

Technology-Mediated Teacher Professional Development and Sustainability of Outcomes: A Systematic Literature Review

Farid Fachrudin

Educational Technology Program, Universitas Negeri Jakarta, Indonesia

ABSTRACT

Technology-mediated teacher professional development (TPD) has expanded rapidly across education systems worldwide, a shift the COVID-19 pandemic accelerated beyond any planned timeline. Yet a foundational question remains unresolved: do these programs produce lasting changes in how teachers teach? This systematic literature review examines the conditions, mechanisms, and evidence gaps associated with sustainability of outcomes in technology-mediated TPD across diverse global contexts (2015–2026). Following the PRISMA 2020 protocol—comprising identification, screening, eligibility assessment, and inclusion—158 initially retrieved records were systematically reduced to 42 peer-reviewed studies involving in-service teachers with measurable outcomes and a minimum three-month post-program follow-up, which were then synthesized using a realist, CMO-configurational approach. We developed the T-STEP (Technology-mediated Sustainability in Teacher Education and Practice) Framework, the first CMO-grounded synthesis of technology-mediated TPD sustainability across global contexts. Four configurational pathways emerged: high-impact programs integrating coaching and leadership support sustained classroom change beyond 12 months ($n=14$); partial-transfer programs using online-only delivery without coaching showed gains of three to six months before decay ($n=12$); passive delivery programs produced no significant practice transfer ($n=9$); and equity-gap programs achieved conditional sustainability through culturally responsive design ($n=7$). Across all pathways, school culture and leadership support proved stronger predictors of sustainability than program design or technology type. Findings offer actionable implications for TPD designers, institutional leaders, and policymakers pursuing evidence-grounded approaches to sustained teacher learning.

Keywords: *teacher professional development, technology-mediated learning, sustainability, blended learning, classroom practice, systematic literature review*

1. INTRODUCTION

Teacher professional development (TPD) occupies a central position in every serious account of educational quality improvement (Desimone & Garet, 2015; Popova et al., 2022; Tonga et al., 2022). Over the past decade, digital platforms, blended learning environments, and AI-enhanced tools have fundamentally altered what TPD looks like in practice (Bragg et al., 2021; Kahraman & Bicen, 2022). The COVID-19 pandemic sharpened this transformation into something far more abrupt: within weeks, teacher training migrated from face-to-face settings to online and hybrid formats at a scale no one had planned for (Cespedes, 2025; Gurgenidze et al., 2022). The transition was rapid, compelled by circumstance rather than by design and it created, almost by accident, a global natural experiment in technology-mediated

teacher learning (Lay et al., 2020; Portillo & Lopez de la Serna, 2021). Yet the transformation in what teachers must know and be able to do did not begin with, and has not ended alongside, the pandemic: the continuing diffusion of artificial intelligence, adaptive platforms, and data-driven instructional tools keeps generating new competency demands that outlast any single crisis-driven adoption cycle (Kahraman & Bicen, 2022; UNESCO, 2023).

That experiment has not, on balance, yielded encouraging results. The central question—does technology-mediated TPD produce lasting changes in how teachers teach?—remains unanswered in ways that should trouble anyone who funds or designs these programs. Theoretically, this gap is unsurprising: Desimone's (2009) core conceptual framework for studying professional development effects distinguishes teachers' knowledge and belief change from the far more demanding step of enacted classroom practice change, and Bandura's (1997) social cognitive theory further specifies that practice change requires not only new knowledge but a shift in teachers' self-efficacy beliefs about their capacity to enact it. The gap between short-term knowledge acquisition and sustained classroom practice change is therefore not a technicality; it reflects a fundamentally different theory of what professional development is for (Guskey, 2002; McChesney & Aldridge, 2021). Evidence from large-scale evaluations paints a sobering picture. A randomized controlled trial of a national teacher training program in China found no measurable effect on student achievement despite substantial investment with passive delivery identified as the primary culprit (Loyalka et al., 2019). The same pattern has surfaced in Sub-Saharan Africa (Muhuro & Kangethe, 2021; Nkanyani et al., 2024) and South Asia (Nedungadi et al., 2024), where short-term gains typically erode within six months of program completion.

