general and particularly in science encompasses collaborative learning in a guided manner, questioning strategies, active involvement of students, effective teacher-students and students-students interaction and appropriate sequencing in the delivery of lessons, which positively affect students’ learning (Blomeke et al., 2016). Regardless of discipline, several high-impact teaching strategies mentioned above, were found to be significantly contributing to students’ performance (Hamilton & Hattie, 2022). Arguably, participatory pedagogies play a fundamentally important role in teaching science for students to not only grasp the abstract concepts but retain the information and make logical links within and across content, hence augmenting

Teacher Factors

Teacher characteristics shape both the quality of instruction and the learning opportunities available to students, making them essential controls when estimating the unique effect of pedagogical quality. These characteristics are commonly grouped into acquired attributes, such as academic qualification, professional preparation, and teaching experience, and sociodemographic attributes, including gender (Coenen et al., 2017).

Academic and professional qualifications influence teachers' content knowledge and pedagogical repertoire, although findings remain mixed. While several studies report positive associations between academic qualification and student achievement (Gess-Newsome et al., 2019; Oppermann et al., 2021), others find no significant effect (Coenen et al., 2017; Zhang et al., 2022). Professional preparation and ongoing development can enhance teachers' capacity to engage students and implement effective instructional strategies (Bhai & Horoi, 2019; Li et al., 2021), yet evidence from some contexts shows limited or no direct relationship with achievement (Jacob et al., 2017; Teodorović et al., 2021). These inconsistencies underscore the need to control for qualification-related factors when examining the contribution of observed classroom practice.

Teaching experience is another widely studied predictor. Early-career experience is often associated with substantial gains in teacher effectiveness, but this effect tends to plateau after the first decade (Hsu et al., 2007). Explanations include the need for continuous updating of subject knowledge and pedagogical skills in rapidly evolving fields (Liu et al., 2010) and the potential stagnation associated with routine teaching (Karakose et al., 2023). Given its non-linear relationship with achievement, experience must be accounted for to avoid confounding its effects with those of pedagogical quality.

Teacher gender, a key sociodemographic characteristic, has produced inconsistent findings globally (Li et al., 2021; Larry & Wendt, 2021). However, studies in several developing contexts, including Pakistan, indicate a performance advantage for students taught by female teachers (Joshi et al., 2022; Lee et al., 2018). Including gender in the analytical model is therefore necessary to ensure that any observed differences in student outcomes are not misattributed to pedagogical quality.

In summary, these teacher-level variables are included not as competing explanations but as necessary controls that allow this study to more precisely estimate the unique contribution of observed pedagogical quality to students' science achievement.

Classroom Factors

Classroom-level conditions shape both the opportunities teachers can create and the ways students engage with science learning. In this study, two variables, medium of instruction and grade level, are included as essential controls to ensure that the estimated effect of observed pedagogical quality is not confounded by structural features of the learning environment.

The medium of instruction is widely recognised as a key quality indicator, though evidence remains mixed. A consistent theme across research is that teaching in a familiar or first language enhances classroom interaction, supports higher-order questioning, and improves conceptual understanding in science (Fung,

2020; Prinsloo & Harvey, 2018; Pun & Macaro, 2019). Students in classrooms where teachers use the local language often demonstrate stronger performance (Fernández-Sanjurjo et al., 2019; NEAS, 2016). However, inconsistency between the language of instruction and the language of assessment can hinder students' ability to interpret test items and construct appropriate responses (Mose, 2019). Given these dynamics, medium of instruction must be accounted for when examining the contribution of classroom interaction patterns to learning outcomes.

Grade level is another important classroom variable. Although achievement is often assumed to improve as students' progress through school, numerous studies, including those from Pakistan, report a decline in science performance in higher grades (Ainley et al., 2020; NEAS, 2016; Rosario et al., 2013). This pattern is frequently attributed to decreasing motivation, self-efficacy, and interest in science as students grow older (Potvin & Hasni, 2014; Toma et al., 2019). Because these motivational shifts can influence how students respond to instructional practices, controlling for grade level is necessary to avoid misattributing grade-related differences to pedagogical quality.

In summary, these classroom factors are included not as competing explanations but as contextual conditions that allow the study to more precisely estimate the unique effect of observed pedagogy on students' science achievement.

Student Factors

Student characteristics shape how learners engage with instruction and respond to pedagogical practices, making them essential controls when estimating the unique effect of observed pedagogical quality. In this study, two widely examined predictors, gender and age, are included to account for baseline differences in learning outcomes.

Student gender has long been associated with variation in classroom participation and academic performance, yet findings remain inconsistent. Some studies report an advantage for girls (Shah et al., 2020), others for boys (Lau & Lam, 2017), while several identify no meaningful differences (Sarouphim & Chartouny, 2016; Tsai & Yang, 2015). Given this variability, gender must be controlled to ensure that any observed differences in science achievement are not misattributed to pedagogical quality.

Age is another important predictor, with large-scale assessments, including those in Pakistan, consistently documenting an inverse relationship between age and science performance (NEAS, 2016; TIMSS, 2020). Although counterintuitive, this pattern is frequently explained by declining motivation, interest, and self-efficacy as students grow older (Toma et al., 2019). These attitudinal shifts are strongly linked to achievement (Wang & Liou, 2017; Zhang & Bae, 2020), suggesting that age operates as a proxy for underlying motivational processes that influence how students respond to instruction.

By including gender and age as student-level controls, the study ensures that the estimated contribution of pedagogical practices reflects differences in teaching rather than pre-existing differences among learners.

Home Factors

Home-level conditions play a significant role in shaping students' learning trajectories, and controlling for them is essential when estimating the unique effect of observed pedagogical quality. Among the various home-related predictors, such as socio-economic status, parental involvement, and parental occupation, parental education consistently emerges as the strongest and most stable indicator of student achievement (Fischer et al., 2020; O'Connell, 2019; Wilberg & Rolfsman, 2019). For this reason, it is included as the key home-level variable in this study.

The positive association between parental education and student performance is often explained through the quality and nature of informed parental involvement. More educated parents tend to hold positive attitudes toward schooling, maintain higher expectations, and engage more proactively with their children's learning (Jeynes, 2015; Xiang & Chiu, 2021). They are also more likely to monitor school progress, understand curricular demands, and provide academic support when needed (Symeou, 2007; Mahuro & Hungi, 2016). These behaviours create a home environment that reinforces learning and shapes how students respond to classroom instruction.

By accounting for parental education, the study ensures that differences in science achievement are not confounded by variations in home support, thereby allowing a more precise estimation of the contribution of pedagogical practices enacted inside the classroom.

School Factors

School-level conditions shape the organisational environment in which teaching and learning occur, making them essential controls when estimating the unique effect of observed pedagogical quality. Two structural features, school system (public vs. private) and school location (urban vs. rural), have been widely examined in relation to student achievement.

