Article

Volume 20 (2026), pp. 227-250
American Journal of STEM Education: Issues and Perspectives
eISSN 30.3-1190| Print ISSN: 3069-0072
https://doi.org/10.32674/s4xywt33

Pedagogical Quality of Science and Mathematics Teachers Across Pakistan: Evidence From a Nationwide Observational Study

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

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

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

Afaq Ahmed
Texas A&M University, United States
https://orcid.org/0000-0001-6382-1197

ABSTRACT

This study examines the pedagogical quality of science and mathematics teachers across Pakistan using observational data. A large-scale, cross-sectional classroom observation study was conducted with 589 science and mathematics teachers across 153 elementary schools. The findings reveal overall weak pedagogical quality among teachers. However, teachers in private schools, female teachers, and those teaching early grades demonstrate relatively better pedagogical practices. Female teachers, in particular, show greater attention to *student interaction and instructional strategies. The study also highlights a tendency among teachers to prioritise content delivery as subject complexity increases, often at the expense of pedagogical quality. The findings have implications for teacher training and professional development in Pakistan, highlighting the need to strengthen pedagogical content knowledge through continuous professional learning and participatory instructional strategies.*

Keywords: Mathematics teaching, Nationwide study, Observational study, Pakistan, Pedagogical quality, Science teaching.

INTRODUCTION

There is a global consensus that teaching quality is crucial for school effectiveness. A major strand of research on teaching quality has used value-added designs, assessing teachers' background characteristics, such as qualifications, certification test scores, and experience, and relating them to student achievement (Buddin & Zamarro, 2009; Chetty et al., 2014; Darling-Hammond, 2015; Hanushek & Rivkin, 2010; Koedel & Betts, 2010; Koedal et al., 2015). Whilst this type of research has attracted policymakers' attention, a key limitation is that it rarely informs policymakers about what happens inside classrooms or about pathways to improve students' learning deficits through pedagogical quality (Boston, 2012; McCaffrey et al., 2003; Pianta & Hamre, 2009). To address this gap, researchers have examined teachers' pedagogical quality to determine which dimensions of mathematics and science classrooms result in better student learning. Research consistently demonstrates that pedagogical quality is directly linked to both teacher effectiveness (Bolyard & Moyer-Packenham, 2008; Blömeke et al., 2022; Hattie, 2009; Hattie & Hamilton, 2022; Konig & Pflanzl, 2016) and student performance (Bhutta & Rizvi, 2022; Blömeke & Olsen, 2019; Burroughs et al., 2019; Crosnoe et al., 2010; Decristan et al., 2016; Emslander et al., 2025; Hattie, 2009; Meyer et al., 2011; Muijs et al., 2014; Rivkin & Schiman, 2015).

Pedagogical quality is a broad term that encompasses both specific and general actions and practices of teachers inside the classroom. In his seminal work, Shulman (1986) refers to this as "pedagogical content knowledge". Several studies have inferred teachers' pedagogical quality from their pedagogical knowledge, assessed through paper-and-pencil tests (König & Pflanzl, 2016; Voss et al., 2011). Although such evidence is cost-effective, it does not help in understanding what is happening inside the classroom and is therefore inefficient for informing policymakers about targeted interventions related to teaching. In other words, there may be differences between what teachers know about pedagogical approaches and what they actually do in the classroom. It is argued that alternative measures, such as classroom observations, can be viable instruments for capturing actual pedagogical quality and can be used to improve instruction and inform teacher education to better prepare teachers (Boston, 2012; Pianta & Hamre, 2009; Stein & Matsumura, 2008). According to Boston (2012, p. 97), observational studies:

> serve the goal of improving instruction, measures of instructional quality must be capable of identifying aspects of classroom practice that influence students' opportunities to learn mathematics and, most critically, of communicating this information to school districts in ways that provide direction for systemic reforms… Reporting the results of each dimension of practice at the school or district level can target areas of strength and improvement and provide usable feedback to enable the school or district to enhance instruction and students' opportunities to learn mathematics. (p.97)

While recognising the significance of observational studies, researchers have developed several instruments to document teachers' pedagogical competencies directly, including, for example, video-based assessment in Germany (Bruckmaier et al., 2016; Dunekacke et al., 2015); the Classroom Assessment Scoring System (Pianta & Hamre, 2009); the Reformed Teaching Observation Protocol for mathematics and science in Arizona (Sawada et al., 2002); the Korean Teaching Observation Protocol for science teaching (Park et al., 2012); and instructional quality assessment in science in the USA (Tekkumru-Kisa, 2020). Studies measuring pedagogical quality through observations vary in the aspects they treat as important factors, such as teacher–child interactions and exposure to instruction (Ottmar et al., 2014; Kang, 2022); classroom management strategies (effective organisation and use of time) (Decristan et al., 2016; Kang, 2022; Lipowsky et al., 2009; Praetorius et al., 2018); a conducive environment (Fauth et al., 2014); and cognitive activation for higher-order thinking (Fauth et al., 2014; Klieme et al., 2009; Lipowsky et al., 2009). Recent work in mathematics education continues to highlight the breadth of observation-based measures used to capture teaching quality, alongside ongoing concerns about validity and fairness evidence across instruments (Gallagher et al., 2025). Existing observational studies also often focus on limited aspects of classroom practice (e.g., classroom management or teacher–child interaction) and may be subject to methodological issues such as convenience sampling (Lee & Santagata, 2020), student ratings that may reflect personal likes and dislikes (Baier et al., 2019), or teachers' self-reports of instructional practices that are influenced by social desirability (van de Mortel, 2008). In addition, researchers have identified that cultural context plays an important role in observational studies assessing teachers' pedagogical quality. Senden et al. (2021), for example, argued that existing instruments on pedagogical quality originated and were implemented in the global North according to its cultural standards and have not been replicated sufficiently in other contexts. This is because factors such as socioeconomic conditions, school support, and teacher development mechanisms vary. For instance, a cross-country study showed that schools with low socioeconomic status have less qualified teachers (Banjerjee, 2016; Liu et al., 2022).

