African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

# Teacher Digital and AI Capability and Classroom Enactment in Primary Education: A Critical Audit of 27 Archived Studies

## Compétences numériques et en intelligence artificielle des enseignants et pratiques de classe au primaire : audit critique de 27 études archivées

A critical methodological review of an archived Scopus–ERIC corpus (2020–2026).

Auteur 1 : Charaf BOULAHDID.
Auteur 2 : Kenza EL KADIRI.

**Charaf BOULAHDID**
**Doctorant — Laboratoire des Sciences Économiques et Politiques Publiques (LSEPP), Faculté d'Économie et de Gestion (FEG), Université Ibn Tofaïl, Kénitra, Maroc.**

**Kenza EL KADIRI**
**Professeure — Laboratoire des Sciences Économiques et Politiques Publiques (LSEPP), Faculté d'Économie et de Gestion (FEG), Université Ibn Tofaïl, Kénitra, Maroc.**

**Review article · Teacher digital and AI capability · Primary education**

**Déclaration de divulgation : L'auteur n'a pas connaissance de quelconque financement qui pourrait affecter l'objectivité de cette étude.**

**Conflit d'intérêts : Les auteurs ne signalent aucun conflit d'intérêts.**

**Pour citer cet article : BOULAHDID .Ch & EL KADIRI .K (2026) « Teacher Digital and AI Capability and Classroom Enactment in Primary Education: A Critical Audit of 27 Archived Studies », African Scientific Journal « Volume 03, Num 39 » pp: 001 – 023.**

**DOI : 10.5281/zenodo.24540165.**
**Copyright © 2026 – ASJ**

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

**Résumé**

Les compétences numériques et celles liées à l'intelligence artificielle des enseignants, la préparation des activités et leur utilisation en classe sont souvent étudiées comme des réalités équivalentes, alors qu'elles nécessitent des preuves différentes. Cette revue critique méthodologique rétrospective examine un corpus délimité issu de Scopus et d'ERIC selon un critère d'inclusion unique : une compétence distinctement étudiée chez l'enseignant et une activité pédagogique identifiable, préparée ou déclarée/observée comme réalisée. La recherche Scopus archivée comprend 123 notices et 57 textes intégraux évalués ; le registre ERIC disponible comprend 31 candidats après application de filtres et 27 textes intégraux évalués. Le corpus analytique réunit 27 études (17 Scopus et 10 ERIC). Les preuves de pratique comprennent trois comptes rendus externes, trois artefacts pédagogiques, 17 descriptions substantielles par les enseignants et quatre déclarations brèves. Vingt études proposent explicitement des recherches empiriques futures. Après exclusion des quatre preuves les moins directes dans une analyse restrictive, 23 études subsistent, dont 17 comportent ces propositions. Le corpus documente rarement une même activité depuis sa préparation jusqu'à sa réalisation en classe et sa réutilisation ultérieure. Un cadre de suivi par activité est proposé pour de futures études ; il ne constitue pas un modèle causal testé. La recherche est archivistique, la requête ERIC originale n'est pas entièrement documentée, et les décisions finales d'éligibilité ainsi que les évaluations méthodologiques ne disposent pas d'une validation indépendante.

Mots-clés : enseignement primaire ; compétences numériques des enseignants ; littératie en IA; pratiques pédagogiques réalisées ; provenance des preuves ; revue méthodologique.

**Abstract**

Teacher digital and AI capability, lesson preparation and classroom use are often treated as equivalent, although they require different evidence. This retrospective critical methodological review audits a bounded Scopus–ERIC corpus using one eligibility rule: a distinct teacher capability construct and a separable instructional activity, recorded as preparation or reported/observed use. Scopus yielded 123 records and 57 assessed full texts; the available ERIC register contains 31 post-filter candidates and 27 assessed full texts. The analytic corpus includes 27 studies (17 Scopus and ten ERIC). Practice evidence comprises three external records, three instructional artefacts, 17 substantive teacher accounts and four brief self-reports. Twenty reports contain explicit proposals for further empirical research. A four-report low-directness restriction leaves 23 studies and 17 explicit proposals. Across the corpus, measures of teacher capability seldom identify one activity followed through preparation, classroom delivery and later reuse. An episode-linked framework is proposed for future investigation; it is not a tested causal model. The search is archival, the original ERIC query is incompletely documented, and independent final eligibility and quality appraisals are not available.

Keywords: primary education; teacher digital competence; AI literacy; enacted pedagogy; evidence provenance; methodological review

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

## 1. Introduction

Digital tools and AI are increasingly part of primary education, but their use differs considerably between classrooms. Teachers may feel confident with a resource, prepare an activity around it, or use that activity with pupils. These are related experiences, yet they provide different kinds of evidence. This distinction becomes important when "technology integration" is used to cover knowledge, intention, preparation and actual instruction without separating them (Mishra & Koehler, 2006; Redecker & Punie, 2017).

The reviewed studies illustrate this problem. Pflaumer et al. (2021) describe how German primary teachers appropriated adaptive literacy games in their routines. Sothayapetch and Lavonen (2022) examine teachers' accounts of science teaching during disruption in Thailand and Finland. Kim and Kwon (2023) analyse competencies for AI-supported teaching in South Korea. In contrast, Rhodes and Calder (2022) provide direct information about lessons implemented in a digital curriculum project. These studies do not give the same type of answer, even when they all concern teachers and technology.

The evidence also covers crisis-driven instruction and teachers' reflections on remote teaching. Jung et al. (2024) describe digital literacy challenges in remote teaching, while Shiu (2024) examines primary English teachers' learning and adaptation during COVID-19. Their accounts help interpret professional mobilisation, but they should not be treated as independent demonstrations of durable effects.

