

**ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL**Available Online: <https://assajournal.com>

Vol. 05 No. 01. Jan-March 2026. Page#.6689-6705

Print ISSN: [3006-2497](#) Online ISSN: [3006-2500](#)Platform & Workflow by: [Open Journal Systems](#)**A Systematic Review: Investigating the Effect of Teachers' Professional Development on Science Students' Achievement****Bushra Khalid**M.Phil. Education, Department of Education, National University of Modern Languages Islamabad,
PakistanEmail: bushrakhalid278@gmail.com**ABSTARCT**

This systematic review aims to uncover the impact of professional development on the academic performance of science students in the literature published from 2002 to 2025. A systematic search was carried out in the major electronic databases, such as Scopus, Web of Science, ERIC, and Google Scholar, with specific search terms designed around the concepts and items related to professional development, science education, teacher quality, instructional practices, and student achievement. The findings show that professional development is a positive tool for teachers' professional development, teaching practices, and dedication, and enhances the quality of science teaching. Moreover, the review shows that there are three steps in the process whereby professional development leads to student achievement, and these are all interconnected: a strengthening of teachers' knowledge and professional competencies, a strengthening of their classroom instructional practices, and a strengthening of students' academic achievement. Contextual and sustained professional learning also help support teachers in implementing effective teaching and addressing evolving needs of science teaching. Although the majority of the reviewed studies reported benefits for teacher effectiveness and student achievement, some reported limited or non-significant benefits, suggesting that professional development is more effective when designed and implemented well, is of sufficient length, is tailored to the educational context, and is actively engaged in by teachers. Overall, this review provides a comprehensive review of evidence spanning over the last 20 years and demonstrates the importance of continued teacher professional development as a means to improve teacher quality and student achievement in science education. The need for the development of long-term, evidence-based, context-specific professional development programmes to enhance teachers' professional competence, teaching quality, and learning outcomes is highlighted.

Keywords; Professional Development, Science Student, Academic Achievement, Systematic Review.

Introduction

Professional development (PD) is the main policy tool for improving teacher quality after teachers have already entered classrooms (Darling-Hammond et al., 2017), and the quality of teachers is found to be one of the strongest school-based determinants of student learning (Lovett & Graham, 2015). The implementation of effective science instruction demands a high level of both disciplinary science knowledge and pedagogical knowledge, as well as the ability to make scientific concepts concrete as student investigations, reasoning, and constructing meaning (Gess-Newsome et al., 2019). Given the many curricular shifts, changes in assessment, and developments in scientific knowledge that a teacher must face over a career, continuing professional learning has emerged as a key component of efforts to improve science achievement in education systems (Desimone, 2009). For more than 20 years, researchers, evaluators, and policymakers have studied how teachers' professional learning relates to measurable student learning outcomes, resulting in a vast and growing body of evidence that is synthesized in this review.

The modern research programme on teacher PD and science achievement can be traced back to the science education reform movement of the late 1990s and early 2000s, which focused on heavy-duty, sustained professional development that was tied to specific content standards and integrated into a system of reform (Supovitz & Turner, 2000). The National Science Foundation (NSF) in the United States has made this reform vision operational on a national level with its Local Systemic Change through Teacher Enhancement programme and early outcomes of this programme showed a positive correlation between the amount of teachers' involvement in the programme and their use of inquiry-based instructional practices and investigative classroom culture (Banilower, 2002). Around the same time, a large-scale, national probability study of mathematics and science teachers that examined the impact of various design features of PD activities found that the strongest effects on teacher knowledge and classroom practice were associated with PD activities that focused on content knowledge, included opportunities to actively engage in learning, and were coherent with other work teachers were doing in their classrooms (Garet et al., 2001). A three-year follow-up study with the same national sample validated the benefits of PD in increase subsequent use of the practices and found that the benefits were amplified when PD included opportunities for active learning (Desimone et al., 2002).

In the middle years of the 2000s, as evidence increased, researchers started to make sense of the evidence they had found and to pinpoint areas that needed improvement in both PD practice and PD research. Guskey (2002) contended that most of those PD programmes of the time were not effective in altering teachers' beliefs and classroom practice because they ignored the nature of teachers' change and suggested that one way to charter a successful path for PD was to focus on student learning outcomes as the key indicators, not teacher satisfaction, of the effectiveness of the PD program. Borko (2004) noted that most professional development offered to teachers was "silo-mentality" work, that is, focused on a single location and program, and suggested the need for research that would extend beyond one-site and one-programme studies to multi-site and cross-context investigations of teacher learning. In science education, a study of the Local Systemic Change over the course of several years revealed that systematic, content-focused professional development has the potential to support the national vision of science education standards and that there were measurable differences in teachers'

instruction after implementing the LSC (Banilower et al., 2007). At this time, the most expansive review of studies conducted included only those studies that met rigorous causal-design criteria, and was limited to three areas of content (reading, mathematics, and science), but identified a moderate and consistent average effect size across the three content areas of PD on student achievement (Yoon et al., 2007).