Ongoing reforms in teacher education require evidence of teachers' pedagogical characteristics to target the competencies that are crucial yet lacking. This study helps to fill this gap by presenting nationwide findings from an observational study in Pakistan. The paper provides evidence on pedagogical quality from a nationwide study that captures classroom practices in mathematics and science in a holistic way—focusing on multiple aspects of teaching and learning. Rather than rating specific behaviours for a limited time, this study reports instructional quality across the full duration of the lesson (see Senden et al., 2021).

Research Questions

1. What is the quality of mathematics and science teachers' teaching in elementary schools across Pakistan?
2. Does the quality of mathematics and science teachers' teaching vary across grade levels in elementary schools of Pakistan?
3. Does the quality of mathematics and science teachers' teaching vary across the gender of teachers in elementary schools of Pakistan?
4. Does the quality of mathematics and science teachers' teaching vary across public and private elementary schools of Pakistan?

LITERATURE REVIEW

Bruner's Theory of Experiential Learning

Jerome S. Bruner, in search of pedagogy, gave a powerful statement.

> The proper reward [of learning] is that we can now use what we have learned, and can cross the barrier from learning into thinking. Are we mindful of these matters in our conduct of teaching? (p.25).

Bruner's theory critiques teaching as merely presenting material related to a subject and viewing learning as a one-directional process (Takaya, 2008). Instead, teaching should expose students to concepts and ideas that can be tried, experimented with, and/or acted on to examine their usefulness. Learning is seen as a simultaneous process in which new information is acquired (replacing previous knowledge), transformed (analysed by converting information from one form to another), and evaluated (judging whether the new information is adequate) (Bruner, 2006). In his search for pedagogy, Bruner uses the term "recoding push" for effective teaching, asserting that teachers must have the ability to condense and transform material into deeper structures of experience for children. Once children develop a deeper understanding, they can make higher-level connections. For example, Bruner would oppose teaching "addition" and "multiplication" as separate units. Instead, if addition is taught in depth, it should enable students to understand multiplication because multiplication is ultimately based on addition. Bruner (1971) further explains that discovery teaching involves students discovering not only what is "out there" but also what is in their own minds. The theory of learning posits that students' learning should serve them in the future, which occurs in two ways: (i) the applicability of tasks is similar to what they have performed in classrooms, an extension of habits or associations; and (ii) the transfer of principles and attitudes enables students to recognise problems as special cases of an originally mastered idea (Bruner, 2006). This means that what is applicable or inapplicable for students depends largely on teachers' skill in dealing with the phenomenon. In terms of intellectual development, Bruner builds on Piaget's idea that children learn better when concepts are represented in ways that align with how children experience them. Development, therefore, occurs through trial and error. Yet, this development does not happen in a vacuum; it is influenced by the environment as well. Therefore, children need not only challenging tasks but also opportunities that make growth into subsequent stages possible.

Dewey's Theory of Experiential Learning

Dewey also emphasised experiential learning for children as a means of producing growth. Growth was viewed without a fixed direction, to avoid requiring students to meet a predetermined purpose of education (Noddings, 2016). Dewey's concept of experiential learning focused on engagement in the present and on gaining new experiences in schools. Experiential learning suggests that humans respond to stimuli in ways consonant with their purposes (Noddings, 2016). An important aspect of experiential learning is that it goes beyond exposure; an individual must be undergoing a meaningful process. The experiential process is embedded in the social and cultural domain and therefore emphasises personal meaning and social interaction (Dewey, 1975). To build continuity, experiences must be connected to previous experiences. Hence, the teacher's role is to understand children's past experiences and to formulate new activities and experiences so that the child grows. In contrast to trial and error, Dewey calls for reflective experience to deal with uncertainty or doubt-filled situations (English, 2013). From Dewey's perspective, trial and error is limited because it only indicates whether changes were successful or unsuccessful; however, children must also understand why the change occurred. To address this, English (2013) argues that "reflection is an inquisitive form of thinking that holds us in suspense, analyses the limits of given knowledge, and develops new possible ways of seeing the world". In this view of learning, children arrive at new learning through inquiry and make connections with their own experiences.

Student-Centred vs Teacher-Centred Teaching in Mathematics and Science

Teachers' decisions about selecting teaching techniques, materials, and the involvement of students play a crucial role in maximising student learning (Cairns, 2019; Hiebert & Grouws, 2007; Roorda et al., 2011; Yu & Singh, 2016). There is an ongoing discussion about the effectiveness of teacher-centred and student-centred approaches in science and mathematics. In these subjects, project-based learning, inquiry-based learning, problem-solving, and collaborative learning have been advocated as effective student-centred teaching practices (Cairns, 2019; Chi et al., 2018; Hmelo-Silver et al., 2007). These practices have a positive impact on students' achievement in science (Gee & Wong, 2012; Jiang & McComas, 2015; Maxwell et al., 2015; Sabaha & Hammouri, 2010) but a negative association in mathematics (Al-Hammouri, 2005; Jo & Seo, 2021; Sabaha & Hammouri, 2010). The difference between the two approaches is apparent. Student-centred teaching is based on constructivism, which requires careful pedagogical planning by teachers. In contrast, traditional teacher-centred approaches, commonly used in many Asian countries, allow coverage of course content within a limited timeframe, often resulting in better performance in examinations (Tao, Oliver, & Venville, 2013).

Research shows a positive association between teacher-centred teaching and students' academic performance in science and mathematics (Huang & Cribbs, 2017; Liou & Ho, 2016; Lau & Lam, 2017; Tao et al., 2013). However, the role of context needs to be considered in empirical evidence. For instance, students in China may view science teaching as passive and rely on memorisation, while Australian students may have more interactive, practical-based classes due to teacher autonomy in lesson design (Tao et al., 2013). There is still uncertainty among researchers about the most effective approach to teaching mathematics (see, e.g., Yu & Singh, 2016). One cannot claim that a single teaching approach guarantees success in science and mathematics classrooms. Effective delivery of instruction depends on teachers' professional skills (Blömeke, Olsen & Suhl, 2016) and the availability of appropriate resources and a supportive environment (Sabah & Hammouri, 2010).