A substantial body of research shows that students in private schools often outperform their public-school counterparts across core subjects, including science (ASER, 2020; Baum & Riley, 2019; NEAS, 2016; Smith-Wooley et al., 2018). Explanations typically point to differences in teacher accountability, instructional support, resource availability, and school-level management practices (Akhter, 2017; Thapa, 2015). These patterns are also reflected in parental preferences: families who can afford private schooling frequently choose it due to perceptions

of higher instructional quality and stronger academic support (Ahmed et al., 2013; Andrabi et al., 2008).

Public schools, by contrast, often face systemic challenges, including limited facilities and weaker accountability structures, that can constrain teaching quality and student learning (Awan & Zia, 2015). Similarly, urban-rural disparities in infrastructure, teacher supply, and learning resources can influence achievement patterns independently of classroom instruction.

By controlling for school system and location, the study ensures that differences in science performance are not confounded by broader institutional conditions, thereby allowing a more precise estimation of the contribution of pedagogical practices enacted inside classrooms.

STUDY CONTEXT

Pakistan is among the countries that constitutionally own the responsibility of providing free and compulsory education. The article 25-A of the national constitution 1973 states: “The State shall provide free and compulsory education to all children of the age of five to sixteen years in such manner as may be determined by law”. Compulsory education is divided into three broad levels (pre-primary, elementary, and secondary & higher secondary), whereas the elementary level, which is the focus of this study, comprises primary and middle grades (Grade I to VIII). Furthermore, the government and private sectors are the main providers of education across the country, which have distinct teachers’ recruitment policies. In the government sector, the eligibility criterion for recruiting elementary school teachers is a bachelor’s degree (14 years of education) in any subject. Professional qualification, on the other hand, has witnessed various reforms in the last two decades. One of these reforms replaced traditional certifications of teacher education with a more comprehensive degree programme of a longer duration. Nevertheless, the current teaching workforce in the government sector is dominated by the graduates of traditional certification. Contrarily, recruitment guidelines in the private sector vary across schools because of their independent status. In general, ‘low-fee schools’ recruit teachers with lower qualifications for lower wages, while elite schools prefer to appoint qualified teachers with master’s degrees. However, professional qualification is not compulsory to recruit teachers in private schools. In the context of this study, science teachers at elementary grades are those who have been assigned the responsibility of teaching science regardless of their academic background in the subject.

As far as the teaching of science is concerned, regardless of the school system, it is taught as a separate and core subject until elementary grades. The science textbooks for these grades are developed in light of standards determined by the national curriculum, which aims to develop scientifically literate individuals by engaging them in inquiry-based and student-centred learning (Government of

Pakistan, 2006). Almost all private schools use textbooks which are printed in English language. In government schools, on the flip side, there is no consistency in the language of textbooks; for example, in some regions, government schools use science textbooks which are published in English, whereas in some geographical areas, national (i.e., Urdu) or regional (e.g., Sindhi) languages are used as the medium of textbooks. The four content strands (Biology, Chemistry, Physics, and Earth Sciences) are constantly represented in all textbooks.

RESEARCH METHOD

This paper is extracted from a nationwide survey (2018-2022) of pedagogical practices of elementary (Grade 5, 6 & 8) science and mathematics teachers and their students' achievement in these subjects across Pakistan. The three grades were strategically selected to capture a holistic picture of teaching and learning in the country. For example, grade 5 and 8 are the exit points of primary and elementary schooling, respectively. Whereas grade 6 is the entrance point of elementary grades where many students from feeding primary schools move to middle or high schools. Therefore, it was important to consider such variation in grades in order to have a more holistic picture of elementary grades. The focus of this paper is only on science, with an aim to answer the posed research question cited earlier.

Sample

A multi-stage stratified sampling technique was used to recruit 286 science teachers and their 15,287 students enrolled in three target grades from 153 schools of 25 randomly selected districts across six regions of Pakistan (i.e., Punjab, Sindh, Khyber Pakhtunkhwa, Balochistan, Gilgit Baltistan and Azad Jammu & Kashmir). Importantly, proportionate representation based on the population of each region was ensured in the sample selection. Furthermore, the sample represents government (51%) and private (49%) schools almost equally.

Of the 286 teachers, higher representation (56%) from government schools is a manifestation of a larger workforce serving in the government sector. Moreover, females were overrepresented (65%) in the sample, which reflects gender proportion in the teaching workforce in Pakistan (NEMIS, 2018). On average, sample teachers have working experience of 11 years ($SD=9.4$), ranging from 0.5 to 35 years. In terms of academic qualification, nearly three-quarters of the teachers reported to have attained a master's degree or above. It is important to note that the academic qualification shows highest level of degree attained by the teachers, regardless of the subject. Similarly, more than eighty percent of the teachers have attained professional qualifications, including certificate, B.Ed, and M.Ed.

Of the 15,287 students, government schools represent a higher (61%) proportion mirrors the enrolment patterns in the population (Pakistan Education Statistic, 2018). The sample students have almost equal representation across grades. A little more than half (52%) of the students were boys. The mean age of students for the whole sample was 12.8 years ($SD=1.7$). Moreover, a relatively higher proportion of fathers (20%) reported to have attained a higher qualification (bachelor's and/or master's) than mothers (13%).

Data Collection Instruments

Data were collected from two sources, teachers, and their students in target grades using two types of instruments: one designed specifically for classroom observation of teachers and the other intended to assess students' achievement in science. Additionally, demographic information was gathered by adding demographic sections in the classroom observation scale (e.g., teachers' qualifications, class size, lesson duration) and students' achievement tests (e.g., gender, age, parents' qualifications).

Classroom Observation Scale (COS)

In order to assess the pedagogical quality of selected science teachers, their science lessons were observed using a contextually valid and reliable 3-point COS. The rubric refers to three levels of quality including inadequate (1); adequate (2); and good (3) as represented in Appendix A. Inadequate (1) describes classroom practice that does not meet the basic criteria of provision and teaching/learning practices (e.g., unavailability of basic material; no efforts for the active involvement of children; harsh discipline) in an elementary classroom. The adequate (2) ratings usually focus on the provision of some basic material; some use of active methods to involve children; and some instances of high-order questions. The good (3) ratings required a supportive environment for interactions; a range of teaching strategies to maximise children's involvement in active thinking; and group work to enhance collaborative learning.

