REVIEW ARTICLE

Journal of Pedagogical Sociology and Psychology
Volume 8, Issue 3, 2026, e49161
https://doi.org/10.33902/JPSP.202649161

Technology-enhanced calligraphy education: A systematic review and bibliometric analysis of research trends

Enqi Tang[1], Siti Nur Diyana Mahmud[2], Nur Atiqah Jalaludin[3] and Mohamad Sattar Rasul[4]

[1]Universiti Kebangsaan Malaysia, Malaysia & Kunming City College, China (ORCID: 0009-0004-1357-8968)
[2]Universiti Kebangsaan Malaysia, Malaysia (ORCID: 0000-0001-6398-1166)
[3]Universiti Kebangsaan Malaysia, Malaysia (ORCID: 0000-0003-4114-5030)
[4]Universiti Kebangsaan Malaysia, Malaysia (ORCID: 0000-0001-9428-470X)

Corresponding Author: Siti Nur Diyana Mahmud, diyana@ukm.edu.my

Submitted: 9 August 2026 Revised: 16 October 2026 Accepted: 1 November 2026 Published online: 15 November 2026

Abstract

Calligraphy is simultaneously an academic subject, a cultural practice inscribed on UNESCO's Representative List of the Intangible Cultural Heritage of Humanity, and a skill increasingly mediated by digital technology. Research on technology-enhanced calligraphy education has grown quickly, yet remains scattered across engineering, computing and education venues and has rarely been examined through a science, technology, engineering, arts and mathematics (STEAM) lens. This study systematically reviews international research published between 2015 and 2025. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 guideline, 753 records were retrieved from the Web of Science Core Collection, Scopus and the China National Knowledge Infrastructure; after deduplication and dual independent screening, 62 studies met the eligibility criteria. Bibliometric analysis, keyword co-occurrence mapping, appraisal with the Mixed Methods Appraisal Tool and thematic coding were applied. Annual output remained modest until 2019, expanded after 2020 and peaked in 2024, with 82.3% of studies carrying mainland-China or Taiwan affiliations—a concentration replicated across the international and Chinese-language databases. Five themes organize the field: AI-based assessment of calligraphy quality; online and web-based learning; intelligent tutoring, generative AI and pedagogy; immersive learning environments; and robotics and haptic skill training. System-development studies dominate, classroom-based effect evidence is scarce, methodological quality is mostly low to moderate, and only one study operationalizes an explicit STEAM framework. The field is technology-rich but pedagogy-poor. A STEAM integration framework is proposed in which calligraphy functions as the arts component, with science and mathematics made explicit in instructional design.

Keywords: Arts integration; Bibliometric analysis; Calligraphy education; Educational technology; STEAM education; Systematic review

1. Introduction

1.1. Calligraphy Education in the Twenty-first Century

Few school subjects carry the dual identity of calligraphy. It is at once a disciplined art form taught through years of guided practice and a living heritage whose transmission has become a matter of public policy. The inscription of Chinese calligraphy on UNESCO's Representative List of the Intangible Cultural Heritage of Humanity in 2009 formalized what educators across East Asia had long acted upon: handwriting arts embody cultural knowledge that cannot be reduced to language literacy alone (UNESCO, n.d.). Parallel traditions hold comparable status elsewhere, from Japanese shūji and Korean seoye to the Arabic and Persian calligraphic schools, each sustained by its own pedagogy of copying, correction and apprenticeship.

Policy has reinforced this standing. In mainland China, the Ministry of Education issued dedicated guidelines for calligraphy education in primary and secondary schools in 2013, and the 2022 national curriculum standard for arts education subsequently absorbed calligraphy into compulsory arts learning, complete with grade-banded expectations for brush and hard-pen practice (Ministry of Education of the People's Republic of China, 2013, 2022). Japan and Korea maintain calligraphy within their national curricula, and demand extends well beyond schooling: calligraphy courses are a staple of community education for older adults across East Asia, a constituency whose learning needs have grown more visible as societies age.