RESEARCH CONTEXT

Pakistan's education system faces two persistent issues, access and quality, at almost all levels. Despite multifarious challenges, Pakistan is among the countries that constitutionally assume responsibility for providing free and compulsory education to citizens aged 5 to 16 years. At the K-12 level, public-sector schools (managed by the government in terms of curriculum, teachers, finance, and other matters) and the private sector (managed by non-government organisations or individuals, either for profit or not for profit) are the two main providers of education to nearly 51.2 million children (AEPAM, 2018). The former accommodates 56% of children, while 44% are enrolled in private schools. In terms of the teaching force, similar proportions of teachers work in the public (46%) and private (46%) sectors (AEPAM, 2018, p. 10). Regarding teacher gender, the teaching force includes more female teachers (60%) than male teachers (40%), and the difference may reflect private schools' preference for recruiting female teachers (AEPAM, 2018). Apparent differences between public and private schools include the language of instruction, examination systems, and textbooks. Nevertheless, there is a general observation that those who can afford fees prefer to send their children to private schools due to social norms and the pursuit of better quality. Large-scale studies support the latter assumption; for instance, evidence indicates better student performance in core subjects such as science and mathematics in private schools (Bhutta & Rizvi, 2022; NAT, 2016). Recent studies continue to highlight challenges in mathematics and STEM education in Pakistan, including teacher-related and school-level constraints (Numajiri & Kosaka, 2025; Rehman et al., 2025). Despite extensive evidence on student learning, there is hardly any large-scale evidence about the quality of teaching in Pakistani schools. This study attempts to bridge this long-overdue gap in the literature from a developing-country context such as Pakistan.

RESEARCH METHOD

Research design

This study is drawn from a large-scale project aiming to assess teachers' pedagogical practices in science and mathematics across Pakistan. A cross-sectional survey was carried out to assess the quality of teaching pedagogy. Considering the population frame and the nature of the assessment, a cross-sectional survey was deemed a suitable and efficient design for gathering data from a large number of respondents (Fraenkel et al., 2011; Robson & McCartan, 2016). This design was also appropriate for comparing various subgroups (i.e., school type, grades, and gender) (Gorard, 2003). The study was approved by the institutional ethics review committee.

Sample and Sampling

A multi-stage stratified sampling technique was used to recruit teachers from 25 districts, proportionally selected from all six regions of Pakistan [Punjab, Sindh, Khyber Pakhtunkhwa (KP), Balochistan, Gilgit Baltistan (GB), and Azad Jammu & Kashmir (AJK)]. The number of districts sampled was determined by regional population size. For example, Punjab, the largest province by population, represents a larger proportion of the sample at the district level (n=8), followed by Sindh (n=6), KP (n=3), Balochistan (n=4), GB (n=2), and AJK (n=2). From each district, six schools were randomly selected with equal representation from public and private schools. Altogether, 153 elementary schools participated in this study.

Teachers' Assessment Tool - COS

Teachers' pedagogical practices in science and mathematics were observed using a structured observation tool (Classroom Observation Scale - COS). The COS was administered to observe the teaching practices of target teachers for one lesson each. The COS is a 15-item rubric (13 items on a 3-point scale and two categorical items) that encompasses various aspects of classroom quality. These quality indicators were used to observe both science and mathematics classrooms to gauge the extent to which students and teachers participate together in the teaching-learning environment. In this rubric, descriptors are anchored on three points, and specific guidelines are provided for observers for each item. In addition, a demographic questionnaire was administered to collect information on relevant teacher characteristics, including gender, age, experience, and qualifications.

In order to check the internal consistency of the COS, Cronbach's Alpha for both science and mathematics classrooms was computed and found to be in an acceptable range (0.70 for mathematics and 0.80 for science).

Data Analysis

The data were analysed using SPSS v.26. After data cleaning, the mean COS score was computed, and categorical variables were coded (e.g., system of school: public = 1, private = 2; teacher gender: male = 1, female = 2). The data were explored to check the assumptions of parametric tests. The distributions for the school system and teacher gender were found to be skewed; therefore, the non-parametric Mann-Whitney test was used to compare differences between the two groups. For grade level, a one-way ANOVA was used to compare teaching quality across the three target grades. Furthermore, the magnitude of differences was estimated by computing effect sizes (r2 for comparisons of two groups; η2 for more than two groups). Benchmarks for effect sizes were defined as small ($r = 0.1$; $\eta 2 = 0.01$), medium ($r = 0.3$; $\eta 2 = 0.06$), and large ($r = 0.5$; $\eta 2 = 0.14$) (Cohen, 1988; Field, 2017). Additionally, a framework was developed to categorise the quality of classroom practices into three categories (weak = ≤1.50; mediocre = 1.51 – 2.50; good = ≥2.51). Overall teaching quality for science and mathematics was presented separately, along with comparisons across grades, school systems, and teacher gender to address the research questions. To obtain a holistic picture of classroom practice, an item-wise analysis was also conducted.

Altogether, 589 teachers participated in this study, including mathematics (n=303; 51%) and science (n=286; 49%) teachers. The sample includes teachers from all six regions of Pakistan, with the highest proportion from Punjab (32%) and the lowest from AJK (6%), reflecting the size of the teaching force and the number of schools in each region. More teachers from the public system participated than from private schools, and this distribution across the two systems was similar for both subjects ($p>0.05$). Overall, there were more female teachers (n=359; 61%) than male teachers (n=230; 39%). Gender distribution across the two subjects followed a similar pattern (female > male); however, female representation was higher in science (65%) than in mathematics (57%), and the difference was statistically significant ($p<0.05$).

A review of academic qualifications showed that most of the teachers (72%) held a master's degree, with a similar distribution across the two subjects ($p>0.05$). A similar trend was observed across public and private school systems. An overwhelming majority of science and mathematics teachers reported having a professional qualification: most science (53%) and mathematics (46%) teachers had completed a Bachelor of Education (B.Ed.), while similar proportions in both subjects (Science=23%; Mathematics=26%) reported having completed a higher professional qualification (M.Ed.). Further analysis comparing professional qualifications across school systems indicated that a substantial proportion of teachers (over 40%), regardless of subject, in the private system did not have any professional qualification.

Overall, the participating teachers' teaching experience averaged 11 years (SD=9.4) for science and 12 years (SD=9.6) for mathematics. The sample included a wide range of experience in both subjects, from less than one year to 36 years in science and 35 years in mathematics. Comparisons across school systems showed that public-school teachers, regardless of subject, were more experienced ($p<0.001$).

