



# Blended learning in pre-service teacher education: A systematic review of design features, student engagement, and learning outcomes

Ruilian Shi<sup>1</sup>, Arumugam Raman<sup>2</sup> and Mohan Rathakrishnan<sup>3</sup>

<sup>1</sup>Baotou Teachers' College of Inner Mongolia University of Science and Technology, Inner Mongolia, China & School of Education, Universiti Utara Malaysia, Kedah, Malaysia (ORCID: 0009-0009-5197-7237)

<sup>2</sup>School of Education, Universiti Utara Malaysia, Kedah, Malaysia (ORCID: 0000-0001-5351-8944)

<sup>3</sup>School of Education, Universiti Utara Malaysia, Kedah, Malaysia (ORCID: 0000-0002-3360-5343)

Corresponding Author: Arumugam Raman, [arumugam@uum.edu.my](mailto:arumugam@uum.edu.my)

Submitted: 4 August 2026 Revised: 2 October 2026 Accepted: 4 October 2026 Published online: 24 October 2026

## Abstract

Blended learning has become increasingly prominent in pre-service teacher education, yet evidence remains fragmented regarding how particular instructional design features relate to student engagement and learning outcomes. This systematic review synthesised empirical research on blended learning in pre-service teacher education, with particular attention to instructional design, engagement, and academic and professional outcomes. We identified studies through systematic database searching and screened them using predefined eligibility criteria. We coded eligible studies for methodological characteristics, instructional design features, forms of engagement evidence, outcome domains, and direction of findings, and appraised methodological quality using the Mixed Methods Appraisal Tool. The synthesis identified six recurring and overlapping design themes, with alignment and sequencing, interaction and feedback, and autonomy and self-regulation particularly prominent. Engagement was represented inconsistently across the literature, ranging from direct measurement and proxy indicators to instances in which study authors invoked engagement without corresponding empirical indicators. Reported outcomes extended beyond academic achievement to include motivational, self-regulatory, technological, pedagogical, and professional dimensions. Although favourable findings were common, variation in research designs, measures, contexts, and methodological quality limits broad causal conclusions about blended learning effectiveness. The review indicates that blended learning in pre-service teacher education is better understood through the configuration and integration of its instructional components than as a uniform delivery model. Future research should more directly examine the pathways connecting instructional design, engagement, and academic and professional outcomes.

**Keywords:** Blended learning; Instructional design; Learning outcomes; Pre-service teacher education; Student engagement; Systematic review

## 1. Introduction

Blended learning (BL) refers to the purposeful integration of face-to-face instruction and computer-mediated learning activities within a single course or programme. In higher education, BL is widely adopted to increase flexibility and access while preserving interaction, feedback and guided practice. Its value, however, lies less in combining modalities than in design alignment: coherent sequencing across learning phases and explicit links among tasks, assessment evidence, feedback and scaffolding (Ashraf et al., 2021; Boelens et al., 2017). Accordingly, BL is better understood as a course-level pedagogical design than as a descriptive label for combining online and face-to-face delivery modes.

In pre-service teacher education (PSTE), BL is consequential because teacher candidates must build pedagogical content knowledge while learning to enact technology-enhanced teaching in authentic contexts. Its effectiveness should therefore be evaluated not only in terms of academic achievement but also through professionally meaningful outcomes, including teaching self-efficacy, reflective practice, digital pedagogical competence and readiness for technology-integrated teaching (Biberman-Shalev et al., 2023; Mishra & Koehler, 2006).

In the post-COVID-19 period, BL adoption accelerated as programmes redesigned coursework and practicum learning to sustain mentoring, collaboration and feedback across settings (Carrillo & Assuncao Flores, 2023). Although meta-analyses report overall benefits, heterogeneity in interventions, comparators and measures limits design-specific guidance (Ma et al., 2023; Schmid

et al., 2023). A synthesis that distinguishes enacted design features from broad modality labels is therefore needed to examine which aspects of BL design are reported alongside student engagement and educational outcomes in PSTE.

Student engagement is frequently proposed as a potential pathway connecting BL design with learning outcomes. Engagement is generally conceptualised as comprising behavioural, emotional and cognitive dimensions, but its operationalisation varies considerably across studies. Some studies measure engagement directly through validated scales, observations or explicitly defined behavioural indicators, whereas others use participation, task completion, platform activity or reflective outputs as proxy indicators. In other cases, engagement is invoked in interpreting reported learning outcomes without independent measurement. The Community of Inquiry framework can help explain interaction and teaching presence, while the TPACK framework addresses the integration of technological, pedagogical and content knowledge; however, neither framework substitutes for the direct measurement of student engagement (Bergdahl et al., 2024; Buntins et al., 2021; Garrison et al., 2000; Mishra & Koehler, 2006). Consequently, engagement is often proposed rather than empirically tested as a pathway connecting BL design with learning and professional outcomes.

Existing syntheses illuminate different aspects of BL and technology-enhanced teacher education but leave several questions unresolved. Reviews have documented definitional inconsistency, rapid growth in blended and flipped learning and generally positive outcome patterns. However, they rarely connect recurring design features with comparable indicators of student engagement and professionally meaningful outcomes in PSTE (Ashraf et al., 2021; Han & Røkenes, 2020; Short et al., 2021). Meta-analyses have estimated overall effects, but high heterogeneity and the limited number of experimental studies restrict conclusions about the pathways through which particular design features may influence engagement and outcomes (Ma et al., 2023; Schmid et al., 2023). Table 1 summarises these gaps.

Table 1  
*Gaps identified in prior syntheses of blended learning in PSTE*

| Prior synthesis        | Type & scope                                                                            | Included evidence (n) | Takeaway for the present review                                                                                                               |
|------------------------|-----------------------------------------------------------------------------------------|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| Ashraf et al. (2021)   | Review of reviews examining cross-disciplinary BL                                       | 57 reviews            | Identifies definitional and reporting inconsistencies but provides limited evidence linking PSTE design, engagement and outcomes.             |
| Short et al. (2021)    | Systematic mapping review of preparation for K-12 blended teaching                      | 88 primary studies    | Documents rapid growth but provides limited consolidation of recurring design features or comparable engagement and outcome indicators.       |
| Han and Røkenes (2020) | Scoping review of flipped classrooms in teacher education                               | 33 primary studies    | Reports generally positive outcomes, but engagement is heterogeneous and frequently inferred rather than directly measured.                   |
| Eggers et al. (2021)   | Systematic review of self-regulation strategy interventions in higher-education BL      | 21 primary studies    | Reports promising findings, but its broader higher-education scope and heterogeneous interventions limit PSTE-specific guidance.              |
| Ma et al. (2023)       | Meta-analysis of BL effectiveness in PSTE                                               | 15 primary studies    | Estimates overall effects, but limited experimental evidence and diverse outcomes restrict design-specific and mechanism-related conclusions. |
| Schmid et al. (2023)   | Meta-analysis comparing online, blended, flipped and classroom instruction for teachers | 107 primary studies   | Reports broad effects, but substantial heterogeneity limits PSTE-specific conclusions about design and underlying mechanisms.                 |

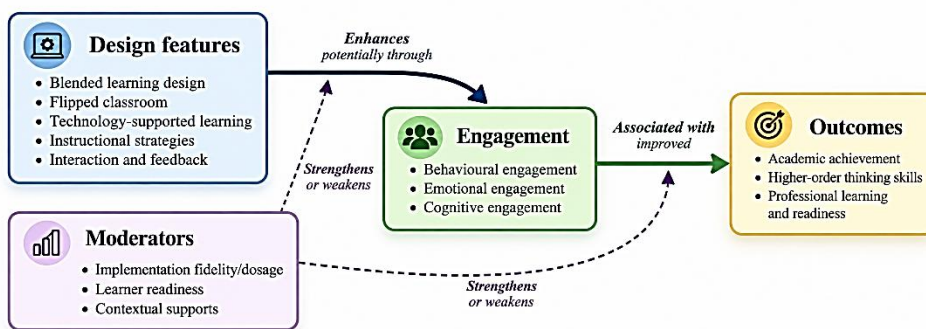
Note. BL = blended learning; PSTE = pre-service teacher education.