The classroom observation scale comprised 15 items in a rubric format. Of the 15 items, thirteen items assess various aspects of classroom quality including physical set-up (e.g., indoor space, classroom display), classroom interaction (e.g., teacher-child interactions, supervision), and teaching and learning methods (e.g., peer learning/group work, active involvement of students). These quality indicators were used to gauge the extent to which students and teachers participate together in the teaching-learning process. Whereas the remaining two items in the COS were defined as checklists in order to record a variety of active methods (e.g., questioning, demonstration, role-plays) and classroom management strategies (e.g., whole class, group/pair work). Additionally, COS included a section to

capture demographic information. In the present study, Cronbach's alpha coefficient for COS was found to be within an acceptable range (0.79).

Science Achievement Tests (SATs)

Contextually valid and reliable Science Achievement Tests (SATs) were adapted for assessing students' performance. The weightage of items in terms of content (i.e., Biology, Physics, Chemistry, and Earth sciences) and cognition levels (i.e., knowledge, application, and reasoning) was aligned with National Curriculum 2006. The test items included Multiple Choice Questions (MCQs) and Constructed Response Questions (CRQs). All SATs consisted of 30 questions except grade five which had 20 items. Considering the variety of languages of textbooks in sample schools, three variants of SATs were administered (i.e., English, Urdu, Sindhi). The questions in the SATs represent a broad spectrum of difficulty ranging from 'very difficult' to 'very easy'. The Cronbach's Alpha, computed for the three SATs, indicated acceptable internal consistency among items (G5=0.60; G6=0.82; G8=0.69).

Data Collection Procedure

The study commenced in the mid of 2018 and was concluded in 2022. The fieldwork for this study was carried out between October 2018 and January 2020, following all ethical considerations committed to the university's Ethical Review Committee (4200-IED-ERC-16). In order to maintain consistency in implementing the COS, two research associates were trained to observe science teaching across the country. These two research associates developed inter-rater reliability with the first of author and reached an acceptable threshold of 80% before they make independent observations (Graham et al., 2012; Reynolds et al., 2009). Each participant teacher was observed once while teaching a science lesson and scored based on COS items. Formal observations of each lesson lasted for 35 to 40 minutes, depending on the individual schedule of the schools. The COS rubric was used for scoring items. It provides space to assign a numeric score for each item. The keywords (e.g., floor swept, lesson started with a quiz, monitoring of all groups, interacting with a few children, appreciate children's work, children's questions not invited) were written next to each item in the comment box for any given score to be fully explained. SATs were administered by fifty research assistants, hired locally in each district. The research assistants were oriented to administer SATs in a standardised manner under the supervision of research associates. Importantly, the lead researchers visited various regions randomly to monitor the data collection process.

Analysis

The data obtained in this study were analysed using SPSS V.25. Students' responses in each test were re-coded in a way that each correct response received one mark, and each incorrect response was coded as zero. Using this coding scheme, all SATs, save grade 5, carried 30 marks. The maximum score assigned to the Grade 5 SAT was 20. The total raw scores were converted to mean percentage to equate scores across grades. Similarly, the overall mean score was computed for COS in order to present pedagogical quality in science classrooms across Pakistan. The mean percentage SATs score, and mean COS score were used as an outcome and the main explanatory variable in the multivariate model, respectively.

After an initial exploration of teachers' and students' data, Multiple Linear Regression (MLR) was employed to answer the research question, which refers to one outcome (i.e., students' SAT score) and multiple explanatory (e.g., pedagogical quality, teachers' qualification, school system, class level, geographical regions) variables or factors (Field, 2013). Bivariate correlation was an initial step to explore associations between outcome variables and potential factors targeted in this study. After exploring the association through bivariate analysis six MLR models were developed separately for six baskets of factors (i.e., teacher, class, students, parents, school, and region). Those variables that appeared significant in the bivariate correlation were added to the six multivariate models. Score on pedagogical quality remained constant for each model. Finally, a parsimonious regression model was built using those variables which retained significance in the six individual models. The quality of classroom pedagogy was entered in the parsimonious model as a main variable to explore its contribution to students' achievement, while controlling for other baskets of factors. Residuals for the model were examined for the assumptions of normality, linearity, and homoscedasticity in order to reduce the chances of drawing biased conclusions from the results (Allison, 1999; Field, 2017; Tabachnick & Fidel, 2019). Further, multicollinearity was also assessed using values of Variance Inflation Factor (VIF) and Tolerance (Allison, 1999; Field, 2017). All the assumptions were met. This paper presents results of the parsimonious model.

RESULTS

The section is organised in two sub-sections to present results of descriptive analysis followed by findings of multivariate modelling.

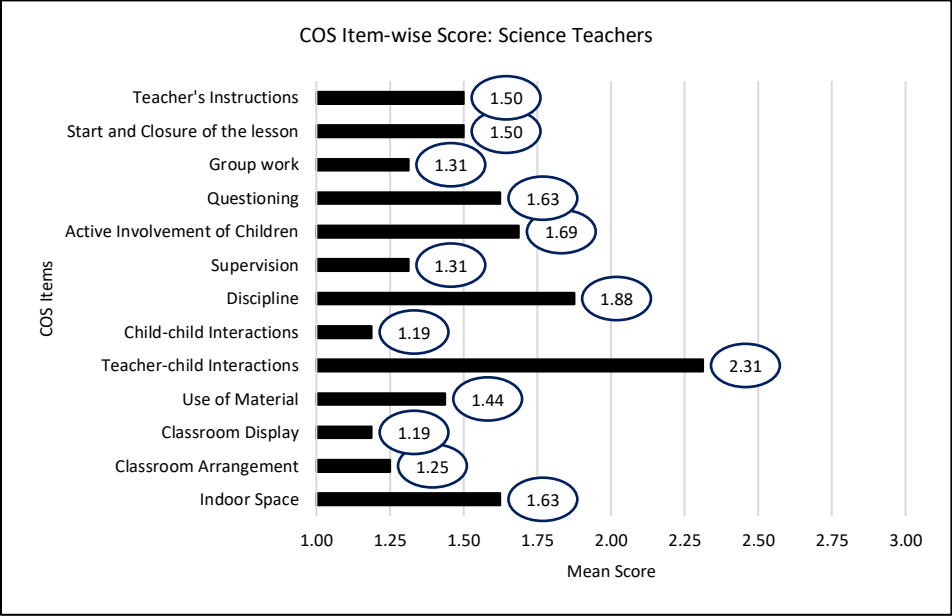
Descriptive Analyses of Students' Learning Outcomes and Pedagogical Quality

Students' performance in science was low nationwide, with an average score of 34% (SD = 16) on the SAT, indicating limited mastery of curriculum-aligned concepts. Performance varied across content domains: students showed comparatively stronger understanding in biology (M = 40, SD = 21), followed by physics (M = 33, SD = 19) and earth science (M = 29, SD = 33), while chemistry emerged as the weakest area (M = 26, SD = 22). A similar pattern appeared across cognitive domains. Students performed best on knowledge-level items (M = 38, SD = 19), with performance declining on application (M = 35, SD = 19) and reasoning tasks (M = 26, SD = 21). These patterns suggest that while basic recall is relatively secure, students struggle to apply or extend scientific ideas, an important insight for instructional design and curriculum alignment.