RESULTS

Quality of Classroom Practices in Science and Mathematics

Observation revealed that the classroom practices of science (M=1.28; SD= 0.24) and mathematics (M=1.28; SD= 0.20) teachers were similar as measured on a 3-point scale. Table 01 presents the percentage of lessons by ranges of COS total scores across subjects.

Table 1: Percentage of lessons by ranges of COS total score across subjects

Subject/Range	Science	Mathematics
Weak practices [<1.50]	247 (86%)	265 (87%)
Mediocre practices [1.50 – 2.50]	39 (14%)	38 (13%)
Good practices [>2.50]	0 (0%)	0 (0%)

Placing these results within the framework, it should be noted that, regardless of subject, most observed lessons in this sample had a COS score below 1.5 (weak classroom practices). It is also worth noting that a small number of observed lessons achieved a score that qualified as "mediocre practices" (1.5-2.5), as defined in the framework. Furthermore, none of the teachers demonstrated "good classroom practices" in either subject (>2.5). Comparisons between science and mathematics showed almost identical results. For instance, most lessons in both science (n=247; 86%) and mathematics (n=265; 87%) were classified as weak practice. A small number of lessons fell in the middle band of quality, with a negligible difference between mathematics (n=38; 13%) and science (n=39; 14%). Item-wise results provide a more detailed picture, as presented in Table 2.

Table 2: Item-wise analysis of classroom practice

COS-Items	Science	Mathematics
Indoor Space*	**1.41 (0.62)**	1.32 (0.54)
Classroom Arrangement	1.07 (0.29)	1.03 (0.18)
Classroom Display*	**1.10 (0.30)**	1.03 (0.18)
Use Of Material*	**1.22 (0.44)**	1.11 (0.31)
Teacher-Child Interactions	1.85 (0.72)	1.89 (0.66)
Child-Child Interactions	1.09 (0.30)	1.11 (0.32)
Discipline	1.82 (0.40)	1.78 (0.44)
Supervision*	1.10 (0.31)	**1.28 (0.49)**
Active Involvement of Children*	1.27 (0.50)	**1.36 (0.53)**
Questioning	1.23 (0.46)	1.18 (0.40)
Peer Learning / Group Work	1.04 (0.19)	1.01 (0.11)
Start and Closure of the Lesson	1.21 (0.44)	1.15 (0.39)
Teacher's Instructions*	1.25 (0.55)	**1.42 (0.68)**

*Mean (Standard deviation); *significant difference*

Mediocre practice (1.5-2.5): An overall comparison of COS item scores across the two subjects revealed that only two items (i.e., teacher–child interaction and discipline) fell in the category of mediocre practice. A consistent pattern was observed across the two subjects. A relatively high score in teacher–child interaction indicated that teachers interacted with children in a supportive manner in some instances. Generally, teachers demonstrated supportive interactions with students, but not with all students (e.g., outspoken children were given more attention). Moreover, many teachers were observed to maintain sufficient control in their classes without using harsh discipline techniques such as corporal punishment or excluding children from activities for long periods. However, there were hardly any examples of teachers using progressive techniques such as praising students' positive behaviour (discipline).

Weak practices (<1.5): Of the thirteen COS items, the remaining eleven fell in the category of weak practice. These items span infrastructure, classroom interaction, and teaching–learning strategies. All four infrastructure items (indoor space, classroom arrangement, classroom display, and use of material) fell in the weak-practice category. However, within this band, science classrooms demonstrated an edge over mathematics classrooms, with significant differences for all items ($p<0.05$) except classroom arrangement. This suggests that science classrooms were relatively tidy, had some subject-related displays, and provided some materials (e.g., chart paper, videos, models) that teachers used to teach science. That said, the low average scores indicate that only a small fraction of classes demonstrated better infrastructure, while the majority remained in the lowest band. Frequency-level analysis showed that, regardless of subject, most classes remained in the lowest band, with only a small fraction reaching the mediocre level. Interestingly, a very small fraction of science classrooms was assigned the highest score (3 on a 3-point scale) in all four items. While none of the mathematics classrooms qualified for the highest score, indoor space was an exception, where a very small fraction reached that level. This frequency pattern not only explains the significant differences between classroom practices in the two subjects but also helps to explain the overall low scores on these items. Additionally, two items related to classroom interaction (child–child interaction and supervision) fell in the weak-practice category in both science and mathematics classrooms. Unlike the infrastructure items, these interaction items slightly favoured mathematics classrooms, with a significant difference only for supervision ($p<0.05$). This suggests that a fraction of mathematics teachers attempted to provide feedback to some students on their work, clarify tasks, and help solve questions. The low mean score for child–child interaction indicates limited opportunities for peer interaction in both science and mathematics classrooms. In other words, there were hardly any instances where teachers provided opportunities for students to interact with each other for peer learning. Frequency distributions showed that many classes remained in the lowest band, with higher frequency in science classrooms. Regardless of subject, a fraction of classrooms approached the mediocre level, with a slight edge for mathematics. For both subjects, a negligible fraction qualified for the highest score (3 on a 3-point scale) in supervision; however, only science classrooms (though a very small proportion) reached the highest level for child–child interaction.

Moreover, all five items related to teaching–learning strategies fell in the category of weak practice in both science and mathematics classrooms. Within this band, science classrooms showed a slight edge over mathematics in three items (questioning, peer/group learning, and start and closure of the lesson), although the differences were not statistically significant. Conversely, comparisons favoured mathematics classrooms for active involvement of children and teachers' instruction, with statistically significant differences. These results indicate that, compared with science, more mathematics teachers provided opportunities for some students to engage in thinking processes for short periods. Similarly, more mathematics teachers showed some clarity in giving instructions for tasks compared with their counterparts in science classrooms. Frequency distributions indicated that, regardless of subject, a small proportion of classrooms qualified for the highest score (3 on a 3-point scale) on most items, except peer learning/group work. Nevertheless, as in the other categories, the majority of classes remained in the lowest band in both science and mathematics. For active involvement of students and teachers' instructions, more science classrooms fell in the lowest performing category compared with mathematics, whereas mathematics had more classrooms in the middle band. In contrast, for questioning, peer learning/group work, and start and closure, mathematics had slightly more classrooms in the lowest band compared with science. Unfortunately, peer learning/group work had the highest representation in the lowest band for both subjects. Put differently, much remains to be desired in implementing cooperative learning to prepare students to be engaged citizens in the 21st century.