Research on this problem is not new. Claro et al. (2024) reviewed digital-competence measures. Avci et al. (2026) assessed their use across K–12 studies. Forsström and Amdam (2026) examined relational and institutional conditions. Kokoç et al. (2026) mapped in-service teachers' AI literacy. Lee (2026) considered teacher intervention in AI-supported lessons. Abounajem (2026) compared reported and enacted digital competence, while Patiño Hernández et al. (2026) examined AI-responsive teacher judgement. More directly, Mensah et al. (2026) compared planning knowledge with teachers' accounts of science teaching. Xiu et al. (2026) found that generative AI was used more often for preparation than for student-facing lessons. Walan and Vinerean-Bernhoff (2026) followed a primary STEM professional-development project that included classroom implementation. These studies mean that a general planning–practice gap cannot credibly be presented as a new discovery. The remaining contribution is an audit of what each source of evidence allows a reader to infer: a prepared resource, a reported or observed lesson, and sustained reuse are different outcomes. We ask whether the archived record can connect these outcomes for the same identifiable lesson, and what a future study would need to document to test that connection.

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

We examine 27 studies selected under the same eligibility rule and distinguish teacher capability, instructional preparation, reported delivery, external records and later reuse. Our contribution is an audit of the warrant for each inference, not a claim that the preparation–practice gap has never been examined. The conclusions apply to this bounded archive rather than to the field as a whole.

### 1.1 Problem statement and questions

A teacher can know how to design an activity without using it in class; an activity delivered once may never become part of a regular teaching routine. Even repeated delivery cannot, on its own, establish gains in pupil learning. We therefore examine where instructional follow-through may occur or break down. This raises a theoretical question about continuity under different school conditions and a methodological question about how the same activity can be documented over time.

The guiding question is: How far does the 27-study archived record support inferences from teacher capability to preparation, delivered teaching and later reuse, and where does the evidence stop? Four questions follow. RQ1: Which measures of teacher capability and professional resources appear? RQ2: Which distinctions between preparation and delivery are documented, reported only or still uncertain? RQ3: Which contextual conditions are described rather than formally tested? RQ4: Which source-located limitations and author-proposed studies inform a feasible lesson-level research design?

## 2. Methodology

### 2.1 Review type and design

This study is a retrospective critical methodological review of an archived empirical corpus. It does not claim a new or exhaustive systematic search. PRISMA 2020 provides the terminology used to describe archived selection stages and subsequent amendments (Page et al., 2021). PRISMA-S provides a benchmark against which search-provenance gaps are disclosed (Rethlefsen et al., 2021). SWiM informs the transparent reporting of numerical findings that cannot be pooled (Campbell et al., 2020). The thematic comparison draws on Thomas and Harden (2008). No prospective protocol registration is recorded; the PRISMA and SWiM references do not amount to a claim that this review fulfils all items of those guidelines. Research published after the August 2026 searches was used only to assess the scope of the novelty claim. These comparator articles were not added to the 27-study analytic corpus.

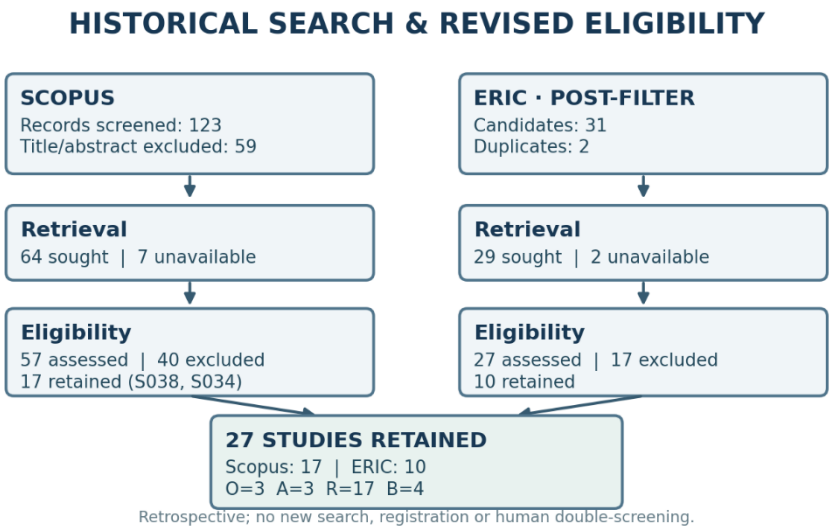
### 2.2 Information sources and selection

The archived Scopus search (8 August 2026) combined primary education, teacher digital or AI capability, and pedagogical use (Supplementary Appendix A). It returned 123 records.

Title/abstract screening excluded 59; 64 full texts were sought, seven were unavailable, and 57 were assessed. Applying the eligibility rule in Section 2.3, 40 full texts were excluded and 17 retained. The ERIC record (18 August 2026) begins with 31 post-filter candidates; after two duplicates and two unavailable reports, 27 full texts were assessed, 17 excluded and ten retained. The resulting analytic set comprises 27 unique studies (17 Scopus; ten ERIC). Figure 1 gives the selection arithmetic; the decision ledger and source locators appear in Supplementary Appendices D, F and K.

The complete original ERIC search string and unfiltered hit count are not archived. The ERIC figure of 31 therefore describes candidates after interface filters, not the complete database yield. No new database search or Web of Science search was undertaken. The resulting flow is a reconstruction of the documented selection, not evidence of exhaustive retrieval.

Figure 1. Selection accounting for the archived Scopus and ERIC records under the uniform eligibility criterion (17 + 10 = 27 studies). ERIC candidate counts begin after interface filtering.




### 2.3 Single eligibility criterion and evidence classification

Eligible empirical reports had to (1) examine in-service primary teachers or a separately analysable primary-teacher subgroup; (2) measure or qualitatively analyse a distinct teacher digital or AI capability (X); and (3) document a separate digital- or AI-supported instructional activity (I), either as an identifiable prepared resource (I-P) or as a concrete classroom action reported as completed (I-D). For I-D, external documentation, a substantive account or an explicit retrospective action item is admissible. Confidence, intention, training attendance and statements about skills acquired do not qualify on their own. Evidence-source codes O, A, R

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

and B describe directness, not study quality; a brief self-report (B) neither establishes independent observation nor proves sustained use or effectiveness. E01 marks ineligible or inseparable teacher populations, E03 a missing distinct instructional activity, and E04 a missing distinct teacher capability construct. Pupil attainment was not required for inclusion. The source-located operational decisions are documented in Supplementary Appendices B, D, F and K.