Classroom observations revealed consistently low pedagogical quality (M = 1.28, SD = 0.24 on a 3-point scale), pointing to a system-wide pattern rather than isolated weaknesses, as illustrated in figure 2. Although teachers generally maintained supportive interactions and avoided harsh discipline, the instructional core was characterised by teacher-centred, whole-class delivery. Opportunities for active learning, peer interaction, problem-solving, or student-initiated inquiry were rare. Material use was minimal, and feedback tended to be confirmatory rather than constructive.

Rather than variation across individual indicators, the broader pattern reflects a narrow instructional repertoire dominated by lecture-style teaching with limited engagement of students in meaning-making processes. This systemic instructional profile provides essential context for interpreting the multivariate results, particularly the contribution of observed pedagogical quality to students' science achievement.

Figure 2
Item-Wise Analysis of the Observational Scale



Predictors of Student Performance

Table 1 presents the results of a parsimonious model developed to explore factors (i.e., teacher, class, parents, students, school, and region) that predict students’ performance in science, over and above the pedagogical quality of the classroom.

Table 1
Factors Predicting Students’ Learning Outcomes in Science

Predictors		Regression coefficient	Beta coefficient
Pedagogical quality	(Constant)	20.921	-
	Classroom practice score (3-point scale)	4.918	0.066***
Teacher	Gender (Male=1; Female=2)	0.600	0.019
	Academic qualification (Bachelor =1; Masters=2)	-0.422	-0.012
	Professional qualification (No=0; Yes =1)	-3.204	-0.084***

Predictors		Regression coefficient	Beta coefficient
Class	Teaching experience (in years)	-0.116	-0.069***
	Instructional time (in minutes)	0.028	0.015
	Consistency in language (Inconsistent=0; Consistent=1)	5.502	0.177***
	Grade 6 [^]	3.909	0.123***
	Grade 8 [^]	-6.017	-0.179***
Parents	Father Education ^{^^} (High School)	0.288	0.009
	Father Education (<i>Bachelor</i>)	1.566	0.031**
	Father Education (<i>Master</i>)	1.963	0.037**
	Mother Education ^{^^} (High School)	-0.128	-0.004
	Mother Education (<i>Bachelor</i>)	4.547	0.070***
Student	Mother Education (<i>Master</i>)	6.728	0.099***
	Student Gender (Boy=1; Girls=2)	1.355	0.044**
	Student Age	-0.671	-0.072***
School	School Type (Public=1; Private=2)	7.131	0.217***
Region	Punjab ^{^^^}	8.576	0.266***
	Sindh	0.059	0.002
	Khyber Pakhtunkhwa	1.029	0.018
	Gilgit Baltistan	2.146	0.033**
	Azad Jammu & Kashmir	-1.759	-0.026*
Adjusted R ²		0.315	
Analysis of Variance		134.34***	
Reference Category: [^] Grade 5; ^{^^} No or basic schooling; ^{^^^} Balochistan			
*p<0.05; **p<0.01; ***p<0.001			

Overall, the model explained 31.5% of the variance ($F=134.34$; $R^2 = 0.315$) in the total mean science scores. The multivariate model showed that the quality of classroom practice has positively contributed ($\beta=0.066$; $p<0.001$) to predicting students' performance, over and above other covariates. More precisely, students would learn better in classrooms where teachers play an active role in: delivering the content using appropriate resources; encouraging them to take the initiative either individually or collaboratively; looking for opportunities to scaffold students' learning by inviting their ideas; and appreciating their efforts.

Furthermore, of the four predictors entered in the basket of teacher factors, only professional qualification ($\beta= -0.084$; $p<0.001$) and teaching experience ($\beta= -0.069$; $p<0.001$) have contributed significantly to students' performance. The

negative coefficient for both variables indicates that those teachers who have no professional qualification and are relatively ‘younger’ in the profession tend to yield better learning outcomes.

Of the three variables added in the basket of class factors only two have contributed in the model. These include consistency in the use of language and grade level. The variable “consistency in the use of language” explains synchronisation in the use of language across textbook, classroom teaching, and assessment and it has contributed positively to students’ performance ($\beta = 0.177$; $p < 0.001$). This intriguing finding unveiled that a consistent use of the ‘same language’ (regardless of native or second language) in three media would lead to better learning outcomes. Interestingly, students in grade 6 ($\beta = 0.123$; $p < 0.001$) have demonstrated better performance as compared to grade 5, while the performance declines as the students progressed to grade 8 ($\beta = -0.179$; $p < 0.001$).

Results of the variables in the basket of parents’ qualification revealed that parents’ education has contributed positively to their children’s performance. Evidence exhibited an upward trend which is more visible beyond secondary school education of parents. Obviously, there are considerably higher gains in children’s scores for those parents who have attained a university-level qualification. Mothers’ education at bachelor ($\beta = 0.070$; $p < 0.001$) and master ($\beta = 0.099$; $p < 0.001$) levels of qualification seems to have made more contribution as compared to fathers (Bachelor, $\beta = 0.031$; Master, $\beta = 0.037$; $p < 0.001$ for both).

A scan of the rest of the baskets of factors revealed that girls have significantly performed better than boys ($\beta = 0.044$; $p < 0.01$). In contrast, a negative coefficient $\beta = -0.072$; $p < 0.001$) for ‘age’ explains that students’ conceptual understanding declines with an increase in their age in years. Moreover, enrolment of students in private schools predicts better performance than their counterparts in government schools ($\beta = 0.217$; $p < 0.001$). It is important to note that system of school is one of the strongest predictors of students’ performance after region. Of the five regions, compared against Balochistan, Punjab has demonstrated the strongest positive contributions in students’ achievement ($\beta = 0.266$; $p < 0.001$) followed by GB ($\beta = 0.033$; $p < 0.01$). Contrarily, a negative coefficient in the case of AJ&K ($\beta = -0.026$; $p < 0.05$) illustrates low performance of students in this region as compared to Balochistan.

DISCUSSION

This large-scale study provides robust evidence that pedagogical quality is a significant predictor of students’ science performance in Pakistan. Classrooms characterised by participatory, engaging, and scaffolded instruction were associated with higher achievement, reinforcing international findings that emphasise the centrality of instructional quality in shaping learning (Akiba & Liang, 2014; Jacob et al., 2017; Teodorović et al., 2021). These results underscore

the importance of strengthening everyday teaching practices as a core strategy for improving science learning.