A smaller but instructive body of work points toward conditions under which technology-mediated TPD does sustain. Studies from Argentina, New Zealand, Belgium, and high-performing PISA systems identify a consistent cluster: embedded coaching, active learning design, coherence with school context, and distributed leadership support (Albornoz et al., 2020; Fischer et al., 2018; Tonga et al., 2022; Van Peteghem et al., 2025). The field does not lack descriptions of these enabling conditions. What it lacks is a mechanistic account of how they interact with technology to produce or undermine sustainability particularly in low-resource and culturally diverse settings (Burgess et al., 2022; Nguyen et al., 2026).

Existing systematic reviews of TPD (Firestone et al., 2020; Koellner et al., 2022; Lay et al., 2020; Philipsen et al., 2019) have moved understanding forward considerably. Three gaps nonetheless persist. First, most reviews treat online delivery and technology-mediated design as equivalent, obscuring how different forms of technological mediation enable or constrain sustainability mechanisms in distinct ways (Bragg et al., 2021). Second, program features are typically analyzed in isolation, without attending to the contextual conditions that determine whether those features are actually activated (McChesney & Cross, 2023). Third, evidence from low-resource, rural, and indigenous settings remains systematically underrepresented a problem that is not merely descriptive but interpretive, since dominant findings may not

generalize to the contexts where most of the world's teachers work (Burgess et al., 2022; Muhuro & Kangethe, 2021; Ngoveni et al., 2025).

This review addresses these gaps through a realist synthesis lens specifically the Context-Mechanism-Outcome (CMO) configurational logic of Pawson and Tilley (1997), adapted for educational systematic review by Wong et al. (2013) applied to 42 studies spanning 2015–2026. The core argument is this: technology in TPD is not a mechanism. It is a medium through which pedagogically sound or unsound mechanisms are delivered. This review is accordingly organized around one core question: under what configuration of conditions does technology-mediated TPD produce sustained, rather than transient, change in classroom practice? Three sub-questions operationalize this core question:

RQ1: What program design conditions are consistently associated with sustained classroom practice change in technology-mediated TPD?

RQ2: To what extent does the literature document long-term outcomes from technology-mediated TPD, and what contextual factors mediate sustainability?

RQ3: What methodological gaps characterize the evidence base, particularly for low-resource contexts?

2. METHODS

This review employs realist synthesis (Pawson, 2006; Pawson & Tilley, 1997; Wong et al., 2013), combined with thematic synthesis (Thomas & Harden, 2008) and narrative synthesis techniques for presenting configurational patterns (Popay et al., 2006), following established procedures for conducting systematic reviews in the social sciences (Booth et al., 2016). The T-STEP (Technology-mediated Sustainability in Teacher Education and Practice) Framework organizes evidence across three analytical layers corresponding to CMO logic: Layer 1 (Context) captures enabling and constraining conditions—school culture, leadership support, digital readiness, and resource environment; Layer 2 (Mechanism) captures the processes through which technology mediates professional learning—program design, technology mediation type, and transfer conditions; and Layer 3 (Outcome) captures observable effects at teacher and student level across documented follow-up periods, including sustained classroom practice change, teacher self-efficacy shift, and student achievement gains.

This review followed PRISMA 2020 reporting guidelines (Page et al., 2021). Four databases were searched on 8 May 2026: Scopus (primary), Web of Science, ERIC, and Google Scholar (supplementary), covering January 2015 to May 2026. The Scopus search string combined terms for teacher professional development, technology delivery modes (online, blended, AI-enhanced, digital platform), and sustainability outcomes (sustained, long-term, classroom practice, student achievement). Screening followed four sequential PRISMA stages, summarized in Table 1: identification (158 records retrieved from the four databases), screening (title/abstract review against the eligibility criteria in Table 2, reducing the corpus to 89 records), eligibility assessment (full-text review against inclusion/exclusion criteria), and

inclusion (confirmation of the final 42-study corpus). Google Scholar and citation-tracking of key prior reviews supplemented the primary Scopus search to guard against omission of relevant studies not indexed under the core search string (Greenhalgh & Peacock, 2005).

Table 1. PRISMA 2020 flow diagram of study identification, screening, eligibility, and inclusion.