Comparison across Grades

Results presented in Table 3 provide a comparative overview of classroom quality in science and mathematics across the target grades. The results follow the same pattern observed for the overall COS scores. Regardless of subject, the scores for all three grades fell within the range of weak practice (<1.5). Grade 5 showed a slight edge over the other grades; however, the differences were not statistically significant ($p>0.05$) in either subject.

Table 3: Overall quality of science and mathematics teaching across grades

Grade level	**Science**	**Mathematics**
Grade 5	1.32 (0.25)	1.31 (0.23)
Grade 6	1.26 (0.25)	1.25 (0.20)
Grade 8	1.27 (0.21)	1.29 (0.17)
Difference	$F(2, 283)=1.844$; $p>0.05$	$F(2, 300)=1.80$; $p>0.05$

Mean (Standard Deviation)

In other words, students had limited opportunities to participate in the teaching–learning process (e.g., asking questions, solving problems, working with peers) in science and mathematics classrooms across grades. That said, a scan of minimum and maximum scores showed that the pedagogical quality of a fraction of classrooms qualified for the mediocre band in all three grades: Grade 5

(Science=18%; Mathematics=16%), Grade 6 (Science=10%; Mathematics=12%), and Grade 8 (Science=13%; Mathematics=12%). Notably, the highest-scoring teachers were teaching Grade 5 in both science (M=2.38) and mathematics (M=2.08). While this fraction in the mediocre band is an encouraging starting point, the overall scores call for urgent attention from policymakers to enhance teachers' pedagogical skills through appropriate professional development opportunities.

Comparison Across System of School

Comparison of COS scores by school system revealed the same pattern observed for the overall mean COS scores in both science and mathematics, as presented in Table 4.

Table 4: classroom practices of science in public and private schools

School type	Science	Mathematics
Public	1.25 (0.20)	1.24 (0.17)
Private	1.31 (0.28)	1.33 (0.22)
Difference	$U = 14628.5$, $z = -1.34$, $p = 0.18$; n.s	$U = 21{,}807.5$, $z = 4.11$, $p<0.001$; r=0.21; r^2=0.04

Mean (Standard Deviation); ns = not significant; r= effect size; r^2=variation explained

On average, the COS scores for both school systems fell within the range of weak teaching practice (<1.5). Comparisons favoured private schools; however, only in mathematics was the difference statistically significant, with a small effect size (p<0.001; r=0.21). In other words, the school system explained 4% of the variation in the mathematics COS score. A scan of minimum and maximum scores showed that the pedagogical quality of a fraction of teachers qualified for the mediocre band in both systems: private (Science=18%; Mathematics=25%) and public (Science=9%; Mathematics=4%). The scoring pattern indicates that both systems included some teachers in both subjects who provided at least some opportunities for students to engage in the teaching–learning process (e.g., activities that provoked some thinking and supportive interactions). Notably, the highest-scoring classroom practices in both science (M=2.38) and mathematics (M=2.08) came from the private sector.

Comparison by Teacher Gender

Comparison of teaching quality by teacher gender revealed the same pattern observed for the overall mean COS scores in both science and mathematics, as presented in Table 5.

Table 5: Classroom practice of science and mathematics across gender

Teacher Gender	Science	Mathematics
Male	1.21 (0.18)	1.23 (0.16)
Female	1.32 (0.26)	1.32 (0.22)
Difference	$U = 18{,}212.0$, $z = 4.30$, $p<0.001$; r=0.23; r^2=0.05	$U = 21{,}428.0$, $z = 4.12$, $p<0.001$; r=0.21; r^2=0.04

Mean (Standard Deviation); ns = not significant; r= effect size; r^2=variation explained

Regardless of gender, the average COS score fell within the range of weak teaching practice (<1.5), as per the framework defined for this study. Comparisons of male and female teaching practices in science (p<0.001; r=0.23) and mathematics (p<0.001; r=0.21) favoured female teachers. Put differently, gender explained 5% and 4% of the variation in COS scores in science and mathematics, respectively. A scan of minimum and maximum scores showed that the pedagogical quality of a fraction of classrooms qualified for the mediocre band for both male (Science=5%; Mathematics=4%) and female teachers (Science=18%; Mathematics=21%). Notably, female teachers achieved the highest scores in both science (M=2.38) and mathematics (M=2.08).

DISCUSSION

We examined the teaching quality of mathematics and science teachers across Pakistan at the elementary level through full-lesson classroom observations. We also examined how teaching quality varies by grade level, teacher gender, and school system (public vs private).

The results showed that classroom practices are generally weak in both science and mathematics. Item-wise analysis revealed that some teachers had better interactions with their students while maintaining classroom discipline. However, overall, the quality of classroom infrastructure, teaching and learning practices, and classroom interactions was deemed weak in both subjects. Comparatively, science classrooms were slightly better equipped in terms of tidy space, displays, required materials, and, to some extent, the use of inquiry-based questioning. However, science teachers were less successful in engaging students in active classroom learning compared with mathematics teachers, possibly due to unclear instructions. We also found that private-school teachers performed better than government-school teachers across grades and genders in both science and mathematics, with a statistically significant difference for mathematics. In addition, female teachers were found to have better pedagogical practices regardless of subject. Grade 5 teachers' performance was relatively better than that of teachers in upper grades (Grades 6 and 8). The literature (e.g., Abreh et al., 2018; Banerjee, 2016; Bhutta et al., 2025) has frequently highlighted growing concern over inadequate pedagogical quality in science and mathematics.

Proposed reasons for weak pedagogical practices include passive lectures, a focus on syllabus completion, and exam preparation (Ansari, 2022; Gauthier et al., 2004; Kahiya & Brijlall, 2021; Lorenzen, 2017; Marusic & Slisko, 2012; Mwelese, 2014; Tsui, 2002).