A notable set of findings concerns teacher qualifications and experience. Academic qualification showed no association with achievement, while professional qualification and teaching experience were negatively associated with student performance. Although counterintuitive, these patterns are consistent with studies reporting weak or inconsistent links between formal qualifications and student learning outcomes (Barasa, 2020; Telese, 2012; Wiseman & Al-Bakr, 2013; Zhang & Campbell, 2015). In the Pakistani context, however, this finding must be interpreted with particular caution. Many teacher education programmes have long been criticised for outdated curricula, limited practicum opportunities, and weak quality assurance mechanisms (Ansari, 2026; USAID, 2006, 2010). Although major reforms introduced the B.Ed. (Hons) programme in 2010, only a small proportion of graduates from these newer, more rigorous programmes have entered the teaching workforce. As a result, many “professionally qualified” teachers may hold older credentials that do not reflect contemporary pedagogical standards, while younger or unqualified teachers may be more adaptable, motivated, or aligned with recent instructional expectations. This contextual landscape helps explain why professional qualification may show a negative association in cross-sectional models and underscores the need to avoid overgeneralising this pattern.

The negative association with teaching experience should also be interpreted carefully. One plausible explanation is the rapid evolution of scientific knowledge: without sustained professional development, teachers’ content knowledge can become outdated, given that scientific fields have an estimated half-life of five to ten years (Crow, 1989; Robinson, 2011). Experience alone does not guarantee improved instructional practice; rather, it is the combination of experience and ongoing professional learning that supports teaching effectiveness (Rice, 2010). These findings therefore highlight system-level issues related to professional development and credential quality, rather than suggesting that experience or qualification are inherently detrimental.

Another striking finding relates to language consistency. Students performed better when the same language was used across textbooks, instruction, and assessment. This contributes new evidence to long-standing debates on the medium of instruction. While research consistently highlights the benefits of home-language instruction for participation and conceptual understanding (Fung, 2020; Prinsloo & Harvey, 2018), this study suggests that inconsistency, rather than the choice of language itself, may be particularly detrimental. When classroom instruction occurs in one language but assessments are administered in another, students may struggle to interpret questions even when they understand the underlying concepts. This highlights the importance of language alignment in multilingual systems.

The grade-level pattern also warrants attention. Students in Grade 6 outperformed those in Grade 5, but performance declined sharply by Grade 8. The downward trend is consistent with international evidence showing decreasing achievement and motivation as students progress through schooling (Ainley et al., 2020; Rosario et al., 2013; Potvin & Hasni, 2014; Toma et al., 2019). Possible explanations include reduced student engagement, increased curriculum difficulty, or less interactive pedagogy in upper grades. The Grade 6 performance bump likely reflects the transitional nature of this stage, when students from multiple primary schools converge into middle schools. Differences in curriculum alignment, assessment practices, or teacher expertise at this transition point may temporarily elevate achievement. Because these explanations remain tentative, the Grade 6 effect should be interpreted as a signal that system-level transitions deserve further investigation.

Parental education emerged as a strong and consistent predictor of achievement, with mothers' education exerting a particularly pronounced influence. This aligns with international evidence linking parental education to informed involvement, positive attitudes toward schooling, and enhanced academic support at home (Jeynes, 2015; Mahuro & Hungi, 2016; Wilder, 2014; Xiang & Chiu, 2021). These findings reinforce the importance of family educational capital in shaping learning trajectories.

Student-level factors also played a role. Girls outperformed boys, echoing studies that report gender advantages for girls in science achievement in some contexts (Shah et al., 2020). The negative association with age suggests that students who are older for their grade may be experiencing academic difficulties or motivational decline, consistent with literature documenting reduced interest in science as students grow older (Toma et al., 2019; Wang & Liou, 2017).

Finally, school and regional disparities were substantial. Students in private schools performed significantly better than those in public schools, a pattern widely documented in Pakistan and internationally (NEAS, 2016; Baum & Riley, 2019; Singh, 2014; Thapa, 2015). This advantage likely reflects a combination of endogenous factors (e.g., teacher accountability, in-house professional development, classroom support) and exogenous factors (e.g., socioeconomic status, parental support, access to tutoring) (Awan & Zia, 2015; Institute of Social and Policy Sciences, 2010). Regionally, Punjab demonstrated the strongest positive association, while AJ&K lagged behind, highlighting persistent geographic inequalities in educational provision and opportunity.

CONCLUSION

This study provides nationwide evidence from Pakistan on how pedagogical quality, alongside teacher, classroom, student, school, and regional factors, predicts students' science achievement. Rather than portraying a bleak picture of

science education, the findings offer a diagnostic lens for strengthening instructional quality and learning outcomes in a developing-country context. As Pakistan moves toward achieving SDG-4, the results underscore that improving science education requires more than curriculum reforms or assessment changes; it demands sustained attention to teaching quality, teacher preparation, and classroom-level supports.

The study highlights several actionable implications. First, the strong association between pedagogical quality and achievement reinforces the need to prioritise teacher preparation, induction, and continuous professional development. Second, the finding that language consistency across textbooks, instruction, and assessment predicts higher achievement offers a pragmatic direction for policy in multilingual systems. Rather than debating the “right” medium of instruction, policymakers should focus on ensuring alignment across learning materials, which directly supports students’ conceptual understanding in science.

To advance science learning, it is essential to mobilise resources for high-quality professional development that equips teachers to foster scientific literacy and 21st century skills. Equally important is the need for rigorous process- and outcome-oriented research to inform policy and practice. Educational policies should be treated as dynamic, evidence-informed processes rather than static directives.

This study has limitations that should guide future research. The cross-sectional design does not allow causal inference, and teachers were observed only once, which may not fully capture typical instructional practice. Although observations were standardised through a validated tool and trained observers, multiple observations across the school year would yield a more comprehensive picture of pedagogical quality. The nested structure of the data could, in principle, have been analysed using multilevel modelling; however, the limited number of cases within school-level clusters made such analysis unfeasible. Future studies could also incorporate additional variables, such as teachers’ content knowledge, beliefs about science teaching, students’ socioeconomic status, motivation, and school leadership practices, to deepen understanding of the mechanisms shaping science learning as well as conduct multilevel modelling to analyse nested data.

Despite these limitations, the study establishes a baseline for future longitudinal and intervention-based research in Pakistan and similar contexts in the Global South. By identifying the factors most strongly associated with science achievement, it provides a foundation for targeted reforms aimed at improving the quality of science teaching and learning.

DECLARATIONS

Conflict of Interest

The authors declare that there are no conflicts of interest associated with this study. Research data can be provided upon reasonable request.

Use of GenAI

An artificial intelligence tool (ChatGPT) was consulted solely for language checking. The article was originally written in mid-2022, prior to the widespread availability of generative AI tools.