The pedagogy that carries this heritage is conservative by design. Learners begin with tracing and grid-guided copying, advance to reproducing model works stroke by stroke, and only

Copyright © 2026 by the author(s). This is an **open-access** article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).

gradually attempt free composition, a sequence that has changed little over centuries. Progress depends on repeated cycles of demonstration and correction in which the teacher sets the learner's work against an exemplar and articulates what is wrong with it: the angle of an entry stroke, the distribution of white space, the tempo of a turn. The knowledge transmitted in these exchanges is tacit as much as explicit, and it presupposes sustained expert attention. Assessment inherits the same character. Judging a piece of calligraphy is an act of trained connoisseurship that weighs proportion, balance, rhythm and vitality together; criteria exist, but they resist the standardization that schooling at scale requires, and they travel poorly across classrooms and examiners.

Yet teaching the subject well at scale remains difficult in many of the settings documented in this review. Studies of school and university calligraphy programmes in China and Japan repeatedly describe a shortage of qualified instructors outside major urban centres, corrective feedback that is difficult to deliver in one-to-many classrooms, and assessment that still rests largely on expert impression rather than shared criteria (Jiang et al., 2016; Ma et al., 2018; Mi, 2018). These frictions, rather than any lack of cultural commitment, define the practical problem space into which educational technology has entered.

1.2. From STEM to STEAM: Integrating the Arts

The movement from STEM to STEAM reframed the arts as a partner in interdisciplinary learning rather than a decorative supplement. Yakman and Lee (2012) articulated an influential framework in which the "A" signals not a fifth silo but a mode of connecting the other four, and Liao (2016) pushed the argument toward transdisciplinarity, where artistic practice becomes the integrative medium itself. A decade of synthesis work has tempered the enthusiasm. Perignat and Katz-Buonincontro's (2019) integrative review of forty-four STEAM studies found conceptual ambiguity, thin documentation of learning outcomes and a near absence of arts-specific achievement measures, while Mejías et al. (2021) catalogued the field's definitional troubles even as they defended its use. Empirical work continues to accumulate, from pedagogical models that raise project competence and motivation (Lin & Tsai, 2021) to design-based studies of cooperative STEAM teaching in primary schools (Li et al., 2022) and analyses of what sustains learners' intention in STEAM courses (Wu et al., 2022). Recent design-based STEM studies of pre-service teachers and secondary students likewise indicate that interdisciplinary learning becomes educationally meaningful only when connections among disciplinary content, design processes and intended outcomes are made explicit, rather than treated as a label applied after the activity (Elmali & Balkan Kiyici, 2025; Fikri et al., 2026). A quieter limitation concerns measurement: outcome research has relied heavily on motivation, engagement and perceived competence, while arts-specific skill gains are rarely assessed directly, so the framework's promise of genuine integration is asserted more often than it is demonstrated (Perignat & Katz-Buonincontro, 2019).

One pattern in this literature has received little comment: the arts that appear in STEAM research are overwhelmingly generic—design tasks, visual arts projects, maker activities—while heritage art forms with codified techniques and cultural obligations, such as calligraphy, rarely occupy the centre of a STEAM investigation. Whether the framework travels well to such forms is an open question, and one that matters most where heritage arts are compulsory school content.

1.3. Technology Meets Calligraphy Education

Four streams of technology have converged on calligraphy education over the past decade. The first applies machine perception to the subject's oldest bottleneck, feedback: convolutional, siamese and transformer architectures now score handwritten characters, detect departures from model styles and estimate the aesthetic quality of imitation work, promising the consistent, criterion-referenced judgement that expert-dependent classrooms cannot always supply (Sun et al., 2023; Yan et al., 2024; Zhang et al., 2017). A second stream recreates the studio rather than the scoreboard, embedding brush practice in virtual, augmented and mixed-reality environments where learners watch strokes executed around them, rehearse pressure control through haptic proxies and study model works at full scale (Loup-Escande et al., 2017; Luo et al., 2023; Wang &