Theories that underlie PD practice today have been developed in the years following 2009. Desimone (2009) offered an influential conceptual framework of five fundamental characteristics of effective PD—content focus, active learning, coherence, sufficient duration, and collective participation—that impact teachers' knowledge and beliefs, which in turn influence instructional practice and ultimately student learning. Professional learning research has come a long way since 2000, but there is still much that is not known about the sustainability of the effects of PD over time, according to a synthesis of a decade of PD research published in the journal *Teaching and Teacher Education* (Avalos, 2011). During this time, additional achievement evidence was collected for science; a multi-year, K-6, science-specific PD programme was tested with over 1,200 elementary teachers and showed a significant positive correlation between the number of hours' teachers received in PD and the students' gains on a high-stakes science test, though this was inconsistent across grade-levels (Shymansky et al., 2012). But not all the evidence from this era was positive: a rigorous federally-funded randomized evaluation of an intensive, two-year, content-focused PD programme for middle-school mathematics teachers showed that the programme influenced some observed aspects of teacher instructional practice, but did not statistically affect teacher content knowledge or student achievement, a first and important sign of the question of what works and what doesn't in PD.

The research literature evolved toward more attention to the methodology of PD and a more focused approach to the knowledge being treated during the PD period. Over the next decade, additional U.S. research informed a refinement of the five features and confirmed the importance of those features, as well as several large randomized trials that used just these features and yet yielded disappointing results (Desimone & Garet, 2015). This is expanded upon by Kennedy (2016) who argued that grouping PD programmes together based on the structural characteristics of the programme – its length, format etc. – masked the factors that were actually linked to programme effectiveness, and suggested that PD programmes should be grouped based on the theory of teacher learning. Kennedy's synthesis of studies with high design standards revealed that a large number of popular features of PD programmes were not, in isolation, linked to PD programme effectiveness. This period has demonstrated one of the most direct connections between science teachers' professional knowledge and student outcomes in science: a two-year PD intervention study which measured the academic content knowledge, general pedagogical knowledge, and pedagogical content knowledge (PCK) separately yielded the finding that teachers' gains were observed in all three constructs, with the strongest gains observed in academic content knowledge, and that the gains made in the former construct showed the greatest predictive power for the subsequent student outcomes (Gess-Newsome et al., 2019).

Randomized trials of large sample sizes, systematic reviews, and meta-analyses have dominated the past six years (2020-2025) of research, trying to reconcile the inconsistent findings from the preceding 20 years. In a four-arm randomized controlled design study of a pedagogy-focused, classroom-

observation-based model of PD, the gains of the students of participating teachers were significantly greater than those of a wait-list control group in mathematics with similar findings in reading and science (Gore et al., 2021). The same systematic review of PD research that examined 125 studies, published at approximately the same time, found that training continued to be the most popular type of PD provided to teachers, but only eleven studies directly tested the impact of PD on student learning, and focused primarily on mathematics and literacy (Ventista & Brown, 2023). Integrated STEM teaching has become increasingly common, and there is also a growing body of research focused on systematic review of K–12 STEM teacher professional development programmes, which catalogued the knowledge foci, typical design approaches, and outcome measures in a large and growing international literature (Huang et al., 2022). Recently, large-scale meta-analytic syntheses have been conducted to quantify these relationships directly: A meta-analysis of experimental studies on mathematics and science PD interventions showed that interventions that led to increased teacher knowledge and classroom instruction were associated with increased achievement for students (Lynch et al., 2025); A parallel meta-analysis of experimental studies focused on mathematics teachers confirmed an overall positive though variable relationship between PD participation and student achievement in mathematics classrooms (Franklin & Chang, 2025); and new evidence using both text-based and video-based assessments of science teachers' pedagogical content knowledge confirmed PCK as a strong and specific predictor of student achievement in science classrooms (She et al., 2025).

As a whole, this 22-year research effort offers a field that has become increasingly sophisticated in the methods used, and yet has not been unified in its findings. Equally, there is strong, positive, and increasingly precise evidence linking PD to teacher knowledge, teacher practice, and student achievement, alongside methodologically rigorous studies with weak, inconsistent, or null effects – sometimes in the same content areas or with similarly rigorous designs. This does not, by itself, prove the failure of PD, but rather indicates that the PD–achievement link is contingent on factors that have not always been measured and reported across studies. What is needed next is a systematic, chronologically exhaustive synthesis that connects over 20 years of evidence specifically on PD's effects on science achievement and identifies conditions under which PD effects are likely to occur. This review addresses that need.

Objectives

1. To examine the effect of teachers' professional development on students' science achievement based on research published between 2002 and 2025.
2. To examine how teachers' professional development effects teachers' professional knowledge and instructional practices and how these contribute to students' science achievement.
3. To identify the programme design, implementation, and contextual factors that influence the effectiveness of teachers' professional development in improving students' science achievement.

Research Questions

RQ1: Based on research published from 2002 to 2025, what is the impact of teachers' professional development on students' science achievement?

RQ2: How does teachers' professional development influence teachers' professional knowledge and practices, and students' achievement in science?

RQ3: What aspects of programme design and implementation, and contextual factors, affect the effectiveness of teachers' professional development in enhancing students' science achievement?

Significance of the Study

This review is of importance for various audiences. The review provides a data-informed basis for policy and ministry decisions regarding the nature of PD funding and mandates, in an era when significant public investments are being made in PD for teachers without a clear commitment to improved student outcomes (Desimone & Garet, 2015). For anyone designing and/or reviewing PD, this literature review summarizes the design elements that have been most consistently linked to the effectiveness of PD: content focus, active learning, coherence, duration, and collective participation (Desimone, 2009; Garet et al., 2001). Pedagogical content knowledge is highlighted as the specific link between PD and science achievement for teacher educators, and as such, has direct implications for the design of PD programmes for science teachers (Gess-Newsome et al., 2019; She et al., 2025). For the research community, the review is a comprehensive, paced summary of more than 20 years of research, which brings together results so far that have been reported in different strands of literature dealing with mathematics PD, general teachers' PD, or science-specific PD, and pinpoints gaps in the evidence base.