REFERENCES

- Ahmed, M., Mahmood, T., Mhyuddin, M. S., & Ghuman, M. A. (2013). Evaluation of school effectiveness of secondary education system: Comparative view of public and private institutions of Pakistan. *Journal of Educational and Social Research*, 3(3).
- Ainley, J., Cloney, D., & Thompson, J. (2020). Does student grade contribute to the declining trend in programme for international student assessment reading and mathematics in Australia? *Australian Journal of Education*, 64(3), 205–226. <https://doi.org/10.1177/0004944120948654>
- Akhter, N. (2017). Factors Affecting Parents' Inclination towards private school system in Pakistan. *Journal of Elementary Education*, 27(2), 49–69.
- Akiba, M., & Liang, G. (2014). Teacher Qualification and the Achievement Gap: A Cross-National Analysis of 50 Countries. In J. Clark (Ed.), *Closing the achievement gap from an international perspective* (pp. 21–40). Springer. https://doi.org/10.1007/978-94-007-4357-1_3
- Ali, S. & Ahmed, A. (2022). *Teaching License in Pakistan: A White Paper*. Karachi: Aga Khan University, Institute for Educational Development. Pakistan. https://ecommons.aku.edu/cgi/viewcontent.cgi?article=1322&context=pakistan_ied_pdck
- Allison, P. D. (1999). *Multiple regression: A primer*. Pine Forge Press.
- Andrabi, T., Das, J., & Khwaja, A. I. (2008). A dime a day: The possibilities and limits of private schooling in Pakistan. *Comparative Education Review*, 52(3), 329–355. <https://doi.org/10.1086/589764>
- Ansari, A. N. (2026, May). *Teacher preparation for quality education: Comparative insights from twelve countries*. Paper presented at the BERA TEAN Conference 2026, Sheffield, United Kingdom.
- Asante, K. (2010). Sex differences in mathematics performance among senior high students in Ghana. *Gender and Behaviour*, 8(2). <https://doi.org/10.4314/gab.v8i2.61947>

- ASER. (2020). *Annual Status of Education Report (ASER) Pakistan 2019*. ASER Pakistan Secretariat, Idara-e-Taleem-o-Aagahi (ITA).
https://asERPakistan.org/document/asEr/2019/reports/national/ASER_National_2019.pdf
- Awan, A. G., & Zia, A. (2015). Comparative Analysis of Public and Private Educational Institutions: A case study of District Vehari-Pakistan. *Journal of Education and Practice*, 6(16), 122–130.
- Barasa, L. (2020). Teacher quality and mathematics performance in primary schools in Kenya. *African Journal of Research in Mathematics, Science and Technology Education*, 24(1), 53–64. <https://doi.org/10.1080/18117295.2020.1734164>
- Baum, D. R., & Riley, I. (2019). The relative effectiveness of private and public schools: Evidence from Kenya. *School Effectiveness and School Improvement*, 30(2), 104–130.
<https://doi.org/10.1080/09243453.2018.1520132>
- Bessudnov, A., & Makarov, A. (2015). School context and gender differences in mathematical performance among school graduates in Russia. *International Studies in Sociology of Education*, 25(1), 63–81.
- Bhai, M., & Horoi, I. (2019). Teacher characteristics and academic achievement. *Applied Economics*, 1–19. <https://doi.org/10.1080/00036846.2019.1597963>
- Bhutta, S. M., & Rizvi, N. F. (2022). *Assessing teachers' pedagogical practices and students' learning outcomes in science and mathematics across primary and secondary school level: A nationwide study (2018-21)*. Aga Khan University, Institute for Educational Development, Karachi, Pakistan.
https://ecommons.aku.edu/pakistan_ied_pdck/317
- Blomeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilsen, & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcome: Relationships across countries, cohorts and time* (pp. 21–50). Springer and IEA.
- Cairns, D. (2019). Investigating the relationship between instructional practices and science achievement in an inquiry-based learning environment. *International Journal of Science Education*, 41(15), 2113–2135.
<https://doi.org/10.1080/09500693.2019.1660927>
- Carroll, J. B. (1963). A model for school learning. *Teachers College Record*, 64, 723–733.
- Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., & Van Klaveren, C. (2017). Teacher characteristics and their effects on student test scores: A systematic review. *Journal of Economic Surveys*, 32(3), 848–877.
<https://doi.org/10.1111/joes.12210>

- Creemers, B. P. (1994). Effective instruction: An empirical basis for a theory of educational effectiveness. In *Advances in school effectiveness research and practice* (pp. 189-205). Pergamon.
- Crow, L.W. (1989). The why, what, how and who of critical thinking and other higher order thinking skills. In Crow, L.W. (Ed.), *Enhancing critical thinking in the sciences*. Second Edition. Houston: Baylor College of Medicine, 9–16.
- Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, 40(3), 291–309. <https://doi.org/10.1080/02619768.2017.1315399>
- Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (2019). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism*, 22(6), 661–674. <https://doi.org/10.1080/13670050.2017.1294142>
- Field, A. (2017). *Discovering statistics using SPSS* (5th ed.). Sage.
- Fischer, C., Foster, B., McCoy, A., Lawrenz, F., Dede, C., Eisenkraft, A., Fishman, B. J., Frumin, K., & Levy, A. J. (2020). Identifying levers related to student performance on high-stakes science exams: Examining school, teaching, teacher, and professional development characteristics. *Teachers College Record*, 122(2), 1-64. <https://doi.org/10.31219/osf.io/83hm2>
- Fung, D. (2020). The impacts of effective group work on social and gender differences in Hong Kong science classrooms. *International Journal of Science Education*, 42(3), 372–405. <https://doi.org/10.1080/09500693.2020.1713419>
- Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D., & Stuhlsatz, M. A. (2019). Teacher pedagogical content knowledge, practice, and student achievement. *International Journal of Science Education*, 41(7), 944–963.
- Government of Pakistan. (2006). *National Curriculum 2006*. Ministry of Federal Education and Professional Training Government of Pakistan.
- Graham, M., Milanowski, A., & Miller, J. (2012). *Measuring and promoting inter-rater agreement of teacher and principal performance ratings*. Online Submission. Center for Educator Compensation Reform. <https://files.eric.ed.gov/fulltext/ED532068.pdf>
- Hacıeminoğlu, E., Ertepinar, H., Yılmaz-Tüzün, Ö., & Çakır, H. (2015). Students and school characteristics related to elementary school students' views of the nature of science. *Education 3-13*, 43(6), 700–721. <https://doi.org/10.1080/03004279.2013.865655>
- Holzberger, D., Reinhold, S., Lüdtke, O., & Seidel, T. (2020). A meta-analysis on the relationship between school characteristics and student outcomes in