Gao, 2024). Robotics constitutes a third stream. Machines that write with brush and ink began as engineering demonstrations and have been reframed as instruments of cultural transmission, with reinforcement-learning systems reproducing master-level stroke dynamics (Wu et al., 2018, 2020b) and haptic devices guiding the learner's hand directly (Fang et al., 2024; Raza et al., 2019). The fourth stream delivers instruction at distance. Online platforms that pair recognition models with personalized feedback have opened calligraphy to self-directed learners (Ch'ng et al., 2025; Wong et al., 2024), a development that coincided with the COVID-19 pandemic and may reflect the pandemic-era shift to remote instruction, and generative AI has most recently entered the construction of multimodal learning resources (Yu et al., 2025). Running across all four streams is a small set of pedagogically explicit designs, among them intelligent tutors for the Arabic and Persian traditions (Mamoun et al., 2015; Sadikoglu & Soroush, 2016) and a single study that names STEAM as its framework (Tang et al., 2025). Evidence from neighbouring arts and creative-learning contexts tells the same story, with digital and AI-supported tools proving most educationally meaningful when embedded in a defined experiential or expressive pedagogy rather than deployed as stand-alone enhancements (Ahdhianto et al., 2025; Dima et al., 2026; Kaşkaya & Ateş, 2025).

Prior syntheses leave this intersection largely unmapped. Bibliometric reviews of calligraphy education have appeared in Chinese-language venues, but their domestic focus and limited international visibility restrict their usefulness to the international research community. The one comprehensive English-language survey adjacent to the topic reviews algorithms for handwritten Chinese character quality evaluation, a methodological catalogue that does not address learning or teaching (Chen et al., 2024). Large-scale reviews of STEAM education, for their part, give no sustained attention to heritage arts. The result is a body of work that is substantial, growing and essentially unsynthesized: no English-language review identified through these searches has yet asked what technology has actually done for calligraphy education, where the evidence lies, or how it hangs together as a field.

1.4. The Aim

This study addresses that gap through a systematic review and bibliometric analysis of research on technology-enhanced calligraphy education published between 2015 and 2025. Five questions structure the inquiry:

RQ 1) What are the publication trends and geographic distribution of research in this field?

RQ 2) Which source outlets and studies exert the greatest influence?

RQ 3) What thematic structure emerges from keyword co-occurrence patterns?

RQ 4) How are technologies integrated into the teaching and learning of calligraphy, across education levels and pedagogical approaches?

RQ 5) What gaps in the evidence base define the future research agenda?

By mapping the field and synthesizing its themes, the study makes three contributions: it provides the first English-language systematic account of the field; it evaluates the extent to which technology integration amounts to genuine STEAM pedagogy; and it proposes a framework for making the science and mathematics of calligraphy explicit in instructional design.

2. Methods

2.1. Review Design

The review followed the PRISMA 2020 guideline for transparent reporting of systematic reviews (Page et al., 2021), combining bibliometric performance analysis in the tradition of Donthu et al. (2021) with an integrative thematic synthesis. The pairing responds to the shape of the field under study. In a small and fast-moving corpus, counts alone can mislead: a run of technically similar papers may dominate the numbers while the studies of greatest pedagogical significance remain numerically invisible. A purely narrative reading carries the opposite risk, missing the structural facts that give the field its character, where it is published, by whom, and along which technical lines it is organized. The bibliometric layer therefore maps production and reception,

while the thematic layer reads every included study for what the technology actually does in an educational setting; the two layers were designed to check each other, and points where they diverge are treated as findings rather than noise. The protocol was defined in advance of screening and is reported in full below; the review was not prospectively registered, a choice noted here for transparency.

2.2. Data Source and Search Strategy

Three databases served as data sources: the Web of Science Core Collection, selected for its quality-controlled metadata and cited-reference fields; Scopus, selected for its broader coverage of education venues and conference literature; and the China National Knowledge Infrastructure (CNKI), added to represent the extensive Chinese-language scholarship of a field centred on a Chinese cultural practice. The Web of Science search was executed in July 2026 and the Scopus and CNKI searches in August 2026, all against titles, abstracts and keywords, combining three concept blocks: calligraphy terms; education terms; and technology or STEAM terms. Publication years were restricted to 2015–2025 and document types to journal articles, conference papers and reviews. Language was restricted to English in Web of Science and Scopus; the CNKI search was conducted in Chinese, and because CNKI indexes the full range of Chinese periodicals including non-academic magazines, an additional source-quality threshold was applied there, retaining only records from Peking University Core, CSSCI, CSCD or EI-indexed journals to approximate the selectivity of the international databases. The full search string appears in Table 1.