Stage	Records (n)
Database search (Scopus, WoS, ERIC, Google Scholar)	158
After deduplication	140
After title/abstract screening	89
Excluded (no measurable outcome)	18
Excluded (non-technology-mediated)	12
Excluded (follow-up < 3 months)	9
Excluded (not peer-reviewed)	5
Excluded (inaccessible full text)	3
Included for final synthesis	42

Cross-case analysis yielded four dominant CMO configurations. Pathway 1 (High impact): supportive context plus coached blended mechanisms plus adequate transfer conditions produced sustained practice change beyond 12 months (n=14). Pathway 2 (Partial transfer): moderate context plus online-only delivery without coaching plus limited follow-up produced short-term gains of three to six months followed by rapid decay (n=12). Pathway 3 (Technology noise): weak context plus passive delivery produced no significant classroom practice transfer (n=9). Pathway 4 (Equity gap): low-resource or indigenous context plus culturally responsive mechanisms produced conditional sustainability (n=7).

Studies were included if they: (1) involved technology-mediated TPD for in-service teachers at any education level; (2) reported at least one measurable outcome (classroom practice, student achievement, or teacher self-efficacy); (3) documented a follow-up period of minimum three months post-program; (4) were peer-reviewed and published in English between 2015 and 2026. Table 2 presents the criteria applied at each stage of screening. The inclusion threshold for follow-up duration, a minimum of three months post-program was set deliberately to exclude studies capturing only immediate post-test effects, which dominated earlier literature but tell us little about sustainability.

Table 2. Inclusion and exclusion criteria applied to the screening process.

Criterion	Inclusion	Exclusion
Study type	RCT, quasi-experimental, mixed-methods, SLR/meta-analysis, qualitative with measurable outcomes	Opinion pieces, editorials, conference abstracts without full paper

Population	In-service teachers, all levels (K-12, HE, TVET)	Pre-service teacher education; student-only outcomes without teacher component
Intervention	Online, blended, AI-enhanced, platform-based TPD	Purely face-to-face TPD without technology component
Outcome	At least one measurable outcome: classroom practice, student achievement, or teacher self-efficacy	Studies measuring only satisfaction or reaction (Kirkpatrick Level 1 only)
Follow-up	Minimum 3 months post-program	Follow-up < 3 months or immediate post-test only
Language	English full-text	Non-English without official translation
Year	2015–2026	Before 2015 (except landmark methodological references)

All 42 included studies were appraised using instruments matched to design: CASP RCT checklist for experimental designs, MMAT Version 2018 (Hong et al., 2018) for mixed-methods, CASP Qualitative for qualitative studies, and AMSTAR-2 (Shea et al., 2017) for systematic reviews. Studies meeting less than 60% of applicable criteria were excluded. Inter-rater reliability at the full-text screening stage was Cohen's $\kappa = 0.84$ (Cohen, 1960; McHugh, 2012), indicating almost-perfect agreement between screeners. Final quality distribution: high ($n=14$), moderate ($n=21$), lower ($n=7$).

3. RESULTS AND DISCUSSION

Forty-two studies met the inclusion criteria, spanning eleven years of publication from 2015 to 2026. The majority (79%) appeared between 2019 and 2026, a distribution that mirrors the acceleration of research interest following the COVID-19 pandemic, when technology-mediated TPD moved from an option to an operational necessity for millions of teachers. Only nine studies (21%) documented follow-up periods exceeding 12 months. That figure matters: it confirms, rather than merely suggests, that long-term longitudinal evidence is absent from a field where sustainability claims are routinely made. SLR and meta-analyses constitute the largest single design category (26%). Table 3 summarizes the distribution of study characteristics.

Table 3. Characteristics of the 42 included studies.

Characteristic	Category	n	%
Intervention type	Online / blended	18	43%
	AI-enhanced platforms	14	33%
	Rural / low-resource blended	10	24%
Study design	SLR / meta-analysis	11	26%

Characteristic	Category	n	%
	Quasi-experimental	10	24%
	RCT	4	10%
	Mixed methods	9	21%
	Qualitative	8	19%
Follow-up duration	3–6 months	18	43%
	7–12 months	15	36%
	> 12 months	9	21%
Geographic context	Global / multi-country	12	29%
	Asia-Pacific	9	21%
	Europe	8	19%
Quality appraisal	Sub-Saharan Africa	7	17%
	Latin America / Other	6	14%
	High	14	33%
	Moderate	21	50%
	Lower	7	17%