Research Gap

Although all the research cited above has been done, there are still several areas that are not covered. Much of the strongest evidence of achievement comes from mathematics PD research or from a combined mathematics and science sample (Garet et al., 2011, Franklin & Chang, 2025; Lynch et al., 2025), which means that conclusions specific to science instruction are not drawn from science-only samples. Second, while some previous systematic reviews of teacher PD have been limited to teacher PD broadly with no disaggregation by the science area (Ventista & Brown, 2023), others have focused on a specific science design feature, science subject, or single science intervention model (e.g., Huang et al., 2022 for STEM programmes specifically), there has been no previous review that chronologically traced the science-specific evidence from 2002 to 2025. Third, the field is lacking in a synthesis that explicitly links the positive findings reported by large-scale meta-analyses (Lynch et al., 2025; She et al., 2025) with the null or limited findings reported by methodologically comparable randomized trials (Garet et al., 2011), with the latter version of PD treated as a uniform intervention. Last but not least, the evidence available is largely from well-resourced education systems, with comparatively little synthesized evidence dealing with how these mechanisms function in the context of science education in much of the developing world, which is often characterized by a high pupil load in low-resource environments. This review aims to fill these gaps by providing a science-focused synthesis of the literature from 2002 to 2025 that explicitly considers the science mechanisms and conditions that explain how PD affects science achievement.

Literature Review

Teacher professional learning research and the science learning outcomes of students can be broadly grouped into four time periods, each with its own methodological or conceptual focus: (1)

foundational reform-era research (early to mid-2000s); (2) consolidation of conceptual frameworks (2007–2015); (3) refined research methods and increased attention to pedagogical content knowledge (2016–2019); and (4) the recent period of large-scale meta-analyses and randomized trials (2020–2025).

Foundational Research and the Science Education Reform Movement (Early-to-Mid 2000s)

The first studies examined here developed directly from the U.S. science education standards movement that envisioned intensive, sustained, content-focused, and systemic science professional development (Supovitz & Turner, 2000). At national scale, this vision was empirically tested in the National Science Foundation's Local Systemic Change through Teacher Enhancement and initial findings in this program showed a strong relationship between the quantity of teacher professional development that teachers received and their use of inquiry-based teaching practices as well as the establishment of an investigative classroom culture (Banilower, 2002). Concurrently, Garet et al. (2001), in a national probability sample of 1,027 mathematics and science teachers, identified for the first time at this scale, which specific features of PD activities predicted teacher learning; their regression analyses resulted in three features found to have significant independent effects: content focus, active learning, and coherence, with the latter two found to operate through the first, rather than independently. Desimone et al. (2002) followed this up with a three-year longitudinal study and found that this PD, involving specific instructional practices, facilitated teachers' subsequent use of the practices, and that active learning opportunities had a larger impact. During this time, Guskey (2002) provided a seminal conceptual critique of PD, arguing that instead of measuring PD by teacher reactions, it more often was measured by teacher satisfaction and proposing a five-level model of PD evaluation that would take a different path toward measuring PD by evidence of changed practice and student learning outcomes. Borko (2004) was able to compile a mid-decade review and found that most existing PD was still intellectually shallow and siloed, and that research needed to go beyond single site case studies and look at comparative multi-site studies to determine what aspects of PD were common across settings.

Consolidating Frameworks: Core Features and Conceptual Models (2007 – 2015)

In the second half of the 2000s, sufficient robust evidence had accrued to the field to enable systematic evidence reviews. Although the number of studies the found was low with only nine studies out of over 1,300 that met the rigor standards of the What Works Clearinghouse, they did find that the average effect size for PD on student achievement was moderate, supporting the overall conclusion that well-implemented PD could have a negative impact on student achievement. In science, an assessment of subsequent cohorts of the Local Systemic Change initiative revealed that the visionary work in the standards for science education had not yet been implemented in classrooms, which was brought closer to implementation the sustained, content-focused PD activities (Banilower et al., 2007). Desimone (2009) then formalized the conceptual model that has served as a foundation for the field since: PD's five core features (content focus, active learning, coherence, duration, and collective participation) influence teacher knowledge and beliefs, which influence instructional practice, which in turn influences student learning, which is moderated by contextual factors associated with students, teachers, and schools. In addition to this theoretical synthesis, Avalos (2011) carried out a stock-taking of a decade of PD scholarship in a top-tier journal and found that the field had come a long way in its understanding of learning, facilitation and collaboration, but left unresolved questions about the sustainability of PD

effects over time. During this time period, direct evidence specific to science achievement continued to grow: Shymansky et al. (2012) reported statistically significant positive associations between the number of hours' teachers spent on PD and the gains of students on high-stakes science assessments, and they also documented grade band differences in the relationships between PD hours and students' science assessment gains. There was some mixed news, however, from a large, federally-funded randomized controlled trial of an intensive two-year mathematics teacher PD programme, which reported gains in certain observed mathematics teacher behaviours, but not in content knowledge or student achievement (Garet et al., 2011) an early and well-designed counter-example that would later require review to be accommodated alongside the widely positive findings reported elsewhere. These five core PD features were given credibility by Desimone and Garet (2015) who directly addressed the tension and confirmed the importance of the five components of PD, but noted that many large randomized trials had been conducted including the five core PD features yet had yielded disappointing results and suggested that a greater focus on implementation quality rather than programme design alone would be more appropriate.