The three concept blocks encode the review's framing and deserve comment. The calligraphy block was kept inclusive of regional terms so that work on Japanese shūji, Korean seoye and Arabic or Persian traditions would be captured rather than filtered out by an anglicized vocabulary. The education block requires teaching, learning or skill transmission to be at stake, committing the review to an educational rather than purely technical reading of the corpus from the outset. The third block pairs conventional educational-technology terms with the STEAM family of terms; the pairing is operational rather than decorative, since a central question of the review is whether technology integration in this field amounts to STEAM pedagogy, and only a search covering both vocabularies could answer it. The start year of 2015 reflects two considerations: deep-learning methods entered calligraphy research in the middle of the decade, and the 2013 ministerial guidelines gave the subject fresh policy visibility in the world's largest calligraphy-education system. Searching three databases guards against the coverage biases that single-database reviews of interdisciplinary fields invite, and the CNKI arm answers directly the objection that an English-only corpus cannot support claims about a field whose largest research community publishes in Chinese: where the retrievals diverge, the difference is itself diagnostic, a point taken up in the results. Because CNKI indexes practitioner magazines alongside peer-reviewed research, Chinese-language records were additionally required to appear in Peking University Core, CSSCI, CSCD or EI-indexed sources, a threshold applied to approximate the selectivity of the international databases.

As a check on the inclusiveness of the calligraphy concept block, both international-database searches were rerun on 13 October 2026 with romanized regional labels (shuji, shodo, seoye, khatt, khat) added to the first concept block and all other parameters unchanged. The expanded strings retrieved 118 records from Web of Science (95 with the original string under identical conditions) and 279 from Scopus (236). Title-level screening of the additional records (66 in total, 47 unique once cross-database overlap is removed) identified none that met the eligibility criteria. Most were Arabic-script recognition or writer-identification studies without an educational context; the remainder were unrelated records retrieved through ambiguous term matches, including public-health studies of khat chewing, a film study referencing the director Terayama Shuji, Kurdish-script digitization, and work on handwriting forensics and spectral imaging. No eligible English-language studies of Japanese, Korean or Arabic calligraphy education were indexed in either database during the review period, a coverage gap taken up in Section 4.5.