RQ1: Conditions Associated with Sustained Classroom Practice Change

The most consistent finding across all 42 studies is also the most practically actionable one: coaching works. Structured, ongoing, individualized support from a knowledgeable other is the single program design feature most strongly associated with sustained classroom practice change (Joyce & Showers, 2002). This holds across intervention types, geographic contexts, and quality levels (Bognar et al., 2024; Desimone & Garet, 2015; Popova et al., 2022). The effect sizes involved are not marginal. Albornoz and colleagues (2020) found that adding a coaching component to an otherwise identical online science TPD program doubled the effect size on teacher practice at 18 months (Cohen's $d = 0.41$ with coaching versus $d = 0.19$ without). Popova et al. (2022) confirmed that programs with embedded coaching components yielded three to four times the sustained impact of programs without it.

The mechanism matters here too, not just the finding. Van Peteghem et al. (2025) describe a process they term "belief-mediated practice change": sustained classroom behavior change was predicted by shifts in teacher beliefs about students' learning capacities, shifts produced through structured dialogue and reflective coaching cycles, not through online content consumption. The technology facilitates access; the coaching does the pedagogical work. This

is consistent with self-efficacy theory (Bandura, 1997): coaching cycles function less as knowledge-transmission mechanisms and more as self-efficacy-building mechanisms, strengthening teachers' belief in their own capacity to enact new practice, a distinction Woodcock and Hardy (2025) draw explicitly in linking sustained engagement with PD to teachers' efficacy beliefs rather than to program duration alone.

Beyond coaching, active learning design and content focus emerged as the second and third most consistently cited design conditions (Desimone & Garet, 2015; Ingvarson et al., 2005). Bognar et al. (2024) confirm that blended delivery is equally effective to face-to-face but only when these elements are present; without them, the delivery mode is irrelevant. Coherence between TPD content and the specific curriculum teachers are responsible for matters more than many program designers recognize. Fischer et al. (2018) found that this content-context coherence predicted 18-month classroom adoption as strongly as program design quality itself, a pattern corroborated across 14 European CPD systems by Antoniou (2016). Teachers whose training matched their actual context-maintained changes at 18-month follow-up; those in misaligned programs showed decay within six months, regardless of how well-designed the program appeared on paper.

Duration and continuity of engagement are necessary conditions, not sufficient ones. Philipsen et al. (2019) established that single-episode or short-burst online programs produce negligible sustained outcomes. What matters structurally is that programs include follow-up touchpoints beyond the initial training event, structured opportunities for teachers to return to the material in context (McChesney & Aldridge, 2021).

Table 4 summarizes these program design conditions and their relationship to sustainability of outcomes. One finding cuts across all four CMO pathways with unusual consistency: school culture functions as the contextual meta-condition within which all program design features either activate or stall (McChesney & Cross, 2023; Schein, 2010). McChesney and Cross (2023) found that positive, collaborative school cultures sustained TPD implementation at 18 months, while competitive or non-collaborative cultures produced rapid regression regardless of program quality. This is not a peripheral moderating variable, it is, in this synthesis, the single most powerful contextual determinant of whether TPD investment translates into lasting classroom change. Lyu (2025) corroborates this from a Chinese EFL context: school culture significantly moderated whether teacher motivation from online PD translated into sustained classroom application, with culture-fit programs showing approximately 40% higher retention of new practices at five months.

Table 4. Program design conditions and their relationship to sustainability of outcomes.

Design Condition	Evidence Strength	Effect on Sustainability	Key Studies
Coaching embedded	Strong (RCT + meta)	Doubles sustained effect size	Albornoz et al., 2020; Popova et al., 2022; You et al., 2025

Design Condition	Evidence Strength	Effect on Sustainability	Key Studies
Active learning design	Strong (multi-SLR)	Required for practice transfer	Desimone & Garet, 2015; Bognar et al., 2024; Ingvarson et al., 2005
Content-context coherence	Moderate (quasi-exp)	Predicts 18-month adoption	Fischer et al., 2018; Tonga et al., 2022; Crescentini & Ostinelli, 2023
Continuity / follow-up	Strong (meta)	Absence causes rapid decay	Philipsen et al., 2019; McChesney & Aldridge, 2021
Leadership endorsement	Moderate-strong	Mediates implementation rate	Youngs & King, 2002; Swayne & Moynihan, 2025
School culture	Strong (longitudinal)	Meta-condition for all features	McChesney & Cross, 2023; Lyu, 2025; Zhou, X., et al., 2024
Technology type	Weak (inconsistent)	Delivery mode secondary to design	Bragg et al., 2021; Wuryaningsih et al., 2019