The cycle of developing the Evidence Base: Pedagogical Content Knowledge and Methodological Scrutiny (2016-2019)

During the period 2016-2019, research has addressed some of the methodological weaknesses found in the preceding years with more focused studies and improved methodology on the specific professional knowledge that should be developed by PD. Kennedy (2016) suggested that the way programmes are classified by the structural characteristics of the PD (e.g. length, format, coaches) is unnecessary as it hides the underlying rationale that leads to programme effectiveness; in her review of 28 rigorously designed studies, she found that many of the popular characteristics of PD programmes were not consistently associated with the success of the programmes, and the specific PD pedagogical strategies used to assist teachers in implementing the new ideas did differentiate between effective and ineffective programmes. In the realm of science education, Gess-Newsome et al. (2019) offered one of the most direct piece of causal evidence available on the mechanism linking PD to achievement outcomes: academic content knowledge, general pedagogical knowledge, and pedagogical content knowledge (PCK) were measured as separate constructs, and there were significant gains in all three constructs after the intervention, with academic content knowledge being the construct most strongly related to subsequent student achievement, and general pedagogical knowledge most strongly related to observed instructional practice. Likewise, Taylor et al. (2017) reported convincing experimental data from a cluster randomized trial of a science professional development program used in 77 schools, 144 teachers, and 2823 students. Analysis of the data showed that students in the analysis of practice programme classrooms had significantly greater achievement gains than students in classrooms of teachers in the content-deepening comparison programme, indicating that linking content to analysis of practice is important.

Darling et al., (2017) compiled 35 methodologically rigorous studies across a variety of subject areas and concluded that effective PD is extended over time, is content-focused, active, collaborative, and includes models of practice; that a number of the more common PD programs found in U.S. schools

did not possess these elements and therefore were unlikely to create greater gains in student achievement than programme designed with these features.

Recent Evidence: STEM Programmes, Randomized Trials, and Meta-Analyses (2020-2025)

In the most recent period of research, larger samples, more rigorous experimental designs, and systematic attempts to quantify PD's effects have been features of many studies. Gore et al. (2021) carried out a four-arm randomized controlled trial of a classroom-observation-based PD model with over 5,000 elementary students, in which students of participating teachers significantly outperformed a wait-list control group in mathematics with parallel outcomes measured in reading and science. Ventista and Brown (2023) undertook a wide systematic review of 125 studies of teacher professional learning, though they identified only 11 that had experimental or quasi-experimental designs, suitable for direct testing of the impact of professional learning on student learning, and determined that when professional learning was continuous and delivered over a longer period of time, there seemed to be a stronger and more consistent relationship between professional learning and student outcomes, compared to a single event. More recent experimental work, however, indicates that professional development does not always lead to student achievement gains. A large cluster randomized trial with 89 schools, 253 teachers and 20,591 students was performed by Kramer et al. (2024). The study results indicated that, in general, teacher professional development that increases the teachers' content knowledge did not result in significantly better science student achievement, while teacher professional development combined with the curriculum intervention that included principles from the cognitive sciences did result in a small but significant positive effect. The results of this study indicate that programme design is important and that it may be necessary to build up the knowledge of teachers if there are to be improvements in learning outcomes. A parallel strand of review literature emerged, which has been structurally mapped by Huang et al. (2022) by systematically summarizing the knowledge context, pedagogical approaches, and technology in the K–12 STEM teacher PD programmes from all around the world, the most frequent skill focus of these programmes, the most common implementation methods, and the most common measures used to assess their effectiveness. Meta-analytic syntheses have recently attempted to quantify the inconsistencies of the field, most recent as a way to address them most recently. In an effort to synthesize 46 experimental studies of PK–12 mathematics and science PD interventions, Lynch et al. (2025) also found that the teacher-level outcomes from the treatment groups were significantly more positive than the outcomes from the control groups, and that programmes with larger effects on teacher knowledge and instructional practice tended to show larger effects on student achievement, thereby providing direct meta-analytic support for the teacher to classroom to student pathway proposed by Desimone (2009). A similar result was found by Franklin and Chang (2025) in a meta-analysis that examined only mathematics teacher PD, which found an overall positive relationship between the programme and students' achievement, noting significant differences in terms of effect size by programme type. Last, she et al. (2025) carried forward the work of Gess-Newsome et al. (2019) on PCK forward and established a strong relationship between science teachers' PCK and their students' achievement by employing both text- and video-based PCK assessments, confirming the importance of PCK as a specific and measurable pathway for science-focused PD to impact learning outcomes.

The literature studied during each of these four periods collectively indicates a field that has progressed from a broad, reform-influenced movement on PD (2000s), to the development of a set of conceptual frameworks that describe the nature of effective PD (2007-2015), to more specific attempts to measure the mechanisms by which PD impacts science achievement (2016-2019), and finally to attempts at large-scale, quantitative analyses that reconcile decades of inconsistent findings (2020-2025). Throughout all four periods, there is a consistent theme: PD which is content-focused, sustained, coherent, collaborative and well implemented is consistently linked with better teacher and student outcomes, especially when rigorously evaluated; PD that is not content-focused, or is not sustained, coherent, collaborative, or implemented well, is often not linked to measurable achievement gains.