The literature presents three main gaps. First, broad modality classifications provide limited information about the specific design features enacted within courses and modules. Second, the strength of evidence supporting the proposed role of student engagement has not been systematically differentiated according to whether engagement was directly measured, represented by proxy indicators, or invoked without corresponding indicators. Third, previous syntheses have focused primarily on academic achievement and have paid less attention to professional outcomes, which are particularly important in teacher education.

To address these gaps, the present review links recurring BL design features to the strength of engagement evidence and to reported academic and professional outcomes across quantitative, qualitative, and mixed-methods studies. Figure 1 presents the analytical logic guiding the synthesis.

Figure 1

*Conceptual representation of potential pathways linking blended-learning design features, student engagement, and learning outcomes*



*Note.* Arrows represent hypothesised relationships derived from the synthesis and should not be interpreted as causal or mediation effects established across the included studies.

The model conceptualises BL as a course-level pedagogical design that creates opportunities for student engagement and contributes to multidimensional outcomes. It also recognises contextual and implementation conditions that may influence the observed relationships among design, engagement and outcomes.

The review contributes to the literature in three ways. First, it consolidates heterogeneous BL implementations into recurring design features that can inform pedagogical practice in PSTE. Second, it evaluates the strength of engagement-related evidence by distinguishing direct measurement, proxy-based evidence and interpretive inference. Third, it extends the synthesis beyond academic achievement to include digital pedagogical competence, self-regulation, collaboration, teaching self-efficacy and professional readiness.

Following the PRISMA 2020 reporting guidelines (Page et al., 2021), the review synthesises 59 empirical studies published between 2021 and 2025. It addresses the following review questions:

RQ 1) What study characteristics and BL configurations are represented in empirical research on pre-service teacher education?

RQ 2) What recurring design features are reported in relation to student engagement and academic and professional outcomes?

RQ 3) How is student engagement conceptualised and operationalised, and how strong is the evidence supporting its proposed role in potential design-engagement-outcome pathways?

RQ 4) What outcome patterns and contextual or implementation conditions are reported across different BL designs?

## 2. Method

This systematic review followed PRISMA 2020 (Page et al., 2021), and the search was reported in accordance with PRISMA-S (Rethlefsen et al., 2021). Before screening began, the reviewers developed a protocol specifying the eligibility criteria, screening and extraction rules, quality

appraisal procedures, and design-engagement-outcomes synthesis framework. The protocol was not prospectively registered. This limitation restricts independent verification that all decisions were made before the results were known. To improve transparency, we report the complete search strategies, eligibility criteria, extraction procedures, coding rules, and study-level decisions in Supplementary Tables S1-S6.

To minimise definitional drift, we operationalised blended learning conservatively as the verifiable integration of online and face-to-face components within the same course or module design. We coded the strength of engagement evidence during synthesis, rather than for eligibility, to calibrate interpretations of potential design-engagement-outcome pathways. We coded learning analytics or log indicators as Proxy unless the original study explicitly defined and analysed them as engagement. Three reviewers independently screened records and extracted data, resolving discrepancies through discussion and consensus. Figure 2 presents the PRISMA flow diagram.

## 2.1. Search Strategy

We searched Web of Science Core Collection, Scopus, and ERIC for English-language empirical journal articles on blended learning in PSTE published between January 2021 and December 2025 (Gusenbauer & Haddaway, 2020). We selected this five-year window to examine recent course designs developed during and after the rapid expansion of technology-supported teacher education associated with the COVID-19 period. It therefore provides a focused account of recent evidence rather than a historical synthesis of all blended-learning research in PSTE. We conducted searches on 31 December 2025, and finalised inclusion decisions on 13 January 2026. Records were deduplicated in EndNote X9. We also examined the reference lists of included studies and key prior syntheses for additional eligible reports. Supplementary Table S1 provides the complete database-specific search strings, applied limits, and yields. We considered language, publication type, database, and date restrictions when interpreting the review's coverage.

## 2.2. Eligibility Criteria and Study Selection

We specified eligibility criteria a priori using PICOS (Amir-Behghadami & Janati, 2020; Higgins et al., 2019). We included studies that examined pre-service teachers, implemented blended learning with verifiable online and face-to-face integration within the same course or module, and reported academic and/or professional outcomes using empirical methods. Quantitative, qualitative and mixed-method studies were eligible. Comparator groups were not required but were recorded when available. We did not include engagement as a criterion because operationalisation was inconsistent. Instead, we coded engagement evidence as Measured, Proxy, or Inferred during synthesis to calibrate interpretations of potential pathways.

We excluded studies when data for pre-service teachers could not be separated from other populations, blended integration could not be verified, reports were non-empirical or out of scope, full texts were unavailable, or articles were not in English. Three reviewers independently screened titles and abstracts and assessed full texts against the eligibility criteria, resolving disagreements through discussion and consensus. Full-text exclusions were coded using a predefined taxonomy (Supplementary Table S2). As shown in Figure 2, we included 59 studies in the final synthesis.

## 2.3. Data Extraction, Quality Appraisal, and Synthesis

We piloted a standardised extraction form and codebook on five studies and refined ambiguous definitions before applying the form to the full sample. Three reviewers extracted study characteristics, design and comparator information, blended-learning features, engagement operationalisation, outcome domains, and effect direction in Excel, and cross-checked the entries. Cohen's kappa was 0.91 for the categorical codes included in the agreement assessment. Because variable-level agreement coefficients were not available, this value indicates overall coding consistency rather than evidence that every category achieved equivalent reliability.

We coded engagement evidence as Measured, Proxy, or Inferred. The Measured category required explicit engagement indicators analysed as engagement, such as validated

multidimensional scales, structured observations or rubric scores, or analytics explicitly defined and tested as engagement. The Proxy category captured structured learning process traces that were not defined or analysed as engagement, such as coded participation or reflective outputs, clicks, views, time-on-task, or access frequency. We used the Inferred category when no engagement indicators were reported, and engagement was invoked only to interpret outcomes. Ambiguous cases were coded conservatively as Proxy or Inferred unless engagement was explicitly defined and tested.

We coded effect direction as a descriptive mapping variable rather than as an estimate of intervention effectiveness. For comparative studies, coding prioritised between-group results. For studies without a comparator, within-group pretest-posttest change was recorded separately because it provides weaker evidence for attribution. When a report did not identify a primary outcome, the reviewers applied the prespecified hierarchy documented in the codebook, beginning with the outcome most closely aligned with the stated study aim. Abstract prominence alone was not treated as sufficient evidence of primacy. We classified effects as positive, null, negative, or mixed. Because these categories combine heterogeneous measures and designs, they indicate the direction of reported findings and should not be interpreted as pooled effects or causal estimates.

We appraised methodological quality using the Mixed Methods Appraisal Tool (MMAT), version 2018 (Hong et al., 2018). In accordance with MMAT guidance, we did not calculate an overall numerical score or exclude studies solely because of appraisal results. We instead used criterion-level judgements to qualify the interpretation of individual findings and pathway-related claims. Supplementary Table S5 reports these judgements, with 'Cannot tell' indicating insufficient reporting. We also developed a review-specific, theme-level confidence appraisal based on four dimensions: methodological limitations identified through MMAT, consistency across studies, directness of engagement measurement, and objectivity of outcome measurement. Supplementary Table S6 reports the decision rules, dimension-level judgements, final confidence classification, and rationale for each of the six design themes. These classifications are interpretive aids and should not be interpreted as a validated certainty-of-evidence assessment. Given the substantial methodological and outcome heterogeneity, meta-analysis was not appropriate. We therefore used thematic synthesis (Thomas & Harden, 2008), descriptive mapping, and cross-study evidence matrices. Supplementary Table S4 provides the evidence-chain mapping.