Table 1
Search strategy and eligibility criteria

Component	*Specification*
Databases	Web of Science Core Collection, including SCI-EXPANDED, SSCI and ESCI (searched July 2026); Scopus (searched 19 August 2026); CNKI (searched 19 August 2026; journal and conference records, 2015–2025). For CNKI, a source-quality threshold (Peking University Core, CSSCI, CSCD or EI) was applied at the screening stage rather than in the database search.
Search field	Title, abstract and author keywords (Web of Science Topic; Scopus TITLE-ABS-KEY; CNKI subject)
Search string (Web of Science)	TS=((calligraph* OR “shufa” OR “brush writing”) AND (education OR teaching OR learning OR curriculum OR pedagogy OR instruction OR classroom OR student* OR school*) AND (“STEAM” OR “STEM” OR “arts integration” OR “art integration” OR interdisciplinar* OR transdisciplinar* OR “digital technolog*” OR “educational technolog*” OR “artificial intelligence” OR “machine learning” OR “deep learning” OR “virtual reality” OR “augmented reality” OR “mixed reality” OR robot* OR “e-learning” OR “online learning” OR “digital media” OR multimedia OR “digital ink” OR haptic*))
Search string (Scopus)	TITLE-ABS-KEY((calligraph* OR "shufa" OR "brush writing") AND (education OR teaching OR learning OR curriculum OR pedagogy OR instruction OR classroom OR student* OR school*) AND ("STEAM" OR "STEM" OR "arts integration" OR "art integration" OR interdisciplinar* OR transdisciplinar* OR "digital technolog*" OR "educational technolog*" OR "artificial intelligence" OR "machine learning" OR "deep learning" OR "virtual reality" OR "augmented reality" OR "mixed reality" OR robot* OR "e-learning" OR "online learning" OR "digital media" OR multimedia OR "digital ink" OR haptic*)) AND PUBYEAR > 2014 AND PUBYEAR < 2026 AND (LIMIT-TO(LANGUAGE, "English")) AND (LIMIT-TO(DOCTYPE, "ar") OR LIMIT-TO(DOCTYPE, "cp") OR LIMIT-TO(DOCTYPE, "re"))
Search string (CNKI)	SU=(‘书法’+‘书写’)*(‘教学’+‘教育’+‘学习’+‘课程’+‘培训’)*(‘人工智能’+‘深度学习’+‘机器学习’+‘神经网络’+‘虚拟现实’+‘增强现实’+‘混合现实’+‘元宇宙’+‘机器人’+‘在线’+‘网络’+‘数字化’+‘数字媒体’+‘多媒体’+‘智慧教育’+‘智能’+‘STEAM’+‘STEM’)
Timespan	1 January 2015 to 31 December 2025
Language	English (Web of Science and Scopus); Chinese (CNKI)
Document types	Article, proceeding paper, review
Inclusion criteria	Studies of calligraphy teaching, learning, assessment or skill transmission in which digital technology or STEAM-framed pedagogy plays a substantive role
Exclusion criteria	Recognition/OCR and heritage digitization without educational context; robotic writing or motion control without pedagogical application; font generation or typesetting without learning-resource context; calligraphy peripheral to the study

Note. Early-access records dated 2026 and retracted publications were removed before screening. Records retrieved from the three databases were deduplicated by DOI and normalised-title matching before screening.

2.3. Eligibility Criteria

Eligibility required an educational or skill-transmission context in addition to a technology component. Studies of calligraphy recognition, optical character recognition or heritage digitization were excluded where no teaching or learning application was present, as were robotics studies concerned solely with motion control and generation studies aimed at typesetting rather than learning materials. Studies in which calligraphy appeared only as a passing example, and studies of clearly non-educational applications, were likewise excluded. The criteria were applied at title-and-abstract level, with borderline records resolved by reading the full abstract and, where necessary, the full text. Two kinds of record tested the criteria most. Engineering papers occasionally gestured at education in their introductions while contributing nothing pedagogical in method or evaluation; these were excluded. Robotic and generative studies posed the harder case, and one rule was applied to them throughout: a study qualified only where skill transmission to learners formed part of the stated purpose and the system embodied a concrete mechanism serving that purpose, whether by reproducing strokes for demonstration, guiding the learner's hand, or generating material for practice. Invoking cultural inheritance as background motivation, without such a mechanism, did not qualify. The 29 robotic writing and motion-control records excluded at screening failed this test, and the robotics studies retained in the corpus pass it; the application of the criteria to every included study is documented in Supplementary Table S1.