Methodology

This review is based on the systematic application of the Preferred Reporting Items for Systematic Reviews and meta-analysis (PRISMA) guidelines for identification, screening and synthesis of relevant literature in a transparent way (Page et al., 2021). The electronic databases used to search for literature include Scopus, Web of Science, ERIC, and Google Scholar and covered studies published between 2002 to 2025.

The terms used in the search were developed based on four conceptual clusters and were connected using Boolean operators: (a) professional development terms (e.g., “teacher professional development”, “in-service training”, “continuing professional development”); (b) science education terms (e.g., “science education”, “science teaching”, “STEM instruction”); (c) teacher quality and practice terms (e.g., “teacher quality”, “pedagogical content knowledge”, “instructional practice”); and (d) outcome terms (e.g., “student achievement”, “academic achievement”, “learning outcomes”).

Studies were only included if they (a) were published from 2002 to 2025 in peer-reviewed journals, technical reports, or reputable research-institute publications; (b) focused on a professional development, in-service training, or teacher-learning intervention that was relevant to science (or mathematics and science together, given the substantial overlap in the PD research base); and (c) reported an empirical or synthesized relationship between PD and teacher knowledge, instructional practice, or student achievement. Those studies that were only preservice teacher education studies, or which didn't report an outcome related to teaching or student learning were excluded. In addition, the key reviews and meta-analyses were hand-searched for further eligible studies following the suggested snowballing methodology for systematic reviews (Page et al., 2021).

A narrative and thematic synthesis was used for this review, as the designs in the literature found were quite varied, and included randomized controlled trials, large-scale correlational studies of national and district-level data, qualitative and mixed-method programme evaluations. Findings were sorted chronologically as presented in the Literature Review and then synthesized thematically based on three patterns that emerged from the literature: PD's impact on teacher knowledge and instructional practice, the interconnected pathway between teacher learning and student achievement, and the programme, contextual, and implementation factors that moderate this pathway.

Findings

Professional Development, Teacher Knowledge and Instructional Practice

In all the literature reviewed, the practice of sustained and well-designed PD is repeatedly linked to increases in the content knowledge, pedagogical content knowledge and classroom instructional practice of teachers. Garet et al. (2001) found that content focus, active learning, and coherence were the PD design features most strongly predictive of teacher learning, and three years later, evidence showed that PD with content focus on specific instructional practices led to increased use of these practices by teachers, especially when coupled with opportunities for active learning (Desimone et al., 2002). Desimone (2009) synthesized this research and later reiterated it by Desimone and Garet (2015), although having the right design features does not ensure success in all contexts. Specifically in the field of science education, pedagogical content knowledge has proven to be an especially significant mechanism; a two-year PD intervention revealed that increase in teachers' academic content knowledge were the highest predictors of students' achievement (Gess-Newsome et al., 2019); and more recent work involving both text-based and video-based assessments of PCK has confirmed a strong relationship between science teachers' PCK and subsequent students' achievement (She et al., 2025). A cluster randomized trial of 77 schools, 144 teachers, and 2,823 students also revealed that students engaged in 1 year of science PD that combined analysis of practice with content deepening demonstrated a 0.52 SD higher gain in science content knowledge than those in content-deepening PD alone (Taylor et al., 2017).

This is the Pathway from Teacher Learning to Student Achievement.

One thread that has emerged from all of the reviewed studies is a three-step pathway that is closely linked—PD results in an increase in teachers' knowledge and professional competencies, which, in turn, substantially impacts classroom instructional practice, which, in turn may increases student achievement. This logic is reflected in the key concepts in the field (Desimone, 2009) and in large-scale meta-analytic studies: A synthesis of experimental mathematics and science PD interventions showed that mathematics and science PD interventions with larger gains in teacher-level outcomes also tended to have larger gains in student achievement, offering direct evidence for the teacher-to-classroom-to-student pathway (Lynch et al., 2025). There are also direct tests of the last step in this chain. Yoon et al. (2007) identified a moderate overall average effect size for reading, mathematics, and science PD interventions that satisfied rigorous causal-design standards. More recent evidence was generally consistent, including a meta-analysis of mathematics teacher PD that showed that there was an overall positive relationship between PD and student achievement, but noted there was a great deal of variation in effect size according to PD programme type (Franklin & Chang 2025), and a four-arm randomized controlled trial of a pedagogy-focused PD model that found that students of participating teachers made significantly higher progress compared with a wait-list control group (Gore et al. 2021).

Programme Design, Context and the Limits of Professional Development

Not all the evidence reviewed indicates benefits; studies reported limited or non-significant impacts are informative on the boundary conditions of the effectiveness of PD. The study, conducted by Garet et al., (2011) was a rigorous randomized evaluation of an intensive, two-year, content-focused PD programme for middle-school mathematics teachers that did not result in a statistically significant impact on teacher content knowledge or student achievement, but did result in increases in specific

observed instructional behaviors. A similar large cluster randomized trial of 89 schools, 253 teachers, and 20,591 students showed that a professional development intervention to boost teachers' content knowledge in science had no impact on science learning compared to business-as-usual, while a curriculum intervention that integrated cognitive-science principles with teacher professional development yielded a small, but statistically significant, positive effect on science learning (Kramer et al., 2024). In the same way a systematic review of 11 experimental and quasi-experimental studies of teacher PD revealed that while teacher PD was the most prevalent type of PD offered, the evidence regarding the connection between a specific type of PD and student learning gains was sparse and inconsistent beyond the areas of mathematics and literacy (Ventista & Brown, 2023). Kennedy (2016) provides a possible explanation for this inconsistency: in her review, many programmes in her sample that had popular structural characteristics, such as a specific length and delivery format, were not consistently related to programme effectiveness, implying that the underlying theory of teacher learning within a programme, and the pedagogical features that supported enactment of that theory, were more important than the structural features themselves.