### 3. Results

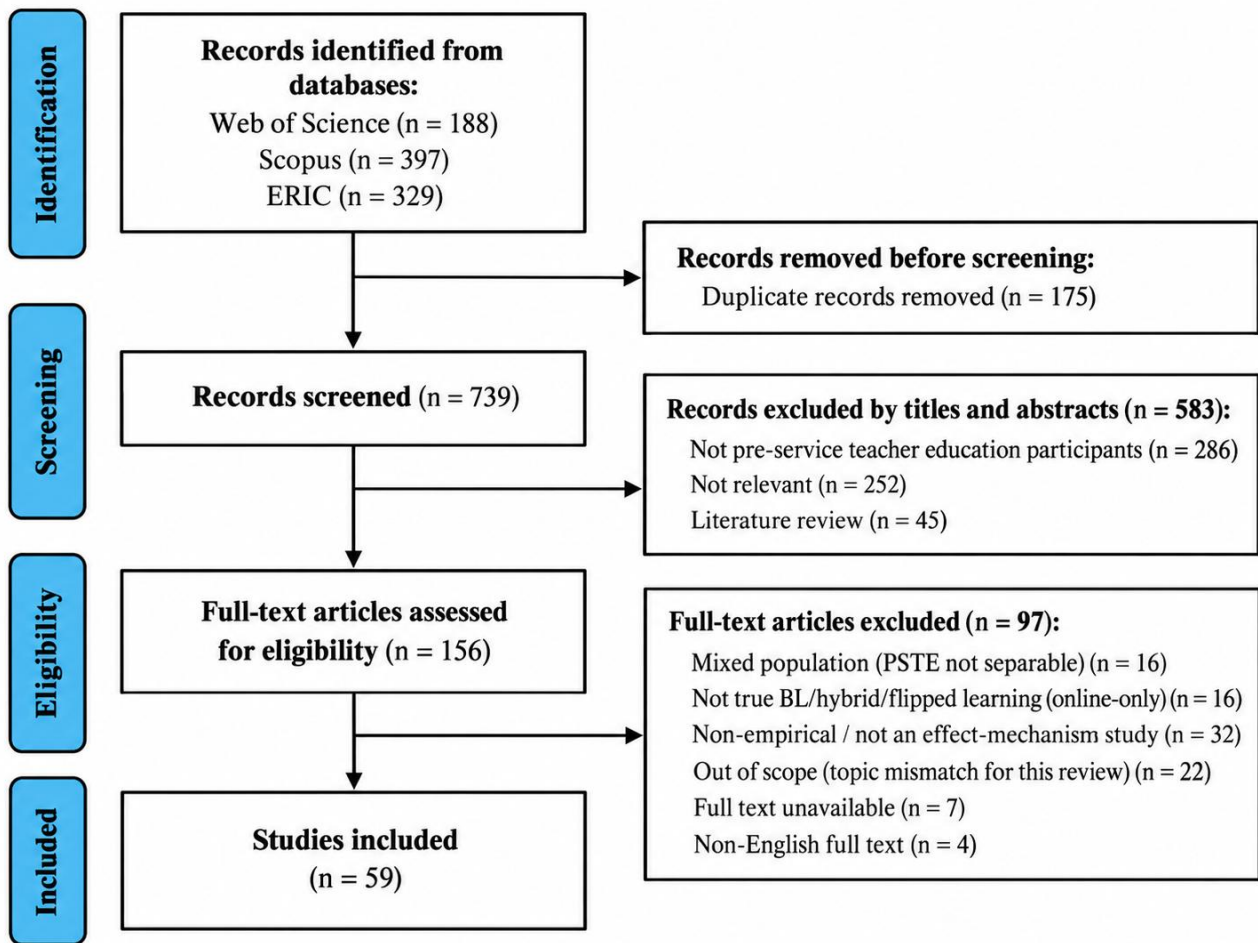
#### 3.1. Study Selection and Analytic Sample

Figure 2 presents the PRISMA flow diagram (Page et al., 2021). Searches in Web of Science, Scopus, and ERIC identified 914 records (188, 397, and 329, respectively). After removing 175 duplicates, we screened 739 titles and abstracts and excluded 583 records (population not relevant to PSTE,  $n = 286$ ; did not examine eligible blended learning,  $n = 252$ ; review articles,  $n = 45$ ). Of 156 full-text articles assessed, 97 were excluded (mixed or non-eligible populations,  $n = 16$ ; online-only designs,  $n = 16$ ; non-empirical or did not meet the outcome criteria,  $n = 32$ ; topic mismatch,  $n = 22$ ; full text not accessible,  $n = 7$ ; non-English articles,  $n = 4$ ). The final analytic sample included 59 empirical studies.

To minimise definitional drift, we operationalised BL conservatively: eligible studies had to demonstrate verifiable online and face-to-face integration within the same course or module and describe links between the two components across tasks and/or assessment. We excluded self-labelled "blended" studies without evidence of such integration. Given heterogeneity in enacted design logics, study designs and comparators, and the operationalisation of outcomes and engagement (Supplementary Table S3), we synthesised findings through a structured narrative synthesis that integrated study-level mapping (Supplementary Table S3) with theme-based evidence-chain aggregation (Section 3.3; Supplementary Table S4), rather than through meta-analysis.

Figure 2

PRISMA flow diagram of the screening and selection procedure for included studies



### 3.2. Descriptive Map of Included Studies

Supplementary Table S3 provides study profiles (study design, comparator, sample, blended learning configuration, engagement evidence type, outcome domains, and main findings), alongside effect-direction coding. The distributions reported below reflect the characteristics of the included studies rather than the representativeness of the wider population.

#### 3.2.1. Temporal trend

Publication output increased toward the end of the window (2021–2025), peaking in 2025 ( $n = 20$ ), indicating increased recent empirical reporting. Supplementary Table S3 (Sheet 2) reports study-level distributions and effect-direction coding.

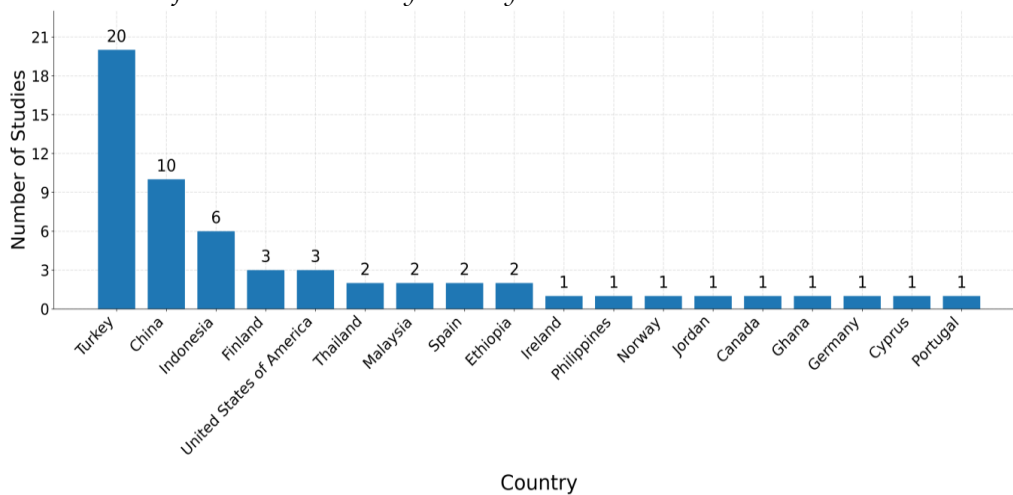
#### 3.2.2. Geographical distribution

Geographical coverage was broad but unevenly distributed. Turkey contributed 20 studies, followed by China and Indonesia, each with six studies. Taiwan contributed three studies and Hong Kong one study. Figure 3 summarises the geographical distribution, with full study-level coding provided in Supplementary Table S3.