Eligibility was assessed retrospectively from archived records. The original Scopus two-column decisions covered 57 full texts (46 agreements; Cohen's $\kappa = 0.616$), but this statistic does not validate the present corpus. A later source-based check of the 13 E03/E04 decisions was supported by AI assistance and was not an independent human second screen; the ERIC record also lacks documented independent double-screening. We do not report a new agreement coefficient. These limitations are detailed once in the reproducibility ledger (Supplementary Appendix J).

The boundary cases demonstrate the distinction between eligibility and evidence quality. Zhou et al. (2022; S038) separately report technological knowledge and self-reported ICT use in class (37%; Tables 4–5). Balakrishnan et al. (2025; S034) analyse digital competence and include a questionnaire statement on PLC-assisted use of online platforms for learning and assessment (Table 3, item 6). Both satisfy the low-directness activity threshold applied to E006 and E007, so all four are coded B. S034 is the weaker of the two because the questionnaire wording does not independently verify completed lessons; none of these B records is treated as external observation. Appendix K provides the source locators.

### 2.4 Data extraction, analytic units and critical appraisal

Extraction covered study design, teacher digital/AI capability, preparation or reported/observed instructional action, evidence-source directness, study limitations and research recommendations. The 27 eligible reports include 20 containing an explicit author proposal for further empirical research under first-analyst coding. S023 is excluded under E04 because it discusses AI-education inequalities and teacher training but does not separately analyse teacher digital/AI capability as the required X construct. None of the revised extraction or future-proposal labels has been independently double-coded.

We assessed each inference against the source design and type of practice evidence, without converting O/A/R/B codes into quality grades. JBI/MMAT working sheets exist, but complete item-level appraisals have not been independently reconciled for the 27-study corpus (Aromataris et al., 2024; Hong et al., 2018). We therefore report no formal overall risk-of-bias

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

or certainty rating. GRADE and GRADE-CERQual were not applied. Study limitations are linked to source locations in the supplement.

### 2.5 Synthesis approach

The review separated teacher capability (X), mobilisation resources (M), preparation (I-P), delivered teaching (I-D), context (W) and outcomes (Y), where reported. Evidence of practice was then grouped by source: external observation or rating, instructional artefact, substantive retrospective account, or brief self-report. Numerical findings remained in their original units. They were not combined across incompatible outcomes or designs. Finally, the synthesis compared stated limitations, future-research proposals and relevant counterexamples. The unit of interpretation was the recorded result or teaching activity, not an assumed article-level effect (Thomas & Harden, 2008; Campbell et al., 2020).

### 2.6 Evidence provenance and robustness checks

The descriptive crosswalk records P = 3 (prepared activities), D = 14 (delivery described or externally documented), D? = 7 (low-specificity or uncertain delivery), and P/D? = 3 (insufficient separation), totalling 27. The two B-coded boundary cases S038 and S034 are assigned D?, not externally observed delivery. O/A/R/B independently describe how practice was evidenced. Neither D nor B establishes that the same named activity was followed through preparation, classroom delivery and later reuse (Supplementary Appendix I).

A sensitivity analysis sets aside all four B-coded studies (E006, E007, S038, S034) consistently. The restricted descriptive subset contains 23 studies (O = 3; A = 3; R = 17), including 17 explicit future-research proposals rather than 20. This is a restriction of evidential directness, not a second PRISMA selection, an effect estimate or a quality appraisal.

The restricted subset permits a check that the proposed gap does not depend solely on brief practice indicators. It cannot establish causal effects or repair incomplete search documentation and the absence of independent final adjudication. The source-by-source distinctions remain central to interpretation.

## 3. Concepts, evidence and paradoxes in the literature

### 3.1 Capability, readiness and teaching action

Teacher digital capability covers knowledge and skill used to select, evaluate, adapt and produce digital resources. AI-related capability may also include evaluating model output, pedagogical judgement and ethical concerns. TPACK distinguishes technological, pedagogical and content knowledge (Mishra & Koehler, 2006); DigCompEdu expands the educator competence domains (Redecker & Punie, 2017); and UNESCO's framework identifies AI-related professional competences (Miao & Cukurova, 2024). None of these frameworks, by

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

itself, documents a completed lesson. Interviews about teachers' self-perceived digital literacy, such as those analysed by Değirmenci Gündoğmuş (2024), also require separation from direct evidence of delivery. Li et al. (2025) illustrate the value of context-rich accounts of digital mathematics use, while remaining a self-reported source for enactment.

Self-efficacy, perceived usefulness and training participation were coded separately as mobilisation resources. Beliefs can influence behaviour but do not establish that behaviour took place (Bandura, 1997). Gejandran and Abdullah (2025), Chen et al. (2024) and Musgrove et al. (2021) provide different associational models. Their outcomes and statistical paths must not be interpreted as the same verified change in teaching.

### 3.2 Preparing a lesson is not delivering it

An artefact can show that a teacher prepared a technology-supported resource. It does not necessarily show how the resource was used with pupils. Hastomo et al. (2026) describe GenAI-assisted materials. Angula and Simuja (2025) document technology-integrated lesson design. Gonzalez and Mohamad (2022) examine virtual-literacy materials. These reports contribute evidence of preparation. By contrast, Rhodes and Calder (2022) document classroom implementation; Wilaphan and Rattanawongsa (2026) use coach ratings; and Muweesi et al. (2021) include classroom observation. Even these stronger sources do not automatically establish a general causal effect on learning.

### 3.3 An evidence continuum, not a hierarchy of research quality

Table 1. Source-of-evidence codes. These categories do not replace JBI/MMAT appraisal.

| Practice evidence | Operational meaning | Main caution |
|---|---|---|
| O — External record | Observation, coach rating or documented implementation | Observer reliability, local scope and design still need appraisal |
| A — Artefact | Lesson product, teaching material or demonstrable preparation | Creation does not establish delivery |
| R — Substantive account | Teacher describes a completed action in some detail | Retrospective interpretation and self-report bias |
| B — Limited report | Brief baseline account or single-use indicator | Low specificity; retain separately in sensitivity analyses |

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

## 4. Development of research and comparative evidence (2020–2026)

### 4.1 Teaching adaptations and digital routines (2020–2022)

Earlier reports in the archived period often focus on how teachers adapted existing teaching routines. Pflaumer et al. (2021) analyse the appropriation of literacy games; Sarı and Keser (2021) examine online teaching experiences; and Sothayapetch and Lavonen (2022) compare science instruction during disruption. Rana et al. (2022) draw attention to rural professional learning and resource constraints. This body of work makes implementation visible, but much of the evidence comes from teacher accounts in particular contexts. These examples describe the composition of the selected archive, not a quantified trend across all published studies. The retained Zhou et al. (2022) study extends this early period with separately analysed teacher technological knowledge and a limited classroom-use item; it cannot demonstrate delivery from direct observation.