Overall, these mixed results indicate that PD is not a guaranteed solution. The moderators of whether PD translates into better practice and ultimately greater student achievement are teachers' active and sustained engagement with PD; its duration and intensity (single workshops vs. sustained multi-year engagement as in Shymansky et al., 2012), alignment with teacher's subject matter and classroom context, and programme design (whether or not core features identified by Desimone, 2009, and Garet et al., 2001 are included). Context-specific PD, which is delivered in a way that is customized to a school's curriculum, student population, and constraints on instruction, seems especially critical for supporting the evolving requirements of science teaching, such as new curriculum standards, integrated STEM teaching, and inquiry teaching (Darling-Hammond et al., 2017; Huang et al., 2022).

Discussion

This systematic review suggests that teacher professional development (PD) is generally, but not uniformly, connected to positive learning outcomes for students in science, with the magnitude and consistency of the PD effect depending on the type of PD and the educational context. The overall picture of the relationship, across the empirical evidence published from 2002 to 2025, is less clear-cut, with not a single study being consistently positive. Professional development that is tightly linked to content, leads to active engagement by the teacher, and offers ongoing contextually-rich support to student learning is more likely to support student learning. Meanwhile, several studies have shown that participation in professional development is not a guarantee of increased student achievement, as several have reported non-significant effects. Overall, this pattern indicates that learning, professional learning, enactment of new knowledge in classroom practice, and conditions for enactment are all important factors in determining the potential of PD to contribute to science achievement.

The positive evidence is especially evident in research specific to science. The cluster randomized trial by Taylor et al. (2017), which involved 77 schools, 144 teachers, and 2823 students, showed that students in schools where teachers completed a one-year-long science content deepening professional development program had a 0.52 SD larger increase in science content knowledge than those in schools where the teachers only completed the science content deepening program. This is a particularly

informative result because both groups were offered PD: Exposure to more content for the students might have been the only thing that matters beyond the structure and integration of the PD in practice. Likewise, Shymansky et al. (2012) presented results from a multi-year K–6 professional development program that correlated sustained teacher professional development with student success at high-stakes science testing. These findings, when taken together, support the interpretation that professional learning can play a role in science achievement when it offers teachers the opportunity to develop the knowledge and expertise that they can use with students in the classroom.

The evidence also shows, however, that positive results are not always seen in all professional development programs. In a cluster-randomized controlled trial of a science professional development intervention to increase the science content knowledge of teachers, involving 89 schools, 253 teachers, and 20,591 students, Kramer et al. (2024) concluded that there was no statistically significant difference between the intervention group and the business-as-usual group in terms of student science learning. Conversely, a cognitive-science-based curriculum intervention with teacher professional development had a small but statistically significant positive impact on content learning for students. This is in contrast to the assumption that simply increasing teachers' content knowledge will yield meaningful changes in students' achievement. Teacher knowledge could be a key intermediate outcome but might also rely on teacher opportunities, instructional supports, resources, and time to incorporate the knowledge into the classroom. In short, the mixed results suggest that the effectiveness of PD is not consistent and cannot be concluded to be either effective or ineffective.

The findings also support a pathway in which professional development is linked to student achievement via changes in the professional knowledge and instructional practice of teachers. This reading aligns with Desimone's (2009) conceptualization of teacher learning and change in instructional practice as two outcomes that facilitate student learning. The pathway is complex, as Gess-Newsome et al. (2019) demonstrate empirically. They examined, in their two-year study, teachers' academic content knowledge (ACK), general pedagogical knowledge (GenPK), pedagogical content knowledge (PCK), classroom practices, and student achievement. Results revealed that various aspects of teacher knowledge were associated with instructional practice and student outcomes in different ways. Specifically, GenPK was associated with teacher practice, and ACK accounted for significant differences in student achievement. The results indicate that PD can have a multi-faceted impact on the achievement of students, not just by directly improving students' achievement, but by also having indirect impacts through multiple and interconnected paths of teacher learning.

In science education, the importance of subject-specific professional knowledge cannot be overlooked. Good science instruction should involve teachers in learning about the concepts they teach, as well as how those concepts can be represented, explained, assessed, and related to students' background knowledge. Thus, PCK is applicable because it is about how disciplinary knowledge is changed into forms that can assist student learning. She et al. (2025) found positive correlations between science teachers' PCK and students' achievement in both text-based and video-based assessments of PCK. This evidence reinforces the argument that teacher subject matter knowledge is important to science learning. The finding should not be taken to mean that professional development alone leads to increased PCK or achievement, however. Instead, it identifies PCK as a plausible mechanism that

connects teacher knowledge, teaching, and learning outcomes. This interpretation is also supported by the findings of Taylor et al. (2017), which show that content deepening produced greater gains for the students when it was followed by an analysis of classroom practices.