RQ2: Long-Term Outcomes and Mediating Factors

One of the most arresting patterns in this review is the mismatch between the timeframes at which outcomes are measured and the timeframes at which educational impact operates. Lay et al. (2020) found that fewer than 30% of online TPD studies documented sustained classroom practice change, the majority measured only satisfaction or knowledge at immediate post-test. Among the 18 studies with follow-up periods of three to six months, the modal pattern is unambiguous: initial gains followed by partial decay. Cespedes (2025) provides a useful comparison point, showing that teachers receiving structured digital mentoring-maintained classroom innovations at six months, while those in unstructured remote PD programs showed rapid decay within the same window.

Nine studies documented follow-up periods exceeding 12 months, a small corpus on which to base claims about long-term sustainability, but instructive, nonetheless. Downey and Luebeck (2025) found that place-attentive professional learning sustained teacher professional growth over 14 months in Canadian rural contexts, and Kurteshi and Rustemi (2025) similarly documented sustained gains in teacher performance from continuous online professional development in a longitudinal study from Kosovo. The most methodologically authoritative longitudinal evidence comes from two RCTs. Albornoz et al. (2020) demonstrated sustained teaching quality gains at 18 months in the coaching condition. Loyalka et al. (2019) document the strongest null result in the literature: a large-scale national program in China with no measurable effect on student achievement despite substantial investment, isolating passive delivery without coaching as the critical differentiating variable, a conclusion Lu et al. (2019) reinforce in a parallel evaluation of TPD programs and student achievement in rural China. Five contextual factors consistently mediate sustainability across the corpus: leadership support, peer learning networks, teacher self-efficacy, contextual coherence between program

content and school curriculum, and infrastructure reliability, particularly in low-resource contexts where access to technology functions as a binary gate rather than a gradient.

RQ3: Evidence Gaps and Equity-Attentive Research

Seven studies in the corpus focused explicitly on low-resource, rural, or indigenous teacher populations, the Equity Gap pathway. These studies share a common diagnostic: conventional technology-mediated TPD designs developed in well-resourced OECD contexts do not fail in low-resource settings because their pedagogical principles are wrong. They fail because their design assumptions are misaligned with contextual conditions (Burgess et al., 2022; Muhuro & Kangethe, 2021). Nguyen et al. (2026) demonstrated that a three-component blended model built around readiness assessment, local relevance, and frugal resource design achieved five-month retention of new instructional practices in low-bandwidth Vietnamese settings. Burgess et al. (2022) documented in an Australian indigenous context that decolonizing TPD through cultural mentorship sustained practice change over 18 months.

Three systematic methodological gaps limit what conclusions can be drawn from the existing evidence base. First, classroom observation data—the most direct and least gameable measure of practice change, and the gold-standard instrumentation exemplified by tools such as the Classroom Assessment Scoring System (Pianta et al., 2008)—appears in fewer than one-third of included studies; most rely on teacher self-report susceptible to social desirability and recall bias. Second, implementation fidelity reporting is consistently inadequate (O'Donnell, 2008), leaving the causal mechanism linking program design to outcomes opaque in ways that make replication and improvement impossible, a pattern Ramli et al. (2024) confirm is a recurring implementation barrier for technology-mediated PD programs more broadly. Third, the near-absence of studies with follow-up periods beyond two years means the field lacks any empirical basis for claims about durability of technology-mediated TPD outcomes at the career level. These are not minor gaps. They are structural deficits in an evidence base that is already being used to justify large-scale policy investments. Read against the T-STEP Framework, all three gaps cluster at Layer 3 (Outcome): without observational, fidelity-linked, and long-horizon data, the Context and Mechanism conditions identified under RQ1 cannot be causally verified as the drivers of sustained practice change, they remain theoretically plausible but empirically under-tested.