One reason for separating teaching activity from teacher attributes is the unit of observation. An interview about using a platform captures the teacher's account of a completed action. It cannot establish how consistently the practice occurred or what pupils learned. An observational case report documents particular classroom events but may not estimate how common they are across a school system. Across these early studies, the relevant question is not which method is intrinsically superior; it is whether the method is suitable for the conclusion being drawn (Pflaumer et al., 2021; Sarı & Keser, 2021; Sothayapetch & Lavonen, 2022).

### 4.2 Competence constructs and contextual conditions (2023–2024)

During 2023–2024, research in this archived sample increasingly specified knowledge constructs and associated professional beliefs. Kim and Kwon (2023) developed elementary AI-teaching competencies; Chen et al. (2024) examined digital competence and perceived autonomy support; and Değirmenci Gündoğmuş (2024) described teachers' perceptions of digital literacy. Shiu (2024) offered a qualitative account of primary teachers' learning during online teaching. These studies illuminate professional capability and reported adaptations. Their methods differ, however, in what they can demonstrate about completed instruction. None of these individual observations should stand for a general period trend.

### 4.3 Generative AI, professional learning and evidence of action (2025–2026)

More recent archived reports examine AI and digital professional learning. Jiang et al. (2026) study AI-TPACK constructs, Hussein et al. (2026) report experiences after Gemini-related professional development, and Hastomo et al. (2026) examine AI-assisted lesson materials. Wilaphan and Rattanawongsa (2026) evaluate coaching with a pre/post design. Their research questions and indicators differ substantially. The growth of AI-related topics in this selected

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

corpus cannot be equated with an increase in independently verified classroom change. Kayınova and Adıgüzel (2026) also describe in-service teachers' views of completed Web 2.0 uses in primary classrooms; these remain retrospective accounts, not observation.

Generative AI makes this inferential distinction especially consequential. The ability to evaluate AI output, write prompts or design resources can be assessed without confirming that the materials reached pupils. Hastomo et al. (2026) provide valuable evidence about the materials teachers produce. Hussein et al. (2026) examine professional development and accounts of transformed practice. By contrast, Wilaphan and Rattanawongsa (2026) use repeated assessments but lack an untreated comparison. These are complementary windows onto practice, not interchangeable demonstrations of AI effectiveness. The same distinction applies to mathematics and AI-TPACK constructs examined by Jiang et al. (2026). Balakrishnan et al. (2025) further show how a professional-learning-community survey can describe technology use while leaving the actual lesson-level sequence unverified.

### 4.4 Studies across educational contexts

The cross-context studies show why opportunity conditions should not be treated as a residual variable. Rural accounts in Nepal and Namibia document limits related to learning opportunities and resources (Rana et al., 2022; Angula & Simuja, 2025). Teachers in New Zealand report locally specific integration practices (Niu & Niranjan, 2025), and studies from Türkiye and South Korea examine different pedagogical and AI-competency settings (Alptekin & Taneri, 2025; Kim & Kwon, 2023). Liu and Zhang (2025) describe urban–rural AI-education inequalities in China, but their school-level analysis lacks a separate teacher-capability X construct and is therefore an excluded contextual comparator (E04), not evidence within the included 27-study synthesis. Contrasting these settings motivates future cross-school tests, not a pooled "context effect".

## 5. Theoretical foundations and analytical framework

### 5.1 Teacher knowledge frameworks

TPACK and related frameworks show why effective use of technology requires integration of several domains of professional knowledge. DigCompEdu and UNESCO also emphasise resource judgement, ethical responsibility and learner support (Mishra & Koehler, 2006; Redecker & Punie, 2017; Miao & Cukurova, 2024). These frameworks justify examining capability as a multidimensional construct. They do not imply that knowledge flows directly into instruction.

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

### 5.2 Mobilisation and opportunity conditions

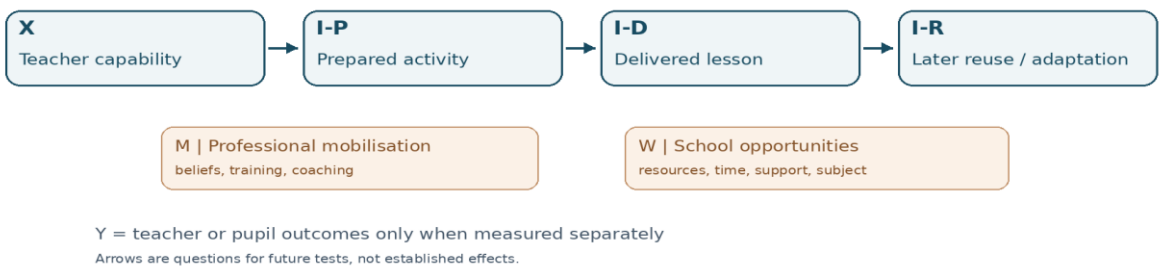
A mobilisation perspective asks what helps teachers act on their knowledge. Self-efficacy provides a theoretical basis for distinguishing belief in capability from execution of a task (Bandura, 1997). Professional development, coaching, feedback and collaboration can create opportunities to practise; access, time, leadership and connectivity can constrain those opportunities. In the empirical record, these resources are not always measured separately or linked to an observed action. Context is therefore treated as a plausible boundary condition, not a statistical moderator unless an individual study tests an interaction.

### 5.3 A construct–action–warrant view of evidence

The analytic framework distinguishes capability (X), mobilisation resources (M), preparation (I-P), delivered teaching (I-D), and later reuse or adaptation (I-R). Contextual conditions (W) are recorded across these stages. Outcomes (Y) are examined only where measured. A separate check asks whether each inference is supported by the source design. The framework organises evidence. It has not been tested as a causal model.