It's also clear that design affects the potential effectiveness of professional development. Certain features of effective professional learning have been identified: content focus, active learning, coherence, adequate duration, and collective participation (Garet et al., 2001; Desimone, 2009). These features can offer teachers the chance to explore the content of the topics, relate new practice to existing practice, work collaboratively with others, and incorporate new practice in classroom contexts. However, these features are not a guarantee of student achievement. Kennedy (2016) showed that common factors highlighted as important (such as programme length and delivery format) were not always linked to programme effectiveness, and that the underlying approach to teacher learning was important. This interpretation aligns with the science-specific results in Taylor et al. (2017) that showed that content deepening paired with analysis of practice led to higher student gains than content deepening alone. Therefore, the pedagogical aspects and goal of professional learning may be more important than structural features when taken alone.

Duration should also be considered a facilitating condition and not a determining factor of effectiveness. Ongoing professional learning can offer teachers the chance to build subject expertise, try out new methods, get feedback, be reflective, and get better at teacher practice. These processes may be particularly important when professional development aims to change complex science teaching practices. But more time does not equal more success for students. During an intensive two-year mathematics professional learning evaluation, Garet et al. (2011) did not find any statistically significant differences in student performance or teacher content knowledge, even though there was significant mathematics professional learning provided. Even though this study is focused on mathematics and not on science, the evidence provided here is relevant to the comparative context and suggests that duration and intensity should not be considered separate predictors of success in mathematics, but rather should be taken in conjunction with the quality of activities to engage students in learning, teacher engagement, the focus of instruction, and conditions for implementation.

Variation in outcomes is also explained by implementation quality and contextual fit. Professional learning should involve more than just absorbing information; teachers should be given the chance to make sense of new content in relation to their current knowledge, to experiment with new practices, to get feedback on their practice, and to make adjustments for their students and classrooms. Engagement of teachers should be seen as an active element of professional learning and not just as attendance. Meanwhile, teachers engage in professional learning in defined curriculum, institutional, and classroom settings. Professional learning can lead to sustained instructional change, which can be enhanced or hindered by curriculum requirements, available resources, support within schools, student characteristics, and classroom conditions. Huang et al. (2022) underscored the significance of the context in interpreting the meanings of professional learning outcomes, and Darling-Hammond et al. (2017) emphasized the need for contextually relevant professional learning. This evidence implies that PD should not be a uniform or standardized program, especially in science education where teachers may be faced with significantly different curricular and resource environments.

Overall, the findings indicate that programme design, teacher learning, instructional practice, implementation, and context all interact and contribute to the variability in the effectiveness of teacher professional development. Professional development that provides relevant teacher knowledge, ties the knowledge to classroom practice, allows for opportunities for active engagement and application, and includes conditions that promote sustained instructional change is more likely to yield positive student outcomes. On the other hand, professional development can enhance teacher knowledge or specific teacher practices, but not lead to measurable student achievement gains if teaching is not applied in the classroom or when application of professional learning and contextual factors limit the extent to which the professional learning is implemented. The evidence suggests that while the question of whether professional development is effective is not the most important, it is under what conditions and how it is effective to enhance science learning.

The results have implications for professional development policy and practice. In addition to immediate information delivery, professional learning should continue, focus on content, be experiential, and take into account the context of teachers' classrooms. Professional development in science education should enable teachers to build their subject matter knowledge, gain knowledge of PCK, and reflect on how the subject matter knowledge can be implemented in classroom practices. Programmes should also include active learning, analysis of classroom practice, collaboration, feedback and opportunities for classroom-based application. These methods are more in line with the mechanisms identified in the reviewed evidence, rather than methods that are driven by attendance or professional development hours.

The priorities for future studies are also noted in the review. More science-specific longitudinal and randomized research is required to explore changes in teacher knowledge, PCK, and instructional practice over time and the relationship of these changes to subsequent student outcomes. Further studies are also needed to explore factors that could moderate the effects, including programme design, length of the programme, implementation fidelity, teacher factors, school support, curriculum factors, and educational context. There is a need for more under-researched educational settings to identify the degree to which research findings from specific educational systems can be generalised to other settings. Future systematic reviews and meta-analyses should be able to build on the growing experimental evidence and evaluate the effects of the programmes as well as variations as a function of programme and contextual features. This kind of research could contribute to the progress of the field beyond simply measuring the effectiveness of teacher professional development to determining what works, for whom, under what circumstances, and why.

Conclusion

A systematic review was conducted to integrate over 20 years of research (2002–2025) about the connection between teachers' professional development and students' academic achievement in science classrooms. The evidence suggests that continuous, well-designed PD supports teachers' professional learning, practices, and commitment and works along a pathway of interconnectedness: better teachers' professional learning results in better teachers' practices which in turn produce better student learning. However, the effectiveness of PD is not assured; PD programme design, programme implementation, programme time, contextual fit and the active involvement of teachers all have an

impact on the success of a specific PD initiative. This conclusion has remained remarkably stable over the last two decades, as the methods used by the field have evolved, from the earliest assessment work of the reform era to the blending of conceptual frameworks and to the current era of large-scale randomized trials and systematic reviews and meta-analysis. Policymakers and school leaders aiming to boost science achievement should therefore prioritize the development of longer term and more evidence-based and context-relevant professional learning programs over shorter term, generic professional development sessions, to enhance teachers' professional competence, improve teaching and learning and ultimately support improved achievement outcomes for science students.

Recommendations

This review suggests the following recommendations for policy makers, school leaders, teacher educators and researchers.