Discussion: Reframing Technology as Mediator, Not Mechanism

The first analytic contribution of this review is a reframing. The conventional research question does online or AI-enhanced TPD work? is, this synthesis suggests, the wrong question. Technology delivery mode is not consistently associated with sustainability of outcomes once design and contextual factors are controlled (Bognar et al., 2024; Bragg et al., 2021). What does sustain is the activation, through whatever channel, of six design conditions: content focus, active learning, coherence, duration, collective participation (DuFour & Eaker, 1998), and coaching (Desimone & Garet, 2015). Technology is the medium, not the message. Framed through the CMO logic underpinning this synthesis (Pawson & Tilley, 1997),

technology operates at the level of Mechanism, not Context or Outcome: it is the delivery channel through which pedagogically activating conditions are, or are not, triggered, which is precisely why technology type shows a weak and inconsistent direct relationship with sustainability in Table 4. This has a direct implication for what counts as genuine innovation in TPD design: the current generation of AI-enhanced and adaptive platforms (Kahraman & Bicen, 2022; Zhou, Q., et al., 2024) advances teacher competency development only to the extent that it is engineered to deliver these six mechanisms—for example, through automated coaching prompts or adaptive content sequencing—rather than merely digitizing passive content delivery under a new label.

The second contribution is what we term the contextual primacy thesis: contextual conditions, particularly school culture, leadership, and infrastructure, function as binding constraints on whether well-designed program mechanisms produce sustained outcomes (Hargreaves & Fullan, 2012; McChesney & Cross, 2023), a pattern Zhang et al. (2024) independently confirm in a cross-national comparison of the factors that make TPD effective. The clearest evidence for this comes from a direct comparison: the Loyalka et al. (2019) China trial and the Albornoz et al. (2020) Argentina trial deployed broadly comparable program architectures and produced opposite results. The differentiating variable was not technology, content, or coaching in isolation, it was the contextual configuration in which those features operated.

The third contribution lies in what Pathway 4 implies for the field's research agenda. The seven equity-gap studies represent contexts where the majority of the world's teachers practice yet they constitute fewer than one-fifth of this corpus. The dominant evidence base, built disproportionately on OECD contexts, likely overstates the generalizability of TPD effectiveness claims. The field needs to reckon with that gap, not treat it as a footnote.

4. CONCLUSION

This review set out to answer three questions about technology-mediated teacher professional development. The answers, taken together, tell a coherent story, though not an optimistic one for business-as-usual approaches to TPD design. This systematic literature review synthesized 42 studies of technology-mediated TPD sustainability (2015–2026) through the T-STEP Framework and CMO configurational analysis. Three findings are empirically robust across the corpus.

First, coaching is the single most powerful design condition for sustained classroom practice change doubling effect sizes and appearing as the critical differentiating variable between effective and ineffective programs across contexts. Second, school culture and leadership support are stronger predictors of sustainability than technology type or program design establishing contextual primacy as the binding principle of technology-mediated TPD effectiveness. Third, the evidence base systematically underrepresents low-resource and

indigenous contexts, limiting the generalizability of dominant effectiveness claims to the minority of contexts where they were studied.

For TPD designers, the practical implication is clear: invest in coaching cycles, content-curriculum coherence, and follow-up touchpoints; treat technology as infrastructure, not curriculum. For institutional leaders, school culture and distributed leadership yield larger dividends than platform investment. For policymakers in low- and middle-income countries, technology-mediated TPD requires parallel investment in cultural alignment and frugal design without which sustainability remains aspirational. These findings directly support SDG 4c's emphasis on qualified, well-supported teachers as the foundation of quality education systems.

Future research should prioritize longitudinal designs with 24–60-month follow-up using classroom observation, equity-attentive research in Sub-Saharan, South Asian, and Pacific contexts, and rigorous evaluation of emerging AI-enhanced TPD programs. Implementation fidelity reporting should be adopted as a precondition for publication of TPD outcome studies.

5. REFERENCES

- Albornoz, F., Anauati, M. V., Furman, M., Luzuriaga, M., Podesta, M. E., & Taylor, I. (2020). Training to teach science: Experimental evidence from Argentina. *World Bank Economic Review*, 34(2), 393–417. <https://doi.org/10.1093/wber/lhy026>
- Antoniou, P. (2016). Continuing professional development for teachers across Europe. *European Journal of Teacher Education*, 39(2), 148–161.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bognar, B., Matic, L., & Sablic, M. (2024). Effectiveness of blended versus face-to-face professional development for mathematics teachers. *Mathematics Teaching-Research Journal*, 16(1), 45–63.
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). SAGE Publications.
- Bragg, L. A., Walsh, C., & Heyeres, M. (2021). Successful design and delivery of online professional development for teachers: A systematic review of the literature. *Computers & Education*, 166, 104158. <https://doi.org/10.1016/j.compedu.2021.104158>
- Burgess, C., Bishop, M., & Lowe, K. (2022). Decolonising teacher professional development: Lessons from an Australian indigenous education context. *Discourse: Studies in the Cultural Politics of Education*, 43(6), 900–914.
- Cespedes, M. (2025). Digital mentoring as a sustainability mechanism in teacher professional development. *Journal of Professional Learning*, 12(1), 1–18.