The framework also separates two forms of validity. Construct validity asks whether X really represents technological and pedagogical capability rather than confidence, general attitude or mere training exposure. Inferential validity asks whether the recorded I-P or I-D evidence permits claims about delivery, improvement or persistence. The latter cannot be borrowed from the former. This distinction helps interpret associations in models such as those of Musgrove et al. (2021) and Gejandran and Abdullah (2025), where the statistical path is meaningful within the observed variables but not evidence of a fully observed classroom sequence.

Figure 2. Proposed episode-linked coding framework. I-P is an instructional artefact, I-D is delivered teaching, and I-R is later reuse or adaptation. M and W are recorded separately. Arrows are questions for future tests, not causal findings.




## 6. Synthesis of findings and theoretical gap

### 6.1 Composition of the evidence

The 27-study corpus contains 17 Scopus and ten ERIC reports. Evidence-source codes are O = 3 (E004, S080, E010), A = 3 (E003, S048, S092), R = 17 and B = 4 (E006, E007, S038, S034).

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

O identifies external records or assessments; A, instructional products; R, substantial accounts of completed activity; and B, brief self-reported indicators. These four categories exhaust the 27 studies, but describe evidence source, not risk of bias or instructional effectiveness.

The reconciliation follows a single, transparent threshold for weak enacted-use indicators: a teacher must report a distinct classroom action, however briefly, while confidence or a claim of newly acquired skill alone does not qualify. The historical twelve E03 changes comprise ten retained exclusions and two reversed decisions; the later S023 E04 exclusion remains. These are documented first-analyst decisions, with independent human adjudication pending. Thus 27 is the internally consistent working denominator for the article and its supplement, not a claim of final two-reviewer consensus.

Table 2. Selection accounting for the 27-study Scopus–ERIC corpus; ERIC figures are post-filter at identification.

| Selection stage | Scopus | ERIC | Combined |
|---|---|---|---|
| Identified / candidates | 123 | 31 post-filter | ERIC begins post-filter |
| Excluded at TA / duplicate removal | 59 | 2 | Procedures differ |
| Full texts assessed | 57 | 27 | 84 |
| Full-text exclusions | 40 | 17 | 57 |
| Included under the uniform eligibility rule | 17 | 10 | 27 |

The activity-stage labels independently classify the 27 reports as D = 14, D? = 7, P/D? = 3 and P = 3. A D label denotes documented or reported delivery, not necessarily an observed classroom; D? identifies a weaker or ambiguous account. None of these labels demonstrates repeated use or measured pupil improvement. The study-level crosswalk is given in Supplementary Appendix I.

It would be inaccurate to say that classroom practice is never observed. Rhodes and Calder (2022) report an implemented digital-curriculum project, Muweesi et al. (2021) draw on classroom observations, and Wilaphan and Rattanawongsa (2026) use externally assessed teaching. These constitute three O-coded reports. The three A-coded reports contain instructional products, not the same level of evidence of delivered lessons. Seventeen R-coded reports rely on richer teacher accounts; four B-coded reports give limited use indicators. This

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

variation matters because a stronger competence score and an observed teaching change are different outcomes, even if both concern “integration”.

### 6.2 Why estimates cannot be pooled

Jiang et al. (2026) report an AI-TPK coefficient of $\beta = 0.29$ in a construct model. Gejandran and Abdullah (2025) report an indirect coefficient of $\beta = 0.048$ ($p = .008$) for a different self-report pathway. Wilaphan and Rattanawongsa (2026) report a pre/post change from a mean of 2.22 to 2.58, without an untreated comparison that would isolate an intervention effect. Rahayu et al. (2024) examine self-assessed TPACK after training; Suryanti et al. (2026) report a brief teacher-rated digital integration indicator. These figures describe different constructs and estimands: a model coefficient, a model-based indirect path, and an uncontrolled before–after mean. Even a common rescaling would not turn them into comparable measures of observed classroom quality or pupil attainment. We therefore report no pooled effect, standardized vote count, pooled direction, or generalized intervention effect (Campbell et al., 2020).

### 6.3 Strengths of the evidence and limitations acknowledged by study authors

Several recommendations point to different missing links. Pflaumer et al. (2021, PDF p. 16) call for longer study of appropriation. Jiang et al. (2026, PDF p. 13) request longitudinal and classroom observations. Hussein et al. (2026, PDF p. 12) recommend checking interviews against artefacts, plans and observations. Sarı and Keser (2021, PDF p. 16) explain that classroom observation was not possible. Gonzalez and Mohamad (2022, PDF p. 28) reach a similar conclusion for virtual literacy instruction. These are explicit limitations of their designs. They should not be generalised to every study in the corpus.

Other authors identify the challenge at a different point. Hastomo et al. (2026) document the design of GenAI-assisted materials but cannot use the artefact alone to establish classroom delivery. Angula and Simuja (2025) show professional learning through lesson design and call for wider, more complete examination of implementation. Chen et al. (2024) warn about common-method bias and cross-sectional inference. Wilaphan and Rattanawongsa (2026) include external assessment, yet their pre/post study lacks an untreated comparison. These weaknesses are not interchangeable: absence of observation concerns verification, while absence of a comparison group concerns causal attribution. Zhou et al. (2022) explicitly acknowledge reliance on quantified self-reported interviews; Balakrishnan et al. (2025) study professional-learning communities using survey and interview data but not independent verification of the platform-use episode.

The record also contains positive counterexamples. Rhodes and Calder (2022) describe completed lessons within an action-research project. Muweesi et al. (2021) report classroom

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

observation in schools with constrained ICT resources. Wilaphan and Rattanawongsa (2026) use outside assessments of teaching. Pflaumer et al. (2021) follow appropriation across 12 weeks. Thus, practice and time are not wholly absent. The narrower issue is whether a prepared activity can be tracked as the same activity through delivery and later adaptation. Our extraction did not establish such a linked record across all these steps within a common study design.