A key concern is what prevents teachers from teaching effectively in the classroom. It is pertinent to note that recruitment policies differ between public and private schools. While the former often recruits teachers with both professional qualifications and subject master's degrees, the latter typically prioritises subject mastery. In either case, these differences do not appear to translate into substantial differences in observed practice. One reason may be that teachers hold strong beliefs shaped during their own schooling. Unlearning these beliefs remains a significant challenge, especially in a context where teacher training is in continuous experimentation and may make little difference (Ahmed, 2022; Ansari et al., 2024). Even when teachers receive training, what is taught in idealised contexts may not match the realities of public-sector classrooms. Although limited research has been conducted, evidence suggests that teachers who are trained may revert to traditional practices after some time (Westbrook et al., 2009) or face difficulties dealing with complexities and implementing student-centred pedagogies in the classroom (Burn et al., 2010; Ingersoll, 2012). Another explanation for low performance may be poor working conditions, which several studies (e.g., Ali, 2018; Borman & Dowling, 2008; Nguyen et al., 2019) have linked to teacher dissatisfaction. Weak classroom practices may also reflect a disproportionate focus on content knowledge and exam preparation, potentially reducing attention to pedagogical quality (Ansari, 2024; Lorenzen, 2017). This can lead teachers to prioritise syllabus completion and curriculum coverage over developing effective teaching strategies that engage students and facilitate learning. Furthermore, these factors may be compounded by traditional mindsets that emphasise knowledge transmission, limiting teachers' potential to improve their pedagogical approaches and enhance classroom practices (Teig et al., 2019). Meta-analytic evidence indicates that mathematics and science professional development can improve teacher knowledge and classroom instruction, and that larger teacher-level gains tend to be associated with larger improvements in student achievement (Lynch et al., 2025).

The study also shows that private-school teachers demonstrate relatively better classroom practices regardless of subject; however, the difference was statistically significant only for mathematics classroom practices. In line with previous research (Brady & Wilson, 2022; Huang et al., 2022; Qoyyimah, 2018; Taje & Goldring, 2020; Zamir et al., 2017), possible reasons include a preference for subject specialisation, better access to resources, and improved working conditions due to stronger monitoring and accountability mechanisms (Ayeni & Amanekwe, 2018; Nawab, 2020; Pillay et al., 2017; Shernoff et al., 2017; Singh et al., 2020), although private-school teachers receive lower salaries than public-school teachers (Ansari et al., 2026; Zafarullah et al., 2016). However, government schools tend to provide limited professional development opportunities (Nawab, 2020; Singh et al., 2020), and even where training is provided, its quality has been questioned. Moreover, private schools usually appoint subject-specific teachers with specialisation in science and mathematics, whereas government schools often appoint teachers based on general academic qualifications (Zafarullah et al., 2016). This suggests that specialist teachers in private schools may teach these subjects more effectively than their counterparts in government schools. Furthermore, private schools may offer a more supportive environment and implement systematic monitoring processes to ensure teaching standards (Afridi, 2018; Nawab, 2020; Siddiqui & Gorard, 2017). Accountability to parents, who often inquire about their children's academic performance and the quality of teaching, may motivate private schools to maintain higher teaching standards (Afridi, 2018; Niaz, 2017; Siddiqui & Gorard, 2017; Zuilkowski et al., 2018). Despite these differences, the findings of this nationwide study indicate that overall pedagogical quality in both systems remains weak.

Interestingly, early-grade teachers generally exhibited better pedagogical practices than their counterparts in upper grades. These findings are consistent with previous research (Ainley et al., 2020; Cleary & Chen, 2009; Rosario et al., 2013; Saeed et al., 2005; Teig et al., 2022), which found a decline in pedagogical practices across grades (5 > 6 > 8). This trend can be attributed to factors such as teachers' focus on pedagogy versus content, teacher interaction aimed at leading versus directing students, and class size, which tends to increase in upper grades (Martin et al., 2012; OECD, 2016, 2018; Teig et al., 2022). A similar pattern of student-teacher ratios (30:1 and 38:1) was observed in lower and upper grades, respectively (Bhutta & Rizvi, 2022). Grade progression also demands more advanced concepts in science and mathematics, which require subject-specific vocabulary and abstract thinking (Genc & Erbas, 2019; Groves, 2016). This can make it challenging for teachers to balance content delivery and pedagogical quality, often resulting in a focus on content over pedagogy (Fredricks et al., 2018). Yang et al. (2020) argued that teachers in lower grades tend to emphasise pedagogical quality and connect concepts with real-life scenarios more easily, which accelerates student participation. This suggests that lower-grade teachers may prioritise helping students understand and relate to content, whereas upper-grade teachers may focus more on delivering information.

In addition, class size plays a crucial role in teachers' classroom practices. Lower-grade teachers tend to have smaller classes that allow more individual support, while class sizes increase as students progress to higher grades (Blatchford et al., 2011; Blatchford & Russell, 2019; Teig et al., 2022). A similar trend was seen in the sampled schools, with student–teacher ratios of 30:1 in lower grades and 38:1 in upper grades. This may reflect the assumption that older students are more mature and independent learners who require less individual support than younger students (Lau et al., 2018). However, the absence of necessary support can lead to reduced student involvement and achievement, ultimately affecting pedagogical quality.

The analysis also revealed that female teachers, on average, demonstrated better pedagogical quality than their male counterparts. These findings are consistent with the literature (e.g., Coffey & Farinde, 2016; Demetriou et al., 2009; El-Emadi et al., 2019; Hussain et al., 2017; Nadeem et al., 2011), which found that female teachers performed better than male teachers in classroom teaching practices. Proposed reasons include differences in interaction style (supportive, caring, and respectful), choice of pedagogical practices, and attitudes towards the teaching profession (Coffey & Farinde, 2016; Demetriou et al., 2009; El-Emadi et al., 2019; Islahi & Nasreen, 2013). Female teachers are often described as nurturing and expressive, and they may be more likely to encourage student participation and collaboration, whereas male teachers may be seen as more dominant, structured, and task-oriented (Rashidi & Naderi, 2012; Wood, 2009). Male teachers may also prefer lectures and limit student involvement, while female teachers may prefer more collaborative and active learning methods (Islahi & Nasreen, 2013). These differences may relate to underlying differences in communication styles (Mullola et al., 2012; Nelson Laird et al., 2007). It is possible that female teachers build stronger rapport with students by showing respect and developing supportive relationships, whereas male teachers may be perceived as more controlling, which can reduce students' comfort and participation (Demetriou et al., 2009; Hussain et al., 2017; Smith, 2010). Cultural norms may also shape teacher-student relationships and classroom dynamics. Overall, the association between gender and teaching practices suggests that female teachers may have a greater capacity to empathise and connect with students; consequently, they may be preferred for the profession in some contexts (Demetriou et al., 2009; Islahi & Nasreen, 2013). These findings suggest that female teachers may be more likely to engage students in classroom activities and support their learning.