#### 3.2.3. Study design types

Study designs were heterogeneous, with intervention-oriented comparisons dominating. Quasi-experimental or experimental designs with comparison groups accounted for 30 studies. One-group pre-post designs were also common (11 studies), which constrain causal attribution and increase reliance on implementation and comparator transparency. Survey, correlational, or

**Figure 3**  
*Distribution of included studies by country*



*Note.* Turkey aggregates studies labelled as Turkey or Türkiye in the source reports; China, Taiwan, and Hong Kong are reported separately according to the study settings recorded in Supplementary Table S3.

structural equation modelling (SEM) designs appeared in five studies. Qualitative or case-based designs appeared in seven studies, and design-based or action research in four. Learning analytics-based observational studies were rare (two studies). Comparator specification and baseline clarity are recorded in Supplementary Table S3.

#### 3.2.4. Blended learning configuration

Configuration labels were unevenly reported. Flipped learning was most common (33 studies). Hybrid variants appeared in two studies; HyFlex and rotation or station models appeared in one study each; massive open online course (MOOC)-based designs appeared in three studies. Nineteen studies were coded as other or unspecified. Because subtype labels were reported inconsistently, the synthesis prioritised enacted design features (Section 3.3) and treated configuration as contextual metadata (Supplementary Table S3).

#### 3.2.5. Engagement evidence and interpretive invocation

Engagement was inconsistently operationalised across the 59 studies. Thirteen studies provided empirical indicators relevant to engagement. Five directly measured engagement using indicators explicitly defined and analysed as engagement, while eight reported proxy indicators such as participation, task completion, learning-process traces, or reflective activity without explicitly operationalising these indicators as engagement. A further seven studies reported no engagement indicators but invoked engagement in the original study authors' interpretation of their findings. We treated these cases as interpretive invocation rather than empirical engagement evidence. The remaining 39 studies contained neither engagement indicators nor substantive engagement interpretation. Behavioural, emotional, and cognitive dimensions were retained only where supported by the underlying engagement coding and were not aggregated across inference-only cases as though they represented measured engagement. Study-level engagement classifications are reported in Supplementary Table S3, while theme-level direct-engagement counts are reported in Supplementary Table S4.

### 3.3. Thematic Synthesis of Design Features, Engagement, and Outcomes

Outcome reporting was extensive but uneven across domains: attitudes and motivation were reported in 47 studies, knowledge and achievement in 37, higher-order thinking and application in 27, and teamwork and communication in 12. These outcome domains were not mutually exclusive. These frequencies indicate domain coverage rather than effect magnitude. We observed mixed findings in 11 studies (Supplementary Table S3, Sheet 2).

Interpretation of potential engagement pathways was calibrated according to the strength of engagement evidence described in Section 2.3. Claims were stated most strongly when engagement was operationalised and analysed with an interpretable comparison and more cautiously when supported only by proxy indicators. Where engagement was invoked without indicators, it was treated as explanatory framing rather than empirical evidence of an engagement pathway.

The six themes were developed through an iterative, code-based thematic synthesis (Thomas & Harden, 2008). Reviewers first coded reported design features, engagement indicators, outcomes, and implementation conditions at study level. They then grouped related descriptive codes into candidate themes, compared them against the extracted evidence, and refined them to distinguish core design patterns from contextual conditions. The authors recorded the final definitions and inclusion rules in the codebook and applied them across the 59 studies. Because a study could contain several design features, assignments were non-mutually exclusive. Theme 2, interaction and feedback cycles, appeared in 54 studies and is therefore treated as a cross-cutting design dimension rather than a discrete configuration directly comparable with the other themes. Supplementary Table S4 provides the coding rules, study assignments, discrepant-case decisions, and examples supporting each theme. Table 2 summarises the theme-level links with direct engagement measurement, outcome-domain coverage, and mixed findings.

Across themes, reporting completeness varied, particularly for sequencing, facilitation routines, assessment criteria, and engagement operationalisation. These reporting gaps limited how explicitly we could map design features, engagement evidence, and outcomes.

### 3.3.1. *Coherent alignment and sequenced integration as the design backbone*

This theme captures blended designs that link preparation, enactment, and consolidation through explicit sequencing and alignment between tasks and assessment across online and face-to-face components (e.g., structured phases, aligned materials, and planned consolidation through reflection and feedback uptake) (Eren & Dökme, 2022; Reffiane et al., 2022).

Engagement evidence was limited within this theme ( $n = 11$ ). Engagement was represented mainly through proxy indicators or interpretive invocation, with no direct engagement measurement, which limits claims about potential engagement pathways (Hayward et al., 2022; Seoane & Jiménez, 2022). Outcomes most often involved knowledge or achievement and attitudes or motivation, with some higher-order application gains reported alongside aligned tasks (Bekele et al., 2025; Hayward et al., 2022). Studies also reported mixed findings when sequencing, dosage, fidelity, or comparator logic was under-specified, reducing interpretability (Kaya & Aydemir, 2025; Yeh et al., 2025).

### 3.3.2. *Interaction quality and feedback uptake as potential engagement pathways*

This theme identifies interaction quality and feedback uptake as potential pathways through which blended-learning designs may support engagement. Across the included studies, peer and instructor exchanges were organised around learning artefacts, feedback-informed revision, and connections between online contributions and in-class sense-making or task refinement (Kesler et al., 2024; Tumelius & Kuure, 2024). These recurring design features were reported alongside engagement indicators and learning outcomes, but the available evidence does not establish interaction or feedback as a causal mechanism.

Where engagement was directly operationalised, it was most consistently reflected in behavioural engagement indicators (e.g., participation quality, task commitment) and was less consistently reflected in emotional indicators (e.g., perceived support or value) (Kesler et al., 2024; Wang et al., 2025). Outcomes commonly involve attitudes, motivation, and achievement, with higher-order application gains most often reported when feedback cycles require revision, justification, and uptake. Mixed findings occur when interaction is shallow, facilitation is uneven, or comparators overlap, weakening counterfactual inference (Cabı & Türkoğlu, 2025; Sun et al., 2022).

Table 2  
Evidence chain map from design themes to engagement and outcome domains (non-mutually exclusive study-theme assignments)

| Design theme                                | Studies,<br>n | Engagement,<br>n | Knowledge/<br>achievement,<br>n (%) | HOT/<br>application,<br>n (%) | Teamwork/<br>communication, n<br>(%) | Attitudes/<br>motivation,<br>n (%) | Mixed<br>outcomes,<br>n (%) |
|---------------------------------------------|---------------|------------------|-------------------------------------|-------------------------------|--------------------------------------|------------------------------------|-----------------------------|
| Theme 1. Alignment and sequencing           | 11            | 0                | 8 (72.7)                            | 5 (45.5)                      | 3 (27.3)                             | 6 (54.5)                           | 2 (18.2)                    |
| Theme 2. Interaction and feedback cycles    | 54            | 5                | 35 (64.8)                           | 27 (50.0)                     | 12 (22.2)                            | 43 (79.6)                          | 9 (16.7)                    |
| Theme 3. Flipped classroom                  | 35            | 3                | 23 (65.7)                           | 18 (51.4)                     | 6 (17.1)                             | 30 (85.7)                          | 7 (20.0)                    |
| Theme 4. Autonomy and self-regulation       | 7             | 2                | 3 (42.9)                            | 0 (0.0)                       | 1 (14.3)                             | 7 (100.0)                          | 3 (42.9)                    |
| Theme 5. Practice and project-centred tasks | 19            | 1                | 10 (52.6)                           | 14 (73.7)                     | 7 (36.8)                             | 16 (84.2)                          | 1 (5.3)                     |
| Theme 6. Contextual and learner differences | 11            | 3                | 3 (27.3)                            | 3 (27.3)                      | 2 (18.2)                             | 11 (100.0)                         | 2 (18.2)                    |

Note. n represents the number of study-theme assignments. Percentages are calculated using the number of studies assigned to each theme as the denominator. HOT = higher-order thinking. Design themes and outcome domains were not mutually exclusive; a study could contribute to multiple themes and outcome domains. Accordingly, theme rows should not be summed or treated as independent groups, and percentages indicate descriptive outcome-domain coverage rather than comparative effect estimates. Direct engagement measurement includes only studies in which engagement was explicitly operationalised and analysed. Mixed outcomes refer to studies classified as Mixed according to the effect-direction rules described in Section 2.3. Although retained as a numbered theme, Theme 2 is interpreted as a cross-cutting design dimension because interaction and feedback features appeared across most included studies.