### 6.4 Evidence-threshold sensitivity analysis

Under the inclusive working rule, 27 studies are retained (O = 3, A = 3, R = 17, B = 4), and 20 have source-located explicit future-research proposals. A conservative descriptive restriction excluding all four B-coded reports yields 23 (O = 3, A = 3, R = 17) and 17 explicit proposals. The controlled-design proposal tag declines from two to one because E007 is omitted; the pedagogical/pupil-outcome tag declines from five to three after omitting E007 and S034. These are changes in descriptive tags, not inferential robustness tests. The sensitivity analysis does not alter the 27-study selection flow.

The methodological conclusion does not depend on the decision to admit the two borderline reports. Even after all B sources are excluded, only three studies are coded O and three A, while the remaining 17 supply substantive teacher accounts. The distinction between documentation strength and the quality of an individual research design remains essential. A new reconciled risk-of-bias exercise would be needed to make stronger quality statements.

### 6.5 Author-proposed research directions

Explicit future empirical work is proposed in 20 of the 27 reports under first-analyst coding. The initial 18 included studies are supplemented by Zhou et al. (2022), who propose more complete TPACK measurement and quantitative–qualitative research, and Balakrishnan et al. (2025), who call for examination of PLC effects over time, including pupil learning outcomes. Selected non-exclusive themes are broader sampling/comparison (9 studies), school conditions (9), longitudinal follow-up (6), multiple evidence sources (4), pedagogical/pupil outcomes (5), and controlled designs (2). The Zhou proposal additionally concerns mixed-method instrument development, which is not represented by these six selected theme headings. The counting is descriptive, not independently double-coded; individual sources appear in the supplement.

These recommendations do not all ask the same question. Larger samples test transferability. Classroom observation checks whether reported practices occur. Follow-up distinguishes a short-lived change from a routine. Controlled comparisons can identify intervention effects. Combining these designs would permit a new analysis of how planned activities reach the classroom and persist. None of the current counts demonstrates how often this transfer succeeds.

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

Table 3. Selected, non-exclusive empirical research themes among 20 source-identified proposals in 27 studies; first-analyst counts.

| Theme in authors' explicit recommendations | Reports | Interpretation for future design |
|---|---|---|
| Sample diversity and comparison | 9 | Replicate across teacher profiles and school settings |
| Context and institutional conditions | 9 | Measure infrastructure, support, time and school differences |
| Longitudinal follow-up | 6 | Observe how teaching changes and persists after training |
| Multiple evidence sources | 4 | Combine interviews with observation, artefacts or logs |
| Pedagogical and pupil outcomes | 5 | Separate teacher practice from measured pupil outcomes |
| Controlled designs | 2 | Use comparison groups when claiming intervention effects |

### 6.6 Contribution: a source-to-inference audit

We propose an episode-linked approach to studying instructional follow-through. Rather than treating a teacher's overall integration score as the unit of analysis, it follows one identifiable activity. The teacher first prepares a task or material (I-P), may then deliver and adapt it in class (I-D), and may later reuse or revise it (I-R). Capability (X), professional mobilisation (M) and school conditions (W) remain distinct measures. Future studies could test which factors accompany transitions between these stages; the present review does not establish a causal model.

This question is narrower than the familiar knowledge–practice gap. It also goes beyond comparing planning scores with reported use. Mensah et al. (2026) examine that comparison in K–8 science. Walan and Vinerean-Bernhoff (2026) document primary teachers working through planning, teaching and reflection. These studies show that the wider issue is not new. The unanswered question within this archived corpus is whether the same activity can be followed from preparation through delivery to later reuse, while school conditions are recorded. The review cannot establish that no study elsewhere has addressed this question.

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

The contribution is an explicit source-to-inference audit rather than the discovery of a previously unknown gap. It distinguishes teacher-level X from lesson-level I, separates prepared artefacts from reported and externally corroborated delivery, and refuses to equate repeated assessment with later reuse of the same activity. It makes an otherwise implicit inferential boundary visible: a teacher-account study can support an interpretation of professional reasoning without verifying each classroom episode; an artefact verifies preparation without verifying delivery; an observation documents an episode without by itself estimating an intervention effect. The appendix makes those boundaries reproducible as a coding proposal. This is a methodological contribution grounded in a limited archive, not validation of a new theoretical pathway.

### 6.7 Falsifiable research propositions, not findings

P1 — Linked evidence. Tracking a named activity from preparation to observed delivery may reveal differences that competence and intention scores do not capture. P2 — School opportunities. Access, time, support and subject requirements may shape whether a prepared activity is delivered. These conditions need direct measures; they are not assumed statistical moderators. P3 — Continuity. Repeated observations after professional development may distinguish one-time use from routine reuse or adaptation. These propositions require longitudinal or comparative designs and measures specified in advance. They are not causal findings from the present review.

A prospective test would register the activity identifier and date of preparation before instruction, retain the dated artefact, record delivery using a classroom observation or timestamped implementation log, and document reuse or adaptation at a later, explicitly defined teaching opportunity. X, M and W would be measured separately, and an independent reviewer would link events using a prespecified decision rule. At least one counterexample should be sought (a prepared but undelivered lesson, or delivery without later reuse). A comparison group and pupil outcome measures would be added only if an intervention-effect question is pursued. This is a proposed design, not a procedure carried out in the present review.

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

## 7. Conclusion, implications and limitations

In the 27-study archive, capability, lesson preparation and classroom delivery remain supported by different types of records. Three studies are coded as external records of teaching, three as instructional artefacts, 17 as substantive accounts and four as brief self-reports. These are evidence-provenance categories, not risk-of-bias grades. The available measures cannot be pooled into a common intervention or student-learning effect. Nor has the extraction demonstrated a verified sequence of preparation, delivery and later reuse for the same named instructional activity.

The article's contribution is an explicit link between the kind of source and the inference it warrants—not a claim to have discovered the planning–practice gap. Existing studies already examine knowledge, professional development, design, classroom use and implementation. The two additional low-directness reports clarify rather than eliminate the problem: self-reported activity is different from externally verified delivery and from repeated lesson use. The 20 source-located proposals for future empirical work help specify a feasible next study, but their coding is not independently reconciled.

A future study could give each prepared lesson a unique identifier, assess relevant teacher capability, retain the instructional artefact and record classroom delivery before returning to the activity in a later teaching cycle. Interviews could help explain why a lesson was used, adapted or abandoned. School resources, leadership support and available time should be measured before contextual interactions are tested. Pupil outcomes would require a separate, appropriately designed assessment.