- science and maths – evidence from large-scale studies. *Studies in Science Education*, 56(1), 1–34. <https://doi.org/10.1080/03057267.2020.1735758>
- Hsu, Y. S., Wu, H. K., & Hwang, F. K. (2007). Factors influencing junior high school teachers' computer-based instructional practices regarding their instructional evolution stages. *Journal of Educational Technology & Society*, 10(4), 118–130.
- Institute of Social and Policy Sciences. (2010). *Private sector education in Pakistan: Mapping and musing*. Institute of Social and Policy Sciences (I-SAPS). Islamabad: Pakistan. http://i-saps.org/upload/report_publications/docs/1401025704.pdf
- Jacob, R., Hill, H., & Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness*, 10(2), 379–407. <https://doi.org/10.1080/19345747.2016.1273411>
- Jeynes, W. H. (2015). A meta-analysis: The relationship between father involvement and student academic achievement. *Urban Education*, 50(4), 387–423. <https://doi.org/10.1177/0042085914525789>
- Joshi, P. R., Digari, S., & James, M. C. (2022). The difference a female teacher makes: Analysis of girls' school achievement in Nepal. *Educational Studies*, 1–19. <https://doi.org/10.1080/00131946.2022.2051032>
- Karakose, T., Polat, H., Yirci, R., Tülübaş, T., Papadakis, S., Ozdemir, T. Y., & Demirkol, M. (2023). Assessment of the relationships between prospective mathematics teachers' classroom management anxiety, academic self-efficacy beliefs, academic amotivation and attitudes toward the teaching profession using structural equation modelling. *Mathematics*, 11(2), 449.
- Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and teacher education*, 36, 143–152. <https://doi.org/10.1016/j.tate.2013.07.010>
- Larry, T., & Wendt, J. L. (2021). Predictive relationship between gender, ethnicity, science self-efficacy, teacher interpersonal behaviors, and science achievement of students in a diverse urban high school. *Learning Environments Research*, 25(1), 141–157. <https://doi.org/10.1007/s10984-021-09354-1>
- Lau, K., & Lam, T. Y. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education*, 39(15), 2128–2149. <https://doi.org/10.1080/09500693.2017.1387947>
- Lee, J., Rhee, D., & Rudolf, R. (2018). Teacher gender, student gender, and primary school achievement: Evidence from ten Francophone African

- countries. *The Journal of Development Studies*, 55(4), 661–679. <https://doi.org/10.1080/00220388.2018.1453604>
- Li, S., Liu, X., Yang, Y., & Tripp, J. (2021). Effects of teacher professional development and science classroom learning environment on students' science achievement. *Research in Science Education*. <https://doi.org/10.1007/s11165-020-09979-x>
- Liu, O. L., Lee, H. S., & Linn, M. C. (2010). An investigation of teacher impact on student inquiry science performance using a hierarchical linear model. *Journal of Research in Science Teaching*, 47(7), 807–819.
- Mahuro, G., & Hungi, N. (2016). Parental participation improves student academic achievement: A case of Iganga and Mayuge districts in Uganda. *Cogent Education*, 3(1), 1264170. <https://doi.org/10.1080/2331186x.2016.1264170>
- Mose, P. N. (2019). Bilingualizing linguistically homogeneous classrooms in Kenya: Implications on policy, second language learning, and literacy. *International Journal of Bilingual Education and Bilingualism*, 22(4), 459–472. <https://doi.org/10.1080/13670050.2016.1268567>
- National Education Assessment System (NEAS). (2016). *National Assessment Report 2016*. Ministry of Federal Education & Professional Training. Islamabad: Pakistan. <http://www.neas.gov.pk/Download>
- NEMIS. (2019). Pakistan Education Statistics 2017-18. <http://library.aepam.edu.pk/Books/Pakistan%20Education%20Statistics%202017-18.pdf>
- Nilsen, T., Gustafsson, J. E., Blömeke, S. (2016). Conceptual Framework and Methodology of This Report. In: Nilsen, T., Gustafsson, JE. (eds) Teacher Quality, Instructional Quality and Student Outcomes. *IEA Research for Education*, 2. Springer, Cham. https://doi.org/10.1007/978-3-319-41252-8_1
- O'Connell, M. (2019). Is the impact of SES on educational performance overestimated? Evidence from PISA survey. *Intelligence* (75), 41–47.
- OECD. (2019). *PISA 2018 assessment and analytical framework*. Paris: OECD. <https://doi.org/10.1787/b25efab8-en>
- Oppermann, E., Hummel, T., & Anders, Y. (2021). Preschool teachers' science practices: associations with teachers' qualifications and their self-efficacy beliefs in science. *Early Child Development and Care*, 191(5), 800–814.
- Potvin, P., & Hasni, A. (2014). Analysis of the decline in interest towards school science and technology from grades 5 through 11. *Journal of Science Education and Technology*, 23(6), 784–802. <https://doi.org/10.1007/s10956-014-9512-x>
- Prinsloo, C. H., & Harvey, J. C. (2018). The differing effect of language factors on science and mathematics achievement using TIMSS 2015 data: South

- Africa. *Research in Science Education*, 50(6), 2207–2226. <https://doi.org/10.1007/s11165-018-9769-9>
- Pun, J., & Macaro, E. (2019). The effect of first and second language use on question types in English medium instruction science classrooms in Hong Kong. *International Journal of Bilingual Education and Bilingualism*, 22(1), 64–77. <https://doi.org/10.1080/13670050.2018.1510368>
- Reynolds, C. R., Livingston, R. B., & Willson, V. L. (2009). *Measurement and assessment in education* (2nd ed.). Pearson Education Inc. USA.
- Rice, J. K. (2010). The impact of teacher experience: Examining the evidence and policy implications. *National center for analysis of Longitudinal data in education research (CALDER)*. <https://files.eric.ed.gov/fulltext/ED511988.pdf>.
- Robinson, S. R. (2011). Teaching logic and teaching critical thinking: revisiting McPeck. *Higher Education Research & Development*, 30(3), 275–287. <https://doi.org/10.1080/07294360.2010.500656>
- Rosario, P., Núñez, J. C., Valle, A., González-Pienda, J., & Lourenço, A. (2013). Grade level, study time, and grade retention and their effects on motivation, self-regulated learning strategies, and mathematics achievement: a structural equation model. *European Journal of Psychology of Education*, 28(4), 1311–1331.
- Sarouphim, K. M., & Chartouny, M. (2016). Mathematics education in Lebanon: Gender differences in attitudes and achievement. *Educational Studies in Mathematics*, 94(1), 55–68.
- Shah, Z. A., Malik, M. A., & Akhtar, J. H. (2020). Exploring the impact of demographic variables gender, parental education and locality on science achievement at 8th and 9th Grades. *Bulletin of Education and Research*, 42(1), 185–198.
- Singh, A. (2014). Test score gaps between private and government sector students at school entry age in India. *Oxford Review of Education*, 40(1), 30–49. <https://doi.org/10.1080/03054985.2013.873529>
- Smith-Wooley, E., Pingault, J. B., Selzam, S., Rimfeld, K., Krapohl, E., von Stumm, S., Dale, P. S., Young, T., Allen, R., & Plomin, R. (2018). Differences in exam performance between pupils attending selective and non-selective schools mirror the genetic differences between them. *Science of Learning*, 3(3), 1–8. <https://www.nature.com/articles/s41539-018-0019-8>
- Symeou, L. (2007). Cultural capital and family involvement in children's education: Tales from two primary schools in Cyprus. *British Journal of Sociology of Education*, 28(4), 473–487. <https://doi.org/10.1080/01425690701369525>
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2019). *Using multivariate statistics*. Pearson.