### 3.3.3. Flipped enactment patterns as a prominent implementation configuration

This theme identifies flipped enactment as the dominant blended configuration, aligning structured pre-class online preparation with active, collaborative, or problem-focused face-to-face sessions. Because the theme was based on enacted design features rather than reported configuration labels, its 35 study-theme assignments exceeded the 33 studies explicitly labelled as flipped learning. Practices include explicit links between preparation and in-class performance, accountability for pre-class work, and use of class time for analysis, practice, and feedback rather than content replay (Chen, 2024; Erol & Celik, 2022; Fidan, 2023; Günbatar, 2021).

Where engagement was operationalised, evidence most often reflected engagement with tasks across the preparation and enactment sequence, particularly when pre-class tasks were consequential and supported by guidance and feedback (Gu et al., 2022; Wang & He, 2023). Outcomes commonly cluster in motivation and knowledge achievement, with higher-order application gains reported when in-class time is protected for explanation, justification, and coached practice building on preparation (Chen, 2024; Erol & Celik, 2022; Günbatar, 2021). Mixed patterns cluster where fidelity or dosage varies, baseline equivalence is unclear, or comparators overlap (Cabı & Türkoğlu, 2025; Yurtseven Avci et al., 2022).

### 3.3.4. Scaffolded autonomy and self-regulation as potential engagement pathways

This theme foregrounds learner control and flexibility, but benefits are conditional on whether autonomy is scaffolded rather than assumed (Kouhia & Härkki, 2023; Yudit et al., 2024).

In the two studies that directly measured engagement, evidence primarily reflected self-regulatory engagement, including behavioural persistence, emotional commitment, and cognitive effort regulation, especially when scaffolds supported planning and monitoring (Cabı & Türkoğlu, 2025; Wang et al., 2025). Outcomes primarily involve attitudes or motivation; knowledge or achievement gains are less consistent and appear contingent on the specificity of self-management supports (Kouhia & Härkki, 2023; Menggo & Darong, 2022). Mixed patterns occur when flexibility is offered without adequate scaffolding, especially where readiness for self-regulated learning varies, and guidance for planning, monitoring, and feedback uptake is limited (Cabı & Türkoğlu, 2025; Yudit et al., 2024).

### 3.3.5. Practice-based and project-centred designs for professional learning outcomes

This theme covers blended interventions organised around authentic teaching tasks and performance artefacts, where integration is enacted through iterative production, rehearsal, and refinement (Hujjatusnaini et al., 2022; Mualim et al., 2025).

Only one study in this theme measured engagement directly. In the remaining studies, engagement was represented through proxy indicators or inferred from patterns such as sustained involvement, revision-focused effort, and collaborative participation (Hujjatusnaini et al., 2022; Mualim et al., 2025). Outcomes often include higher-order application gains alongside motivational benefits, with teamwork or communication benefits emerging when collaboration is structurally required (Dolmaci & Acar, 2025; Hujjatusnaini et al., 2022; Mualim et al., 2025). Mixed patterns occur when fidelity, exposure conditions, or artefact specifications and assessment criteria are under-reported, limiting attribution across settings (Kitching, 2026).

### 3.3.6. Contextual and learner differences as boundary conditions

This theme treats institutional context, delivery constraints, and learner characteristics as boundary conditions that may help explain heterogeneous outcomes, rather than assuming uniform transfer of design features (Kesler et al., 2024; Ramos et al., 2025).

Direct engagement measurement was limited to three studies in this theme, which linked contextual supports and learner characteristics to engagement and motivational outcomes (Han et al., 2025; Ramos et al., 2025; Wang et al., 2025). Attitudes and motivation were reported across all 11 studies, whereas knowledge and achievement, higher-order thinking and application, and teamwork and communication outcomes appeared in smaller subsets (Günbatar, 2021; Kesler et al., 2024; Tumelius & Kuure, 2024). Mixed patterns clustered where baseline conditions and

comparator logic were unclear, or where constraints reshaped engagement opportunities without sufficient implementation detail (Han et al., 2025; Jong, 2023).

## 4. Discussion

This review synthesised 59 empirical studies of blended learning in PSTE published from 2021 to 2025. Studies reported more favourable outcomes when they described purposeful sequencing, practice-oriented tasks, and feedback linked to revision. These patterns are descriptive rather than causal because study designs, comparators, measures, and implementation reporting varied considerably. Only five studies directly measured engagement. Consequently, the evidence supports plausible design propositions, but it does not establish engagement as a causal pathway linking blended designs with academic or professional outcomes.

### 4.1. Design as the Foundation for Effective Blended Learning

The synthesis identifies instructional design as a central factor differentiating reported outcome patterns in teacher preparation. More consistent positive outcomes were reported when online and face-to-face components were integrated through an explicit sequence, tied to assessment evidence, and assigned clear learning functions that supported rehearsal and reflection. This interpretation is consistent with research identifying interaction, support for learning processes, and the organization of instructional activities as central challenges in blended learning design (Boelens et al., 2017). Qualitative evidence from flipped learning implementations similarly highlights planning, material preparation, instructional processes, and assessment as important dimensions of implementation (Şahin et al., 2020). Moreover, positive outcomes do not necessarily extend across all learning dimensions: Yıldırım and Turgut (2025) reported improved academic achievement among pre-service teachers following a flipped classroom intervention but found no significant effects on self-regulated learning or classroom engagement. In a related teacher-education context, pre-service mathematics teachers participating in an AI-supported practicum described opportunities to develop reflective pedagogical decision-making, respond to diverse learner needs, and reconsider classroom interactions (Danişman, 2026). In this sense, blended learning is best interpreted as a pedagogical opportunity created through deliberate design rather than an inherent advantage associated with the “blended” label.

Teaching presence and scaffolding appear to be important design requirements. Within the Community of Inquiry framework, teaching presence encompasses instructional design, facilitation, and the direction of learning (Garrison & Arbaugh, 2007). Evidence from pre-service teacher education further associates teaching presence with instructor support, relevant instruction, and opportunities for active and collaborative learning (Özüdoğru, 2021). These principles can be operationalized through facilitation routines, planned feedback cycles, and clear expectations for participation and responses. Implementation challenges, including increased workload, uneven digital pedagogical capability, and inadequate technological support, can undermine instructional coherence (Cheng et al., 2026; Rasheed et al., 2020). Related evidence indicates that targeted training can strengthen pre-service teachers' self-efficacy in developing technology-supported instructional content, while structured practice opportunities, modelling, and feedback have been identified as important considerations in blended teacher preparation (Mohafa et al., 2026). To sustain BL beyond pilot delivery, programmes therefore need feasible workload models, reusable facilitation scripts, and reliable institutional infrastructure. Accordingly, mixed or null findings should be interpreted alongside reported instructional design, implementation fidelity, and comparator conditions rather than treated as definitive evidence against BL.