The review is limited to an archived search, without a recoverable complete ERIC query, an unfiltered ERIC yield, a Web of Science search or a prospective protocol. The final 27-study eligibility decisions and first-analyst future-research classifications have not been independently adjudicated. JBI/MMAT appraisals have not been reconciled. These limits restrict generalisability and confidence; no causal or prevalence inference about the wider literature is warranted.

The most informative next step is to observe identifiable instructional episodes across preparation, delivery and later reuse, recording contextual constraints separately. Our contribution is an auditable distinction between those forms of evidence, not a demonstration that one causes another.

Scope of claims. This manuscript is a retrospective critical methodological audit of a bounded 27-study corpus. Its selection flow is reconstructed from archived records and the stated eligibility criterion. Incomplete ERIC search provenance, lack of final independent adjudication

and unreconciled appraisal prevent claims of exhaustive coverage or validated certainty. The episode-linked framework is a testable research proposal, not a demonstrated causal process.

**Declarations and source transparency**

Protocol and registration. This retrospective critical review was not prospectively registered.

Data availability. Search counts, the uniform eligibility rule, study identifiers, source-located decisions, extraction codes and descriptive sensitivity checks are reported in the supplementary material. Original publisher PDFs are not redistributed.

Ethics. This review uses published research only. No new participants were recruited. Ethics approval and individual consent were not required for this documentary analysis.

## References

Abounajem, N. (2026). Evaluating the enactment of teacher digital competence: A convergent mixed-methods approach. Journal of Information Technology Education: Research, 25. https://doi.org/10.28945/5853

Alptekin, Z., & Taneri, A. (2025). Technology integration in pedagogical processes: Digital competence and teaching practices of primary school teachers in Turkey. Discover Education, 4, 351. https://doi.org/10.1007/s44217-025-00646-9

Angula, M., & Simuja, C. (2025). Evolution of TPACK of rural primary school mathematics teachers through technology-integrated lesson design in Namibia. South African Journal of Education, 45(4), Article 2520. https://doi.org/10.15700/saje.v45n4a2520

Aromataris, E., Lockwood, C., Porritt, K., Pilla, B., & Jordan, Z. (Eds.). (2024). JBI Manual for Evidence Synthesis. JBI. https://doi.org/10.46658/JBIMES-24-01

Avci, H., Pedersen, S. J., & Thompson, C. (2026). Digital competence of in-service teachers in K–12 education: A systematic literature review. Review of Education, 14(2), e70172. https://doi.org/10.1002/rev3.70172

Balakrishnan, S. V., Muthukrishnan, P., Prasad, R. D., Boonroungrut, C., & Saroinsong, W. P. (2025). A mixed-method study on professional learning communities in enhancing the digital competence of primary schools English teachers in Malaysia. Computer-Assisted Language Learning Electronic Journal, 26(4), 488–511. https://callej.org/index.php/journal/article/view/618

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. BMJ, 368, l6890. https://doi.org/10.1136/bmj.l6890

Chen, A., Li, W., & Fu, W. (2024). Unleashing digital superheroes: Unravelling the empathy factor in digital competence and online teacher autonomy support. British Journal of Educational Technology, 55(4), 1790–1810. https://doi.org/10.1111/bjet.13433

Claro, M., Castro-Grau, C., Ochoa, J. M., Hinostroza, J. E., & Cabello, P. (2024). Systematic review of quantitative research on digital competences of in-service school teachers. Computers & Education, 215, 105030. https://doi.org/10.1016/j.compedu.2024.105030

Değirmenci Gündoğmuş, H. (2024). Classroom teachers' perceptions of their own digital literacy. Educational Research and Reviews, 19(5), 79-83. https://doi.org/10.5897/ERR2024.4398

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

Forsström, S. E., & Amdam, S. H. (2026). Relational aspects of teachers' professional digital competence: A systematic review of systematic reviews. Humanities and Social Sciences Communications, 13, 834. https://doi.org/10.1057/s41599-026-06793-0

Gejandran, P., & Abdullah, N. (2025). Developing a model to explain teacher effectiveness: The mediating role of gamification. Cogent Education, 12(1), 2521154. https://doi.org/10.1080/2331186X.2025.2521154

Gonzalez, M., & Mohamad, G. A. (2022). Virtual literacy instruction: An investigation of how elementary educators exhibited TPACK during COVID-19 school closures. Journal of Pedagogical Research, 6(5), 54-88. https://doi.org/10.33902/JPR.202216914

Hastomo, T., Widiati, U., Ivone, F. M., & Zen, E. L. (2026). Integrating generative AI in primary English material design: Insights from Indonesian teachers' perceptions. Journal of Elementary Education, 19(1), 77-91. https://doi.org/10.18690/rei5409

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. Education for Information, 34(4), 285–291. https://doi.org/10.3233/EFI-180221

Hussein, A., Daher, W., & Assali, A. (2026). Primary school teachers' I-TPACK competencies after Gemini-based professional development. Frontiers in Education, 11, 1800871. https://doi.org/10.3389/feduc.2026.1800871

Jiang, C., Laknapichonchat, T., Rakthin, C., & Suryanti, S. (2026). The impact of AI on primary school teachers' TPACK transformation challenges in the Chinese education context. International Journal of Learning, Teaching and Educational Research, 25(3), 100-115. https://doi.org/10.26803/ijlter.25.3.5

Jung, J., Choi, S., & Fanguy, M. (2024). Exploring teachers' digital literacy experiences. International Review of Research in Open and Distributed Learning, 25(2), 41-59. https://doi.org/10.19173/irrodl.v25i2.7572

Kayınova, B., & Adıgüzel, T. (2026). Primary school teachers' readiness for generation Alpha learning environments. Contemporary Educational Technology, 18(2), Article ep643. https://doi.org/10.30935/cedtech/18063

Kim, K., & Kwon, K. (2023). Exploring the AI competencies of elementary school teachers in South Korea. Computers and Education: Artificial Intelligence, 4, 100137. https://doi.org/10.1016/j.caeai.2023.100137