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Crescentini, A., & Ostinelli, G. (2023). Professional development of teachers in 14 European countries: A comparative study. *Professional Development in Education*, 49(3), 450–467.
- Critical Appraisal Skills Programme (CASP). (2018). CASP checklists. <https://casp-uk.net>
- Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society, and Education*, 7(3), 252–263.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development. *Educational Researcher*, 38(3), 181–199.
- Downey, J. A., & Luebeck, J. (2025). Place-attentive professional learning in rural and remote contexts. In *Handbook on Rural and Remote Education* (pp. 215–234). Springer.
- DuFour, R., & Eaker, R. (1998). *Professional learning communities at work*. Solution Tree.
- Fischer, C., Fishman, B., & Dede, C. (2018). Exploring fit and function: Using mixed methods to investigate teacher professional development and enactment. *Teaching and Teacher Education*, 71, 195–209.
- Firestone, W. A., Mangin, M. M., Martinez, M. C., & Polovsky, T. (2020). Content knowledge and professional development. In *Teacher Learning in the Workplace*. Springer.
- Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence. *BMJ*, 331(7524), 1064–1065.
- Gurgenidze, D., Gogoladze, N., & Khachidze, M. (2022). Shifting to remote teacher professional development in the context of the COVID-19 pandemic in Georgia. *International Journal of Education and Development Using Information and Communication Technology*, 18(1), 180–194.
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8(3), 381–391.
- Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.
- Hong, Q. N., Pluye, P., Fabregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT)*, version 2018. McGill University.

- Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes and efficacy. *Education Policy Analysis Archives*, 13(10).
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). ASCD.
- Kahraman, O., & Bicen, H. (2022). Examining the effects of artificial intelligence supported teacher professional development training. *Educational Technology Theory and Practice*, 12(2), 55–72.
- Koellner, K., Jacobs, J., & Borko, H. (2022). Situating professional development: The roles of expertise and community. In *Handbook of Research on Science Teacher Education*. Routledge.
- Kurteshi, V., & Rustemi, V. (2025). Impact of continuous online professional development on teacher performance: A longitudinal study from Kosovo. *Educational Process: International Journal*, 14(1), 1–20.
- Lay, C. D., Allman, B., Cutri, R. M., & Kimmons, R. (2020). Examining a decade of research in online teacher professional development. *Frontiers in Education*, 5, 167.
- Loyalka, P., Popova, A., Li, G., & Shi, Z. (2019). Does teacher training actually work? Evidence from a large-scale randomized evaluation of a national teacher training program. *American Economic Journal: Applied Economics*, 11(3), 128–154.
- Lu, M., Loyalka, P., Shi, Y., Chang, F., Liu, C., & Rozelle, S. (2019). The impact of teacher professional development programs on student achievement in rural China. *Journal of Development Effectiveness*, 11(2), 105–131.
- Lyu, C. (2025). Understanding expert EFL teachers' motivation for professional development: How school culture matters. *SAGE Open*, 15(1).
- McChesney, K., & Aldridge, J. M. (2021). What gets in the way? A new conceptual model for the trajectory from teacher professional development to impact. *Professional Development in Education*, 47(5), 837–855.
- McChesney, K., & Cross, J. (2023). How school culture affects teachers' classroom implementation of learning from professional development. *Learning Environments Research*, 26(3), 785–801.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282.