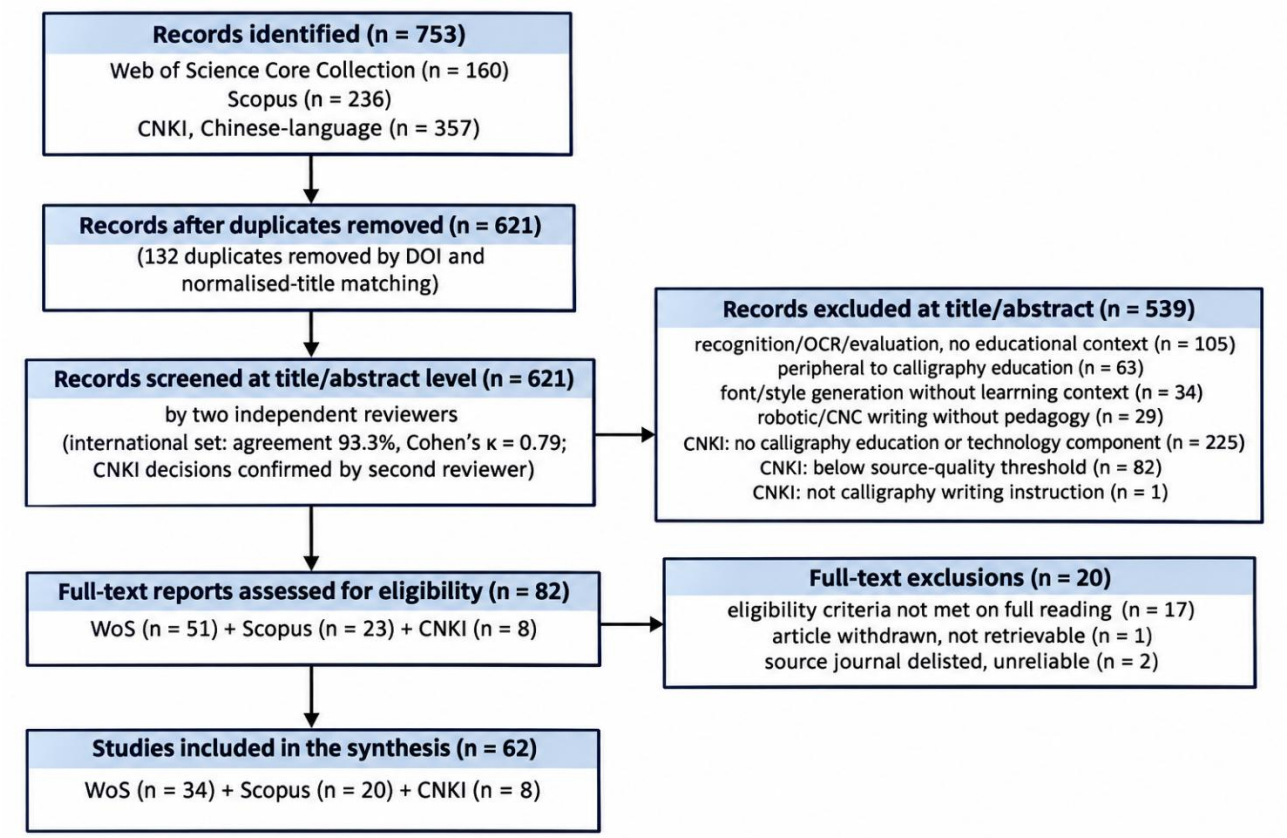
2.4. Screening Procedure

Screening proceeded in two stages under dual review. At the title-and-abstract stage, two reviewers independently applied the eligibility criteria to the 264 deduplicated international records and reached the same decision on 257 of them (97.3% agreement, Cohen's $\kappa = .79$). The seven disagreements were resolved through discussion against the codebook, and two records were reinstated at this point where discussion showed the stricter reading to have been inconsistent with prior inclusions of the same study type. The 357 Chinese-language records were screened with the identical criteria and source-quality threshold, and all Chinese-language inclusion decisions were confirmed by the second reviewer. At the full-text stage, 82 reports were assessed and 20 were excluded for failing the eligibility criteria on full reading. Every candidate record was additionally screened for retraction status and journal indexing; this source-reliability check was applied uniformly to all candidates, with simultaneous delisting from both international databases as its objective criterion, and it was not prespecified in the review protocol. Four of the excluded reports came from a journal later delisted from Web of Science and Scopus amid paper-mill concerns, but each had already failed the eligibility criteria on content grounds, so no theme count or substantive conclusion depends on the delisting rule. The process is summarized in the PRISMA flow diagram (see Figure 1).