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

Kokoç, M., Ağırman, Ş., & Yağmurlugil, B. N. (2026). Mapping in-service teacher AI literacy: A systematic review of empirical studies. Teaching and Teacher Education, 181, 105707. https://doi.org/10.1016/j.tate.2026.105707

Lee, H. (2026). Teacher intervention in K-12 AI-based instruction: A systematic review of processes, strategies, and effects. Smart Learning Environments, 13, 35. https://doi.org/10.1186/s40561-026-00461-1

Li, M., Vale, C., Tan, H., & Blannin, J. (2025). Factors influencing the use of digital technologies in primary mathematics teaching: Voices from Chinese educators. Education and Information Technologies, 30, 12573-12608. https://doi.org/10.1007/s10639-024-13309-3

Liu, Y., & Zhang, Y. (2025). Revisiting techno-optimism: Multidimensional inequalities in AI education across urban, suburban, and rural primary schools in China. Education as Change, 29, Article 18869. https://doi.org/10.25159/1947-9417/18869

Mensah, A., Vo, T., & Ko, U. H. (2026). From planning to practice: Technology integration knowledge and enacted practice in elementary and middle school science. Education Sciences, 16(6), 958. https://doi.org/10.3390/educsci16060958

Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017-1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Musgrove, A., Powers, J., Nichols, B. H., & Lapp, S. (2021). Exploring the role of elementary teachers' TPACK in the adoption of 1:1 computing across subject areas. International Journal of Technology in Teaching and Learning, 17(1), 1-17. https://doi.org/10.37120/ijttl.2021.17.1.01

Muweesi, C., Shizhou, L., Nakonde, J., Jerome Kotira, S., Tomusange, R., & Sserwadda, L. (2021). Discourses in ICT integration: Pedagogical orientations in selected city primary schools in Uganda. Educational Research and Reviews, 16(5), 172-180. https://doi.org/10.5897/ERR2020.4103

Niu, H., & Niranjan, M. R. R. (2025). Teacher digital competence in New Zealand primary schools: A qualitative study of integration practices, challenges, and cultural considerations. The Australian Educational Researcher, 52, 4611-4632. https://doi.org/10.1007/s13384-025-00912-7

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.

African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 39 Décembre 2026

M., Hróbjartsson, 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. https://doi.org/10.1136/bmj.n71

Patiño Hernández, J. F., Chiappe, A., & Ortega González, E. (2026). From digital competence to AI-responsive pedagogy: A scoping review of technology-supported teacher education. Frontiers in Education, 11, 1949296. https://doi.org/10.3389/feduc.2026.1949296

Pflaumer, N., Knorr, N., & Berkling, K. (2021). Appropriation of adaptive literacy games into the German elementary school classroom. British Journal of Educational Technology, 52(5), 1917-1934. https://doi.org/10.1111/bjet.13149

Rahayu, A. H., Widodo, A., Saud, U. S., & Muslim. (2024). Improving technological pedagogical content knowledge (TPACK) of elementary school teachers through training with a collaborative, practical and reflective approach. Mimbar Sekolah Dasar, 11(4), 642–658. https://doi.org/10.53400/mimbar-sd.v11i4.78644

Rana, K., Greenwood, J., & Henderson, R. (2022). Teachers' experiences of ICT training in Nepal: How teachers in rural primary schools learn and make progress in their ability to use ICT in classrooms. Technology, Pedagogy and Education, 31(3), 275-291. https://doi.org/10.1080/1475939X.2021.2014947

Redecker, C., & Punie, Y. (2017). European framework for the digital competence of educators: DigCompEdu. Publications Office of the European Union. https://doi.org/10.2760/159770

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA Statement for reporting literature searches in systematic reviews. Systematic Reviews, 10, 39. https://doi.org/10.1186/s13643-020-01542-z

Rhodes, K., & Calder, N. (2022). Teachers implementing new primary school digital technology areas: What are we teaching now? Waikato Journal of Education, 27(3), 159-173. https://doi.org/10.15663/wje.v27i3.960

Sarı, M. H., & Keser, H. (2021). Classroom teachers' online teaching experiences during the COVID-19 pandemic: The perspective of technological pedagogical content knowledge. Journal of Pedagogical Research, 5(4), 251-269. https://doi.org/10.33902/JPR.2021474706

Shiu, W. H. C. (2024). Teacher learning in Hong Kong: Primary English language teachers' cognitions about online teaching to survive or thrive during COVID-19. Teachers and Teaching, 30(7-8), 1097–1125. https://doi.org/10.1080/13540602.2024.2328017

Sothayapetch, P., & Lavonen, J. (2022). Technological pedagogical content knowledge of primary school science teachers during the COVID-19 in Thailand and Finland. EURASIA

Journal of Mathematics, Science and Technology Education, 18(7), em2124. https://doi.org/10.29333/ejmste/12118

Suryanti, H. H. S., Handini, O., & Sutoyo. (2026). Teacher needs analysis in model development TPACK-based problem learning to improve critical thinking skills. International Journal of Research in Education and Science, 12, 497–511. https://doi.org/10.46328/ijres.5469

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8, 45. https://doi.org/10.1186/1471-2288-8-45

Walan, S., & Vinerean-Bernhoff, M. (2026). Generative AI-enhanced STEM-education: Investigating teachers' experiences from implementation in primary school settings. International Journal of Science and Mathematics Education, 24, Article 123. https://doi.org/10.1007/s10763-026-10747-y

Wilaphan, K., & Rattanawongsa, R. (2026). Development of a professional coaching model to enhance digital technology competency for primary teachers: A mixed-methods study. Suranaree Journal of Social Science, 20(1), e282027. https://doi.org/10.55766/sjss282027

Xiu, R. L., Aguilar, S. J., Macías, A. J., Xing, Y., Al-Sulaiti, R., Junjunia, M., & Jebril, S. T. (2026). An exploratory cross-national study of K–12 teachers' generative AI literacy and classroom enactment. Education Sciences, 16(5), 811. https://doi.org/10.3390/educsci16050811

Zhou, X., Padron, Y., & Waxman, H. (2022). Exploring the relationship between professional development experience and skills in educational technology integration among primary EFL teacher. Contemporary Educational Technology, 14(1), ep328. https://doi.org/10.30935/cedtech/11365