CONCLUSION

In this study, we captured teachers' classroom practices across Pakistan through full-lesson observations. Teaching performance in both science and mathematics was generally weak. However, private-sector teachers performed relatively better than public-sector teachers. Female teachers demonstrated stronger pedagogical practices than their male counterparts. We suggest that this difference may be related to more supportive teacher–student interactions, which may encourage students to share concerns and participate. In lower grades, pedagogy appears to be more prominent; however, as grade level increases, teachers tend to focus more on content.

This study is limited in that it does not explain how pedagogical quality is linked to student outcomes. A causal link still needs to be established through research that examines teachers' background characteristics (e.g., experience, qualifications, training), teaching quality (through observation), and student achievement. This would help bridge the missing link between teachers' backgrounds, what they do in the classroom, and the effects on student development.

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

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

REFERENCES

Abreh, M. K., Owusu, K. A., & Amedahe, F. K. (2018). Trends in performance of WASSCE candidates in the science and mathematics in Ghana: Perceived contributing factors and the way forward. *Journal of Education*, *198*(1), 113-123.

AEPAM. (2018). Pakistan Education Statistics, 2017-18. Ministry of Federal Education and Professional Training, Government of Pakistan.

Afridi, M. (2018). Equity and quality in an education public-private partnership: A study of the World Bank-supported PPP in Punjab, Pakistan.

Akiba M., Liang G. (2014). Teacher Qualification and the Achievement Gap: A Cross-National Analysis of 50 Countries. In: Clark J. (eds) *Closing the Achievement Gap from an International Perspective*. 21-40. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-4357-1_3

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher*, *36*(7), 369-387. https://doi.org/10.3102/0013189x07308739

Ansari, A. N. (2022). *Exploring Social Capabilities and Life Skills of Students and the Role of School in Promoting the Same: A Mixed-Methods Study* [Unpublished Master's Thesis]. Aga Khan University.

Ansari, A. N., Bhutta, S. M., & Ahmad, S. (2024). Teacher Professional Education and Student Learning: A Secondary Analysis of Nationwide Data. *Journal of Contemporary Trends and Issues in Education*, *3*(2), 80-104. https://doi.org/10.55628/jctie.v3i2.138

Ayeni, A. J., & Amanekwe, A. P. (2018). Teachers' instructional workload management and students' academic performance in public and private secondary schools in Akoko North-East Local Government, Ondo State, Nigeria. *American International Journal of Education and Linguistics Research*, *1*(1), 9-23.

Banerjee, P. A. (2016). A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. *Cogent Education*, *3*(1), 1178441.

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*, 1-4.

Blatchford, P., & Russell, A. (2019). Class size, grouping practices, and classroom management. *International Journal of Educational Research*, *96*, 154-163.

Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher–pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and instruction*, *21*(6), 715-730.

Brady, J., & Wilson, E. (2022). Comparing sources of stress for state and private school teachers in England. *Improving Schools*, *25*(2), 205-220.

Coffey, H., & Farinde-Wu, A. (2016). Navigating the journey to culturally responsive teaching: Lessons from the success and struggles of one first-year, Black female teacher of Black students in an urban school. *Teaching and Teacher Education*, *60*, 24-33.

Demetriou, H., Wilson, E., & Winterbottom, M. (2009). The role of emotion in teaching: Are there differences between male and female newly qualified teachers' approaches to teaching?. *Educational studies*, *35*(4), 449-473.

El-Emadi, A. A., Said, Z., & Friesen, H. L. (2019). Teaching style differences between male and female science teachers in qatari schools: Possible impact on student achievement. *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(12), em1800.

Fraenkel, J., Hyun, H., & Wallen, N. (2011). *How to design and evaluate research in education*. McGraw-Hill Education.

Fredricks, J. A., Hofkens, T., Wang, M. T., Mortenson, E., & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, *55*(2), 271-298.

Gauthier, C., Dembélé, M., Bossonnette, S., & Richard, M. (2004). Quality of teaching and quality of education: A review of research findings. *UNESCO. Retrieved online February*, *5*, 2005.

Genc, M., & Erbas, A. K. (2019). Secondary Mathematics Teachers' Conceptions of Mathematical Literacy. *International Journal of Education in Mathematics, Science and Technology*, *7*(3), 222-237.

Gorard, S. (2003). *Quantitative methods in social science research: The role of numbers made easy* (1st Ed.). Continuum: New York.

Government of Pakistan. (2016). *National assessment reports 2016.* Islamabad: National Education Assessment System (NEAS): Ministry of Federal Education & Professional Training.

Groves, F. H. (2016). A longitudinal study of middle and secondary level science textbook vocabulary loads. *School Science and Mathematics*, *116*(6), 320-325.

Huang, X., Lai, C., Wang, C., & Xu, G. (2022). The relationship between informal teacher learning and self-efficacy among public school teachers and private school teachers in China. *Journal of Education for Teaching*, 1-19.

Hussain, S., Farooq, R. A., Mahmood, Z., & Tabassam, R. (2017). Comparison of Male and Female Teachers' Performance Working at Elementary Level in Punjab. *Journal of Elementary Education*, *27*(2), 129-138.

Islahi, F., & Nasreen, N. (2013). Who Make Effective Teachers, Men or Women? An Indian Perspective. *Universal Journal of Educational Research*, *1*(4), 285-293.