- Telese, J. A. (2012). Middle school mathematics teachers' professional development and student achievement. *The Journal of Educational Research*, 105(2), 102–111.
- Teodorović, J., Milin, V., Bodroža, B., Đerić, I. D., Vujačić, M., Jakšić, I. M., Stanković, D., Cankar, G., Charalambous, C. Y., Damme, J. V., & Kyriakides, L. (2021). Testing the dynamic model of educational effectiveness: The impact of teacher factors on interest and achievement in mathematics and biology in Serbia. *School Effectiveness and School Improvement*, 1-35. <https://doi.org/10.1080/09243453.2021.1942076>
- Thapa, A. (2015). Public and private school performance in Nepal: An analysis using the SLC examination. *Education Economics*, 23(1), 47–62. <https://doi.org/10.1080/09645292.2012.738809>
- TIMSS. (2020). *TIMSS 2019 International Results in Mathematics and Science*. International Association for the Evaluation of Educational Achievement. <https://timss2019.org/reports/wp-content/themes/timssandpirls/download-center/TIMSS-2019-International-Results-in-Mathematics-and-Science.pdf>
- Toma, R. B., Greca, I. M., & Orozco Gómez, M. L. (2019). Attitudes towards science and views of nature of science among elementary school students in terms of gender, cultural background and grade level variables. *Research in Science & Technological Education*, 37(4), 492–515. <https://doi.org/10.1080/02635143.2018.1561433>
- Tonga, F. E., Eryiğit, S., Yalçın, F. A., & Erden, F. T. (2022). Professional development of teachers in PISA achiever countries: Finland, Estonia, Japan, Singapore and China. *Professional Development in Education*, 48(1), 88-104. <https://doi.org/10.1080/19415257.2019.1689521>
- Toropova, A., Johansson, S., & Myberg, E. (2019). The role of teacher characteristics for student achievement in mathematics and student perceptions of instructional quality. *Education Inquiry*, 10(4), 276–299. <https://doi.org/10.1080/20004508.2019.1591844>
- Tsai, L., & Yang, C. (2015). Hierarchical effects of school-, classroom-, and student-level factors on the science performance of eighth-grade Taiwanese students. *International Journal of Science Education*, 37(8), 1166–1181. <https://doi.org/10.1080/09500693.2015.1022625>
- USAID. (2006). *Situation Analysis of Teacher Education in Pakistan: Towards a Strategic Framework for Teacher Education and Professional Development*. <https://eric.ed.gov/?id=ED497052>
- USAID. (2010). *Rationalization of Pre-Service Teacher Education Programs in Pakistan*. Pre-STEP USAID: Islamabad.
- Wang, C., & Liou, P. (2017). Students' motivational beliefs in science learning, school motivational contexts, and science achievement in Taiwan. *International Journal of Science Education*, 39(7), 898–917. <https://doi.org/10.1080/09500693.2017.1310410>

- Wilberg, M., & Rolfsman, E. (2019). The association between science achievement measures in schools and TIMSS science achievement in Sweden. *International Journal of Science Education*, 41(16), 2218–2232. <https://doi.org/10.1080/09500693.2019.1666217>
- Wilder, S. (2014). Effects of parental involvement on academic achievement: A meta-synthesis. *Educational Review*, 66(3), 377–397. <https://doi.org/10.1080/00131911.2013.780009>
- Wiseman, A. W., & Al-Bakr, F. (2013). The elusiveness of teacher quality: A comparative analysis of teacher certification and student achievement in Gulf Cooperation Council (GCC) countries. *Prospects*, 43(3), 289–309. <https://doi.org/10.1007/s11125-013-9272-z>
- Xiang, N., & Chiu, S. W. K. (2021). Parents or teachers: whose guidance matters more in students' alignment of educational plan and achievement? A comparative study of Taiwan and Hong Kong. *Asia Pacific Journal of Education*, 1–16.
- Zhang, D., & Campbell, T. (2015). An examination of the impact of teacher quality and "Opportunity gap" on student science achievement in China. *International Journal of Science and Mathematics Education*, 13(3), 489–513. <https://doi.org/10.1007/s10763-013-9491-z>
- Zhang, F., & Bae, C. L. (2020). Motivational factors that influence student science achievement: A systematic literature review of TIMSS studies. *International Journal of Science Education*, 42(17), 2921–2944. <https://doi.org/10.1080/09500693.2020.1843083>
- Zhang, F., Bae, C. L., & Broda, M. (2022). Science self-concept, relatedness, and teaching quality: A multilevel approach to examining factors that predict science achievement. *International Journal of Science and Mathematics Education*, 20(3), 503–529. <https://doi.org/10.1007/s10763-021-10165-2>

Bio

SADIA MUZAFFA BHUTTA, PhD, is an Associate Professor at the Aga Khan University, Institute for Educational Development, Karachi, Pakistan. Her major research interests include classroom teaching, science education, and quantitative research methods. Email: sadia.bhutta@aku.edu

NUSRAT FATIMA RIZVI, is a faculty member at the Aga Khan University, Institute for Educational Development, Karachi, Pakistan. Her major research interests include mathematics education and teacher education. Email: nusrat.fatima@aku.edu

SOHAIL AHMAD, is a Research Specialist and Doctoral Candidate at the Aga Khan University, Institute for Educational Development, Karachi, Pakistan. Alongside this, he is serving as a Research Associate at the REAL Centre, Faculty of Education, University of Cambridge, England. His major research interests include higher education, classroom teaching, and quality assurance. Email: sohail.ahmad2@scholar.aku.edu

ANIL KHAMIS, is a Research Faculty at the Institute for Educational Development, East Africa. His major research interests include education policy and sociology of education.

AISHA NAZ ANSARI, is a PhD Scholar at School of Education, Durham University, England. Alongside this, she has served as a Research Specialist at the Aga Khan University, Institute for Educational Development, Karachi, Pakistan. Her major research interests include teacher education, classroom teaching, and systematic reviews. Email: aisha.naz22@alumni.aku.edu