- Muhuro, P., & Kangethe, S. M. (2021). Prospects and pitfalls associated with implementing blended learning in rural-based higher education institutions in Southern Africa. *Perspectives in Education*, 39(1), 427–441.
- Nedungadi, P., Menon, R., Gutjahr, G., & Raman, R. (2024). Evaluating Holistic Education and Digital Learning model for advancing SDG4. *Cogent Social Sciences*, 10(1), 2358495.
- Nguyen, T. P., Duong, B.-H., Nhat, N. T. H., Vu, T. T. H., & Thu, H. V. T. (2026). Bridging readiness, relevance, and resources: A blended approach to teacher professional development in Vietnam. *European Journal of Education*, 61(1), e12678.
- Ngoveni, M. A., Mphuthi, G. T., Maile, L. M., Pitsoane, E. M., & Letseka, M. M. (2025). Willing but underserved: Interpreting digital training needs of Grade 8-12 teachers in a rural circuit. *Education Sciences*, 15(2), 145.
- Nkanyani, T. E., Mudau, A. V., & Sikhosana, L. (2024). Teaching and learning of physical sciences in rural schools through rural blended learning strategy. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(7), em2469.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
- Pawson, R. (2006). *Evidence-based policy: A realist perspective*. SAGE Publications.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE Publications.
- Philipsen, B., Tondeur, J., Pareja Roblin, N., Vanslambrouck, S., & Zhu, C. (2019). Improving teacher professional development for online and blended learning: A systematic meta-aggregative review. *Educational Technology Research and Development*, 67(5), 1145–1174.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Paul H. Brookes Publishing.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. ESRC Methods Programme.

- Popova, A., Evans, D. K., Breeding, M. E., & Arancibia, V. (2022). Teacher professional development around the world: The gap between evidence and practice. *The World Bank Research Observer*, 37(1), 107–136.
- Portillo, J., & Lopez de la Serna, A. (2021). An international perspective for improving teacher professional development for online and blended learning. *Educational Technology Research and Development*, 69(1), 25–28.
- Ramli, A. F., Rasul, M. S., & Omar, M. (2024). Issues and challenges in the implementation of professional development programs for design and technology teachers. *International Journal of Learning, Teaching and Educational Research*, 23(7), 1–20.
- Schein, E. H. (2010). *Organizational culture and leadership* (4th ed.). Jossey-Bass.
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews. *BMJ*, 358, j4008.
- Swayne, C., & Moynihan, J. (2025). Cultivating coaching cultures in schools: Insights from school leaders. *International Journal of Mentoring and Coaching in Education*, 14(2), 187–203.
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45.
- Tonga, F. E., Eryigit, S., Yalcin, F. A., & Erden, F. T. (2022). Professional development of teachers in PISA achiever countries: Finland, Estonia, Japan, Singapore and China. *Professional Development in Education*, 48(4), 588–604.
- UNESCO. (2023). *Global report on teachers: Addressing teacher shortages and transforming the profession*. UNESCO Publishing.
- Van Peteghem, H. K., Engels, N., Vandecandelaere, M., Plackle, I., & Consuegra, E. (2025). Transforming teacher beliefs through professional learning for powerful learning environments. *Professional Development in Education*, 51(2), 245–262.
- Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: Realist syntheses. *BMC Medicine*, 11, 21.
- Woodcock, S., & Hardy, I. (2025). Teacher self-efficacy, inclusion and professional development practices. *Professional Development in Education*, 51(1), 87–103.
- Wuryaningsih, Susilastuti, D. H., Darwin, M., & Pierewan, A. C. (2019). Effects of web-based learning and F2F learning on teachers' achievement in teacher training program in Indonesia. *International Journal of Emerging Technologies in Learning*, 14(21), 123–147.

- You, H., Park, S., Hong, M., & Warren, A. (2025). Unveiling effectiveness: A meta-analysis of professional development programs in science education. *Journal of Research in Science Teaching*, 62(4), 815–845.
- Youngs, P., & King, M. B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly*, 38(5), 643–670.
- Zhang, H., Lyu, X., & Qiu, Y. (2024). Effective teacher professional development and its influencing factors: A cross-national comparison. *Journal of Teacher Education*, 75(5), 561–578.
- Zhou, Q., Diao, J., Wang, Y., Liu, X., & Zhang, T. (2024). *Strategies for developing TVET teachers' professional competencies*. Springer Nature.
- Zhou, X., Padron, Y., Waxman, H. C., Cheng, W. C., & Acosta, S. (2024). How do school climate and professional development in multicultural education impact job satisfaction and teaching efficacy for STEM teachers. *International Journal of Science and Mathematics Education*, 22(8), 1675–1697.