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
Kahiya, A., & Brijlall, D. (2021). What are the strategies for teaching and learning mathematics that can be used effectively in a multilingual classroom? *Technology Reports of Kansai University; Vol. 63, Issue 5*.
Lau, C., Kitsantas, A., Miller, A. D., & Drogin Rodgers, E. B. (2018). Perceived responsibility for learning, self-efficacy, and sources of self-efficacy in mathematics: a study of international baccalaureate primary years programme students. *Social Psychology of Education*, *21*, 603-620.
Lorenzen, J. K. (2017). *The effect of instructional strategies on math anxiety and achievement: A mixed methods study of preservice elementary teachers* (Doctoral dissertation, The University of Southern Mississippi).
Martin, M. O., Mullis, I. V., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 International results in science*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.
Marušić, M., & Sliško, J. (2012). Influence of three different methods of teaching physics on the gain in students' development of reasoning. *International Journal of Science Education*, *34*(2), 301-326.
Mullola, S., Ravaja, N., Lipsanen, J., Alatupa, S., Hintsanen, M., Jokela, M., & Keltikangas-Järvinen, L. (2012). Gender differences in teachers' perceptions of students' temperament, educational competence, and teachability. *British Journal of Educational Psychology*, *82*(2), 185-206.
Muralidharan, K., & Sheth, K. (2016). Bridging education gender gaps in developing countries: The role of female teachers. *Journal of Human Resources*, *51*(2), 269-297.
Mwelese, J. K. (2014). Effectiveness of Classroom Practices to Achievement in Mathematics. *Journal of Education and Practice*, *5*(25.64-73).
Nadeem, M., Rana, M. S., Lone, A. H., Maqbool, S., Naz, K., & Ali, A. (2011). Teacher's competencies and factors affecting the performance of female teachers in Bahawalpur (Southern Punjab) Pakistan. *International Journal of Business and Social Science*, *2*(19), 217-222.
Naviwala, N. (2016). Pakistan's Education Crisis: The Real Story. *Woodrow Wilson International Center for Scholars*.
Nawab, A. (2020). Monitoring and accountability in professional development of teachers in Rural Pakistan. *Journal of International and Comparative Education (JICE)*, 77-89.

Nelson Laird, T. F., Garver, A. K., & Niskodé, A. S. (2007). Gender gaps: Understanding teaching style differences between men and women. Association for Institutional Research Annual Forum.

Niaz, A. (2017). *An exploration of maternal support to children's academic achievement in low-cost private schools in Lahore* (Doctoral dissertation, University of Oxford).

Pillay, H., Muttaqi, I. A., Pant, Y. R., & Herath, N. (2017). *Innovative strategies for accelerated human resource development in South Asia: Teacher professional development-Special focus on Bangladesh, Nepal, and Sri Lanka*. Asian Development Bank.

Qoyyimah, U. (2018). Policy implementation within the frame of school-based curriculum: a comparison of public school and Islamic private school teachers in East Java, Indonesia. *Compare: A Journal of Comparative and International Education*, *48*(4), 571-589.

Rashidi, N., & Naderi, S. (2012). The effect of gender on the patterns of classroom interaction. *Education*, *2*(3), 30-36.

Robson, C., & McCartan, K. (2016). *Real World Research* (4th ed.). Wiley.

Ryu, J. H., & Sajjad, S. (2018). Quality of teachers' performance evaluation in public and private secondary schools of Karachi, Pakistan. *Pakistan Journal of Education*, *35*(2).

Shaukat, S., & Chowdhury, R. (2020). Teacher educators' perceptions of professional standards: Implementation challenges in Pakistan. *Issues in Educational Research*, *30*(3), 1084-1104.

Shernoff, D. J., Sinha, S., Bressler, D. M., & Ginsburg, L. (2017). Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. *International Journal of STEM Education*, *4*, 1-16.

Siddiqui, N., & Gorard, S. (2017). Comparing government and private schools in Pakistan: The way forward for universal education. *International Journal of Educational Research*, *82*, 159-169.

Singh, A. K., Rind, I. A., & Sabur, Z. (2020). Continuous professional development of school teachers: Experiences of Bangladesh, India, and Pakistan. *Handbook of education systems in South Asia*, 1-27.

Smith, D. J. (2010). *The use and perceived effectiveness of instructional practices in two-year technical colleges* (Doctoral dissertation, uga).

Taie, S., & Goldring, R. (2020). Characteristics of Public and Private Elementary and Secondary School Teachers in the United States: Results from the 2017-18 National Teacher and Principal Survey. First Look. NCES 2020-142. *National Center for Education Statistics*.

Teig, N., & Nilsen, T. (2022). Profiles of instructional quality in primary and secondary education: Patterns, predictors, and relations to student achievement and motivation in science. *Studies in Educational Evaluation*, *74*, 101170.

Teig, N., Scherer, R., & Olsen, R. V. (2022). A systematic review of studies investigating science teaching and learning: over two decades of TIMSS and PISA. *International Journal of Science Education*, *44*(12), 2035-2058.

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

TIMSS. (2020). *TIMSS 2019 International Results in Mathematics and Science*. International Association for the Evaluation of Educational Achievement (IEA). https://timss2019.org/reports/wp-content/themes/timssandpirls/download-center/TIMSS-2019-International-Results-in-Mathematics-and-Science.pdf

Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *The journal of higher education*, *73*(6), 740-763.

Wood, T. D. (2009). *Teacher perceptions of gender-based differences among elementary school teachers*. Saint Louis University.

Yang, X., Kuo, L. J., & Jiang, L. (2020). Connecting theory and practice: A systematic review of K-5 science and math literacy instruction. *International Journal of Science and Mathematics Education*, *18*, 203-219.

Zafarullah, S., Mumtaz, K., Murad, P. U., Abida, S., & Humera, S. (2016). Teachers' time management and the performance of students: A comparison of government and private schools of Hyderabad, Sindh, Pakistan.

Zamir, S., Arshad, M., & Nazir, N. (2017). A comparative study of self-efficacy of public and private school teachers at elementary level. *Journal of Elementary Education*, *27*(1), 23-36.

Zuilkowski, S. S., Piper, B., Ong'ele, S., & Kiminza, O. (2018). Parents, quality, and school choice: Why parents in Nairobi choose low-cost private schools over public schools in Kenya's free primary education era. *Oxford Review of Education*, *44*(2), 258-274.

Bio

SADIA MUZAFFA BHTTA, 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

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

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

AFAQ AHMED, is a PhD Scholar at Department of Teaching, Learning, and Culture, Texas A&M University, United States. Alongside this, he has served as a Research Associate at the Aga Khan University, Institute for Educational Development, Karachi, Pakistan. His major research interests include teacher education, teaching license, and educational policy. Email: afaqahmed.edu@gmail.com