### 4.2. Engagement as a Potential Pathway, with Cautious Interpretation of Evidence Strength

This review treats engagement as a potentially testable pathway rather than an established explanatory mechanism, making stronger claims only where studies explicitly operationalised and measured engagement. Engagement is commonly conceptualised in behavioural, emotional, and

cognitive dimensions (Fredricks et al., 2004), a distinction that is also relevant to blended learning environments. Within the reviewed studies, behavioural engagement was suggested by participation in required preparatory activities and task completion, whereas opportunities for cognitive engagement appeared in activities involving explanation, reflection, and collaborative problem-solving. However, the presence of such activities should not automatically be equated with measured engagement. Emotional engagement was less consistently represented, and references to perceived support or task value did not necessarily constitute direct measures of this dimension. Evidence from blended learning research further illustrates the importance of distinguishing engagement dimensions, as cognitive, social, behavioural, and agency-related engagement can exhibit different patterns (Cubacub & Jimenez, 2025). Because direct engagement measurement was uncommon in the reviewed studies, these links are interpreted as plausible design propositions rather than established causal relationships.

Across most of the evidence base, engagement pathways should therefore be presented as hypotheses rather than tested explanations. Future pre-service teacher education studies of blended learning should employ validated multidimensional engagement measures and triangulate self-reports with observational evidence and theoretically grounded trace or log indicators, alongside relevant performance measures. This recommendation is consistent with evidence that learning analytics research predominantly captures observable behavioural engagement while less frequently addressing its multiple dimensions (Bergdahl et al., 2024). Related research with pre-service teachers has identified nonlinear associations between the frequency of AI-tool use and perceived mental workload, further illustrating why indicators of technology use should not be treated as interchangeable with measures of learners' cognitive experiences (Danışman & Evran Acar, 2026). The distinction between engagement and achievement is particularly important in teacher preparation: Yıldırım and Turgut (2025) reported improved academic achievement following flipped classroom instruction but no significant improvement in classroom engagement. Future studies should also report implementation fidelity and comparator equivalence, including contact time and assessment demands, to support stronger tests of proposed pathways, replication, and cumulative synthesis.

### **4.3. Learning and Professional Outcomes, Implications, and Limitations**

#### *4.3.1. Outcome patterns and conditionality*

Within these evidential constraints, studies frequently reported gains in knowledge and achievement, higher-order thinking, application, reflective practice, pedagogical confidence, and professional readiness. These frequencies describe the direction and coverage of reported findings rather than a pooled estimate of effectiveness. Professionally relevant outcomes were more often reported when authentic tasks, such as microteaching, portfolios, or coached enactment, were combined with structured feedback and revision. Nevertheless, alternative explanations, including selection effects, unequal contact time, comparator overlap, and implementation differences, cannot be excluded.

#### *4.3.2. Implications for practice, programme leadership, and research*

The implications extend across teaching practice, programme leadership, and research. Teacher educators should prioritise coherence over tools by aligning objectives, learning activities, and assessment evidence. Online components can support preparation, while face-to-face time can be protected for feedback-rich practice activities (Biggs, 1996; Garrison & Arbaugh, 2007). Programme leaders should support teaching presence through feasible workload models, stable infrastructure and staff development. They should standardise reusable assets, including facilitation scripts, feedback rubrics and assessment-linked task sequences, to maintain instructional quality beyond pilot implementation. Researchers should strengthen transparency by reporting implementation fidelity and comparator equivalence and by using replicable analytic decision rules. They should also operationalise engagement with validated multidimensional measures and triangulate self-

report measures with trace or log indicators explicitly modelled as engagement, alongside relevant performance-based indicators (Fredricks et al., 2004).

#### 4.3.3. Limitations and confidence in pathway interpretations

The principal limitation concerns the weakness of evidence for proposed engagement pathways. Engagement was directly measured in only five studies; eight used process indicators as proxies, seven invoked engagement without corresponding indicators, and 39 provided neither engagement indicators nor substantive engagement interpretation. Claims concerning a design-engagement-outcomes pathway must therefore remain provisional. Transferability is also limited by the evidence's geographical concentration and the predominance of single-institution studies. The restriction to English-language journal articles published between 2021 and 2025 and indexed in three databases may have excluded earlier research, non-English studies, grey literature, and studies indexed elsewhere, thereby introducing language, database, time-lag, and publication biases. The review should accordingly be understood as a synthesis of recent indexed evidence rather than a comprehensive historical account. Finally, although the protocol was prepared a priori, it was not prospectively registered, limiting external verification of prespecified decisions.

Comparability is also constrained by study design and reporting. Many studies relied on quasi-experiments or one-group pre-post designs and self-reported outcomes, with limited performance indicators or longitudinal follow-up. Heterogeneous blended configurations and incomplete fidelity reporting likely contribute to mixed findings. In particular, mixed or null effects are difficult to interpret when comparator conditions include substantial online elements, or when contact time and assessment demands differ across conditions. Reporting bias cannot be excluded; without meta-analysis, small-study effects cannot be formally assessed, so interpret the observed positive patterns cautiously.

#### 4.3.4. Priorities for future research and reporting

Future research should use designs matched to the questions being examined. Strong comparative or cluster-randomised designs may be appropriate when the aim is to estimate intervention effects and when allocation, sample size, and clustering can be handled rigorously. Design-based and longitudinal approaches may be more appropriate for examining how blended-learning components operate across authentic teacher-education settings. Across designs, researchers should prespecify primary outcomes, document comparator conditions and contact time, report implementation fidelity and null findings, and use validated measures of behavioural, emotional, and cognitive engagement. Self-reports should be triangulated with trace data and performance indicators that are explicitly modelled as engagement. Future reviews should register protocols prospectively and apply established certainty-assessment procedures where appropriate.

These improvements would permit clearer comparisons across settings and more defensible tests of whether engagement mediates or otherwise helps explain relationships between blended-learning design and student outcomes.

#### 4.3.5. Evidence-informed design principles

Drawing on the evidence map presented in Table 2, six evidence-informed design principles can be identified to guide the development and implementation of blended learning courses in pre-service teacher education (PSTE). First, blended learning activities should be purposefully sequenced, ensuring alignment between pre-class, in-class, and post-class activities, their intended learning functions, and assessment evidence. Second, interaction and feedback should be structured as iterative processes that enable student teachers to engage with feedback, revise their work, and progressively improve their understanding and performance, rather than merely receive feedback. Third, flipped classroom designs should prioritise the use of face-to-face instructional time for pedagogical enactment, including rehearsal, modelling, and guided practice. Fourth, learner autonomy should be supported through appropriate scaffolding, combining flexibility with clear instructional structures, pacing support, and prompts that facilitate self-regulated learning. Fifth, practice-centred learning activities should emphasise the production of

teaching artefacts, microteaching experiences, and reflective outputs that connect coursework with practicum expectations. Finally, the implementation of blended learning should be responsive to contextual conditions, including practicum arrangements, technological access, cohort characteristics, and institutional resources, with these conditions reported transparently to support the interpretation and transferability of findings. Given the limited direct evidence concerning engagement pathways, these principles should be understood as evidence-informed design propositions requiring further empirical examination rather than universally validated prescriptions.

## 5. Conclusion

This review indicates that blended learning in pre-service teacher education is best examined as a course-level pedagogical design rather than as the mere use of multiple delivery modes. Across 59 recent studies, studies that coherently sequenced online and face-to-face components, actively facilitated them, aligned them with assessment, and connected them to authentic teaching practice more often reported favourable academic and professional outcomes. However, heterogeneity in research designs and outcome measures, together with direct engagement measurement in only five studies, prevents causal conclusions about effectiveness or the proposed design-engagement-outcome pathways. The review's main contribution is therefore an evidence-informed framework that differentiates design features, the strength of engagement evidence, and reported outcomes. Future research should test these proposed pathways using transparent implementation reporting, validated multidimensional engagement measures, prespecified outcomes, and appropriately designed comparative or longitudinal studies.