- Policymakers and education ministries should invest more in long-term (multi-year) professional development programmes than in short-term (one-off) professional development sessions because research has consistently shown that more sustained professional development has been shown to have more profound impacts on teacher practices and student achievement (Desimone, 2009; Shymansky et al., 2012).
- Instructional coaches and school leaders need to plan or select professional development that incorporates the five core essential elements that have been identified in the literature: content focus, active learning, coherence with school goals, adequate duration, and collective participation (Garet et al., 2001; Desimone & Garet, 2015).
- In the context of science classrooms, teacher education institutions and PD providers should give particular attention to developing science teachers' pedagogical content knowledge (PCK) specifically, as there is a strong and repeated relationship between PCK and student achievement in science classrooms (Gess-Newsome et al., 2019; She et al., 2025).
- Professional development must be tailored to the curricular context, resources, and classroom context of the individual teacher, and should not be a "one size fits all" intervention, especially in classrooms with low resources or high students, teacher's ratio which do not often appear in the evidence base (Darling-Hammond et al., 2017; Huang et al., 2022).
- Because research indicates that programmes with the right design features may fail to yield achievement gains when implemented poorly (Garet et al., 2011, 2010; Kennedy, 2016), programme implementers should ensure that there are procedures in place to monitor implementation fidelity during a programme and teachers' active participation throughout the PD process.
- Future research should be more science-specific, longitudinal, and context-differentiated in order to identify specific PD design features that lead to greater scientific achievement and to expand the body of evidence to under-researched settings, such as low-resource and non-Western settings, as well as to other contexts (Ventista & Brown, 2023).
- Future systematic reviews and meta-analyses should keep reviewing and summarising the accumulating evidence base in the field as new RCTs and large scale studies are released. In

addition to overall effects, moderators (e.g., programme design, duration, implementation quality, teacher characteristics and educational context) should be reviewed in such a review (Lynch et al., 2025).

References

- Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and teacher education*, 27(1), 10-20.
- Banilower, E. R. (2002). Results of the 2001-2002 study of the impact of the local systemic change initiative on student achievement in science.
- Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the national science foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 44(3), 375-395.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, 33(8), 3-15.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. *Learning policy institute*.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational researcher*, 38(3), 181-199.
- Desimone, L. M., & Garet, M. S. (2015). Best practices in teacher's professional development in the United States.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational evaluation and policy analysis*, 24(2), 81-112.
- Franklin, A. V., & Chang, M. (2025). Meta-analysis for math teachers' professional development and students' achievement. *Education Sciences*, 15(9), 1156.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American educational research journal*, 38(4), 915-945.
- Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., ... & Doolittle, F. (2011). Middle School Mathematics Professional Development Impact Study: Findings after the Second Year of Implementation. NCEE 2011-4024. *National Center for Education Evaluation and Regional Assistance*.
- 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.
- Gore, J. M., Miller, A., Fray, L., Harris, J., & Prieto, E. (2021). Improving student achievement through professional development: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education*, 101, 103297.

- Guskey, T. R. (2014). En kortfattad beskrivning om hur en skola i Haninge kommun har arbetat med att implementera formativ bedömning hållbart i verksamheten.
- Huang, B., Jong, M. S. Y., Tu, Y. F., Hwang, G. J., Chai, C. S., & Jiang, M. Y. C. (2022). Trends and exemplary practices of STEM teacher professional development programs in K-12 contexts: A systematic review of empirical studies. *Computers & education*, 189, 104577.
- Kramer, S. L., Scull, J., Porter, A., Massey, C. M., Merlino, F. J., & Baker, J. Y. (2024). Can either using cognitive science principles or improving teacher content knowledge boost student achievement in middle school science?. *Journal of Research in Science Teaching*, 61(7), 1543-1573.
- Kennedy, M. M. (2016). How does professional development improve teaching?. *Review of educational research*, 86(4), 945-980.
- Lynch, K., Gonzalez, K., Hill, H., & Merritt, R. (2025). A meta-analysis of the experimental evidence linking mathematics and science professional development interventions to teacher knowledge, classroom instruction, and student achievement. *AERA open*, 11, 23328584251335302.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372.
- She, J., Chan, K. K. H., Wang, J., Hu, X., & Liu, E. (2025). Effect of science teachers' pedagogical content knowledge on student achievement: Evidence from both text-and video-based pedagogical content knowledge tests. *American Educational Research Journal*, 62(1), 92-135.
- Shymansky, J. A., Wang, T. L., Annetta, L. A., Yore, L. D., & Everett, S. A. (2012). How much professional development is needed to effect positive gains in K–6 student achievement on high stakes science tests?. *International journal of science and mathematics education*, 10(1), 1-19.
- Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of research in science teaching: the official journal of the national association for research in science teaching*, 37(9), 963-980.
- Taylor, J. A., Roth, K., Wilson, C. D., Stuhlsatz, M. A., & Tipton, E. (2017). The effect of an analysis-of-practice, videocase-based, teacher professional development program on elementary students' science achievement. *Journal of Research on Educational Effectiveness*, 10(2), 241-271.
- Ventista, O. M., & Brown, C. (2023). Teachers' professional learning and its impact on students' learning outcomes: Findings from a systematic review. *Social Sciences & Humanities Open*, 8(1), 100565.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). Reviewing the evidence on how teacher professional development affects student achievement. issues & answers. rel 2007-no. 033. *Regional Educational Laboratory Southwest (NJ1)*.