**Artificial Intelligence (AI) Use Disclosure:** The authors declare that the generative artificial intelligence (AI) tool ChatGPT (OpenAI) was used solely to assist with language editing, wording refinement, and editorial compliance checking during manuscript revision. AI did not influence the scientific content, study design, data extraction, data analysis, interpretation of results, or conclusions. The authors critically reviewed and verified all AI-assisted revisions and take full responsibility for the manuscript's final content.

**Author contributions:** Ruilian Shi: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. Arumugam Raman: Methodology, Investigation, Data curation, Validation, Supervision, Writing – review & editing. Mohan Rathakrishnan: Investigation, Data curation, Validation, Supervision, Writing – review & editing.

**Data availability:** All data were derived from published articles. Supporting search, extraction, coding, and supplementary materials are provided in the Supplementary Materials.

**Declaration of interest:** The authors declare no conflict of interest.

**Ethical statement:** Ethical approval was not required for this study, as it was a systematic review based exclusively on previously published literature and did not involve the collection of primary data from human participants or animals.

**Funding:** This research received no external funding.

## References

- Amir-Behghadami, M., & Janati, A. (2020). Population, intervention, comparison, outcomes and study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal*, 37(6), 387. <https://doi.org/10.1136/emered-2020-209567>
- Ashraf, M. A., Yang, M., Zhang, Y., Denden, M., Tlili, A., Liu, J., Huang, R., & Burgos, D. (2021). A systematic review of systematic reviews on blended learning: Trends, gaps and future directions. *Psychology Research and Behavior Management*, 14, 1525–1541. <https://doi.org/10.2147/PRBM.S331741>

- Bekele, A., Melese, W., & Sime, T. (2025). The impact of a blended learning approach on students' academic achievement in Ethiopia: Jimma College of Teacher Education in focus. *Interactive Learning Environments*, 33(10), 5996–6006. <https://doi.org/10.1080/10494820.2025.2491617>
- Bergdahl, N., Bond, M., Sjöberg, J., Dougherty, M., & Oxley, E. (2024). Unpacking student engagement in higher education learning analytics: A systematic review. *International Journal of Educational Technology in Higher Education*, 21(1), 63. <https://doi.org/10.1186/s41239-024-00493-y>
- Biberman-Shalev, L., Broza, O., & Chamo, N. (2023). Contextual changes and shifts in pedagogical paradigms: Post-COVID-19 blended learning as a negotiation space in teacher education. *Education Sciences*, 13(3), 275. <https://doi.org/10.3390/educsci13030275>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- Boelens, R., De Wever, B., & Voet, M. (2017). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review*, 22, 1–18. <https://doi.org/10.1016/j.edurev.2017.06.001>
- Buntins, K., Kerres, M., & Heinemann, A. (2021). A scoping review of research instruments for measuring student engagement: In need for convergence. *International Journal of Educational Research Open*, 2, 100099. <https://doi.org/10.1016/j.ijedro.2021.100099>
- Cabı, E., & Türkoğlu, H. (2025). The impact of a learning analytics based feedback system on students' academic achievement and self-regulated learning in a flipped classroom. *International Review of Research in Open and Distributed Learning*, 26(1), 175–196. <https://doi.org/10.19173/irrodl.v26i1.7924>
- Carrillo, C., & Assuncao Flores, M. (2023). Online teaching and learning practices in teacher education after COVID-19: Lessons learnt from the literature. *Journal of Education for Teaching*, 49(5), 869–881. <https://doi.org/10.1080/02607476.2022.2153018>
- Chen, C. Y. (2024). Flipped classroom with case-based learning for improving preservice teachers' classroom management learning outcomes. *Teaching and Teacher Education*, 152, 104785. <https://doi.org/10.1016/j.tate.2024.104785>
- Cheng, Z., Yang, F., & Zhu, C. (2026). Understanding teacher workload in blended learning: Insights through the job demands-resources model. *Electronic Journal of e-Learning*, 24(1), 125–137. <https://doi.org/10.34190/ejel.24.1.4465>
- Cubacub, P. D. C., & Jimenez, E. C. (2025). The role of blended learning in enhancing student engagement: Evidence from high schools in Micronesia. *International Journal of Didactical Studies*, 6(3), Article 33074. <https://doi.org/10.33902/ijods.202533074>
- Danişman, Ş. (2026). Generative AI supported practicum experiences contribute to the pedagogical reasoning and professional identity development of preservice mathematics teachers. *E-Learning and Digital Media*. Advance online publication. <https://doi.org/10.1177/20427530261465688>
- Danişman, Ş., & Evran Acar, F. (2026). Artificial intelligence use and cognitive resource allocation: Nonlinear associations with mental workload and perceived performance. *Journal of Intelligence*, 14(8), Article 154. <https://doi.org/10.3390/jintelligence14080154>
- Dolmaci, A., & Acar, F. E. (2025). Development of a twenty-first century skills curriculum through the flipped learning approach. *Humanities and Social Sciences Communications*, 12(1), 1666. <https://doi.org/10.1057/s41599-025-05925-2>
- Eggers, J. H., Oostdam, R., & Voogt, J. (2021). Self-regulation strategies in blended learning environments in higher education: A systematic review. *Australasian Journal of Educational Technology*, 37(6), 175–192. <https://doi.org/10.14742/ajet.6453>
- Eren, E., & Dökme, I. (2022). An effective intervention with a blended learning environment for improving cognitive learning and spiritual meaning. *Cultural Studies of Science Education*, 17(4), 991–1012. <https://doi.org/10.1007/s11422-021-10083-3>
- Erol, T., & Celik, S. (2022). Flipped classrooms: Designed and implemented with colleagues collaboration. *Journal of University Teaching and Learning Practice*, 19(3), 13. <https://doi.org/10.53761/1.19.3.12>
- Fidan, M. (2023). The effects of microlearning-supported flipped classroom on pre-service teachers' learning performance, motivation and engagement. *Education and Information Technologies*, 28(10), 12687–12714. <https://doi.org/10.1007/s10639-023-11639-2>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <https://doi.org/10.3102/00346543074001059>
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2–3), 87–105. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)

- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3), 157–172. <https://doi.org/10.1016/j.iheduc.2007.04.001>
- Gu, J., Tang, L., Liu, X., & Xu, J. (2022). Promoting pre-service teacher students' learning engagement: Design-based research in a flipped classroom. *Frontiers in Psychology*, 13, 810275. <https://doi.org/10.3389/fpsyg.2022.810275>
- Günbatar, M. S. (2021). Evaluation of the problem based flipped classroom instruction process in the framework of Community of Inquiry. *Journal of Problem Based Learning in Higher Education*, 9(2), 46–68. <https://doi.org/10.5278/ojs.jpblhe.v9i2.6690>
- Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, 11(2), 181–217. <https://doi.org/10.1002/jrsm.1378>
- Han, H., & Røkenes, F. M. (2020). Flipped classroom in teacher education: A scoping review. *Frontiers in Education*, 5, 601593. <https://doi.org/10.3389/educ.2020.601593>
- Han, H., Rokenes, F., & Krumsvik, R. (2025). The flipped classroom's impact on students' motivation and engagement. *Nordic Journal of Digital Literacy*, 20(3), 158–171. <https://doi.org/10.18261/njdl.20.3.3>
- Hayward, D. V., Mousavi, A., Carbonaro, M., Montgomery, A. P., & Dunn, W. (2022). Exploring pre-service teachers engagement with live models of universal design for learning and blended learning course delivery. *Journal of Special Education Technology*, 37(1), 112–123. <https://doi.org/10.1177/0162643420973216>
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). Wiley. <https://doi.org/10.1002/9781119536604>
- 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>
- Hujjatusnaini, N., Corebima, A. D., Prawiro, S. R., & Gofur, A. (2022). The effect of blended project-based learning integrated with 21st-century skills on pre-service biology teachers' higher-order thinking skills. *Jurnal Pendidikan IPA Indonesia*, 11(1), 104–118. <https://doi.org/10.15294/jpii.v11i1.27148>
- Jong, M. S. Y. (2023). Flipped classroom: Motivational affordances of spherical video-based immersive virtual reality in support of pre-lecture individual learning in pre-service teacher education. *Journal of Computing in Higher Education*, 35(1), 144–165. <https://doi.org/10.1007/s12528-022-09334-1>
- Kaya, O. N., & Aydemir, S. (2025). A new learning environment for improving pre-service science teachers' views of nature of science: Triple blended learning environment. *Education and Information Technologies*, 30(6), 7387–7455. <https://doi.org/10.1007/s10639-024-13014-1>
- Kesler, M., Kaasinen, A., & Kervinen, A. (2024). Teaching science outdoors: Supporting pre-service teachers' skill development with the help of available mobile applications. *Education Sciences*, 14(11), 1218. <https://doi.org/10.3390/educsci14111218>
- Kitching, N. (2026). 'I'll do it tomorrow': Supporting the engagement of primary pre-service teachers through blended, inquiry-based learning. *Irish Educational Studies*, 45(1), 143–161. <https://doi.org/10.1080/03323315.2024.2394396>
- Kouhia, A., & Härkki, T. (2023). Advanced student motivation in craft teacher education: Learning new yarn crafts at university study courses. *FormAkademisk*, 16(4). <https://doi.org/10.7577/formakademisk.5384>
- Ma, K., Zhang, J., Chutiyami, M., Liang, L., & Dong, J. (2023). Effectiveness of blended teaching in pre-service teacher education: A meta-analysis. *Distance Education*, 44(3), 495–521. <https://doi.org/10.1080/01587919.2022.2155617>
- Menggo, S., & Darong, H. C. (2022). Blended learning in ESL/EFL class. *LLT Journal: Journal on Language and Language Teaching*, 25(1), 132–148. <https://doi.org/10.24071/llt.v25i1.4159>
- 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>
- Mohafa, L. G., Makhechane, M., & Qhobela, M. (2026). Pre-service physical science teachers' self-efficacy for sustainable digital pedagogy: Evidence from teaching practice after emergency blended learning. *Sustainable Education and Digital Transformation*, 1(2), e43696. <https://doi.org/10.33902/SEDIT.202643696>
- Mualim, M., Margana, M., Widiantoro, A., Nguyen, H. H., Khristianto, K., Faisal, F., & Roiyasa, N. (2025). Evaluating INSTALL as an instructional model to enhance EFL pre-service teachers' TPACK: A case of

- three Indonesian universities. *Frontiers in Education*, 10, 1580173. <https://doi.org/10.3389/feduc.2025.1580173>
- Özüdoğru, M. (2021). Pre-service teachers' perceptions related to the distance teacher education learning environment and community of inquiry. *Journal of Pedagogical Research*, 5(4), 43–61. <https://doi.org/10.33902/JPR.2021472945>
- 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. <https://doi.org/10.1136/bmj.n71>
- Ramos, A. M., Lee, H., & Mabuan, R. A. (2025). Exploring the relationship among pre-service teachers' e-learning readiness, learning engagement, and learning performance in HyFlex learning environments. *International Review of Research in Open and Distributed Learning*, 26(2), 89–110. <https://doi.org/10.19173/irrodl.v26i2.8165>
- Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A systematic review. *Computers & Education*, 144, 103701. <https://doi.org/10.1016/j.compedu.2019.103701>
- Reffiane, F., Sudarmin, Wiyanto, & Saptano, S. (2022). The effectiveness of hybrid learning model using Ethno-STEM approach to develop critical thinking skills (CrTS). *International Journal of Early Childhood Special Education*, 14(1), 1242–1250. <https://doi.org/10.9756/int-jecse/v14i1.221142>
- 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(1), 39. <https://doi.org/10.1186/s13643-020-01542-z>
- Schmid, R., Borokhovski, E., Bernard, R., Pickup, D., & Abrami, P. (2023). A meta-analysis of online learning, blended learning, the flipped classroom and classroom instruction for pre-service and in-service teachers. *Computers & Education Open*, 5, 100142. <https://doi.org/10.1016/j.caeo.2023.100142>
- Seoane, R. C., & Jiménez, J. E. (2022). Effectiveness of online and blended delivery methods on pre-service teachers' knowledge and beliefs for writing instruction. *Journal of Education for Teaching*, 48(2), 178–196. <https://doi.org/10.1080/02607476.2021.1963191>
- Short, C. R., Graham, C. R., Holmes, T., Oviatt, L., & Bateman, H. (2021). Preparing teachers to teach in K-12 blended environments: A systematic mapping review of research trends, impact, and themes. *TechTrends*, 65(6), 993–1009. <https://doi.org/10.1007/s11528-021-00626-4>
- Sun, F. R., Hu, H. Z., Wan, R. G., Fu, X., & Wu, S. J. (2022). A learning analytics approach to investigating pre-service teachers' change of concept of engagement in the flipped classroom. *Interactive Learning Environments*, 30(2), 376–392. <https://doi.org/10.1080/10494820.2019.1660996>
- Şahin, Ş., Boyacı, Z., Ökmen, B., Danişman, Ş., Eriş Hasırcı, H. M., & Kılıç, A. (2020). Determining visions related to the flipped learning. *Manisa Celal Bayar University Journal of Social Sciences*, 18(3), 62–84. <https://doi.org/10.18026/cbayarsos.600312>
- 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>
- Tumelius, R., & Kuure, L. (2024). Pre-service teachers' professional vision and agency emerging in orchestrating language learning in a hybrid space. *Computer Assisted Language Learning*, 37(7), 1554–1578. <https://doi.org/10.1080/09588221.2022.2092154>
- Wang, T., & He, H. (2023). Pre-service teachers' task engagement in flipped Chinese language and culture learning. *International Journal of Computer-Assisted Language Learning and Teaching*, 13(1), 1–17. <https://doi.org/10.4018/IJCALLT.323649>
- Wang, Z., Chen, X., Zhao, H., Xu, Y., & Wang, P. (2025). How does the online discussion intervention influence learners' engagement, emotions, and motivation in a blended learning environment? Learners' need for cognition matters. *BMC Psychology*, 13(1), 1122. <https://doi.org/10.1186/s40359-025-03466-6>
- Yeh, Y. C., Wang, C. M., Ting, Y. S., & Chiang, J. L. (2025). A visible thinking approach for pre-service teacher training: Effects and mechanisms for professional development in creativity instruction. *Teaching Education*, 36(3), 241–258. <https://doi.org/10.1080/10476210.2025.2485230>
- Yıldırım, M., & Turgut, Y. E. (2025). The effect of the flipped classroom model on pre-service teachers' academic achievement, self-regulated learning skills, and classroom engagement. *Journal of Pedagogical Research*, 9(3), 242–257. <https://doi.org/10.33902/JPR.202535543>

- 
- Yudt, K. E., Sawyer, B. E., & Shera, S. B. (2024). Preservice elementary teachers' mathematical achievement and attitudes: A study of blended learning. *Journal of Mathematics Teacher Education*, 27(3), 355–377. <https://doi.org/10.1007/s10857-022-09565-0>
- Yurtseven Avci, Z., Ergulec, F., Misirli, O., & Süral, I. (2022). Flipped learning in information technology courses: Benefits and challenges. *Journal of Further and Higher Education*, 46(5), 636–650. <https://doi.org/10.1080/0309877X.2021.1986623>