2.5. Data Analysis

Descriptive bibliometric indicators were computed for publication years, countries and regions, institutions, source outlets and citation counts. For the co-occurrence analysis, author keywords and Keywords Plus were standardized through a synonym table that merged spelling variants, acronyms and near-synonyms (for example, variants of generative adversarial network (GAN), and regional calligraphy traditions under calligraphy). Keywords occurring in at least two studies were retained, and the resulting network was partitioned with the Clauset–Newman–Moore modularity algorithm (Clauset et al., 2004), which yields the cluster structure used for interpretation. The network was constructed and partitioned in Python 3.11 with NetworkX 3.4 (Hagberg et al., 2008), and figures were drawn with Matplotlib 3.10.9 (Hunter, 2007). Citation counts were retrieved on 19 August 2026. Thematic coding then classified each included study on technology type, calligraphic tradition, education level or context, and research type, with a short statement of its core contribution; the full coding table is provided as Supplementary Table S1.

Figure 1
PRISMA 2020 flow diagram of study selection



Thematic coding followed an iterative, constant-comparison procedure. An initial codebook was drafted from the eligibility categories and the technology terms of the search strategy, then applied to every included study and revised as recurring distinctions emerged, most importantly the separation of robotics aimed at cultural reproduction from haptic systems aimed at skill transfer, and of assessment research from generic recognition work. Coding reliability was verified independently on the final 62-study corpus: a second analyst, working from the titles and abstracts of the sampled studies and the same codebook, and blind to the primary coding, classified the one-third sample (21 of the 62 studies, selected at fixed intervals) on all four dimensions. Agreement with the primary coding was complete across the 75 judgements returned. The nine cells left undecided owing to title-level ambiguity were excluded from the agreement calculation; for these cells the primary coding was verified against the full texts and retained, so no undecided judgement entered the synthesis unchecked. Co-occurrence ties were counted on the standardized keyword set without fractional weighting, and the Clauset–Newman–Moore partition was preferred because it requires no prespecified cluster number and performs well on networks of this size (Clauset et al., 2004). Because keyword networks in this review are built from a vocabulary that partially overlaps the search string, a sensitivity analysis re-estimated the network under two restrictions: with all query-derived terms removed, and with index keywords only; the resulting structures are reported alongside the primary analysis.

Methodological quality was appraised with the Mixed Methods Appraisal Tool (MMAT), version 2018 (Hong et al., 2018), which applies five design-appropriate criteria within each of five study categories. Every empirical study was first assigned to the category matching its own design: randomized trials to the randomized-trial category, other experiments and quasi-experiments to the non-randomized category, and system-development and technical-validation studies to the quantitative descriptive category, since such studies report performance against benchmark or sampled data rather than intervention effects. Criterion-level judgements were then made against each study's own design and stated aims rather than an external educational benchmark: a criterion that presupposes learner outcomes was recorded as not applicable where the study's

purpose did not include them, and rated unclear only where the report provided insufficient information, so that a technically sound validation study without learner outcomes is not conflated with a poorly conducted one. Conceptual papers, design proposals and reviews are not appraisable under MMAT and were retained as not applicable rather than excluded, since removing them would erase the field's conceptual layer; they are flagged wherever study-type counts are reported. Supplementary Table S3 reports each appraisable study's category, its criterion-level ratings and the reason for every negative or unclear judgement, and the synthesis conditions the strength of its claims on these results.

3. Results

3.1. Study Selection

The three searches retrieved 753 records (160 from Web of Science, 236 from Scopus, 357 from CNKI). After removal of 132 duplicates, 621 records entered title-and-abstract screening, which excluded 539 records: recognition, OCR, heritage digitization or automated evaluation studies without an educational context (n = 105); studies in which calligraphy was peripheral or the application non-educational (n = 63); font generation or typesetting studies without a learning-resource context (n = 34); robotic writing or motion-control studies without a learner-directed mechanism (n = 29); CNKI records lacking either a calligraphy-education or a substantive technology component (n = 225); CNKI records below the source-quality threshold (n = 82); and one Chinese-language record on typing-based literacy rather than handwriting instruction (n = 1). Full-text assessment of the remaining 82 reports excluded a further 20, and the final corpus contains 62 studies, 34 from Web of Science, 20 from Scopus and 8 Chinese-language studies from CNKI, comprising 31 journal articles and 31 conference papers. The corpus is youthful even in its composition: the median publication year is 2023, and the two reviews in the set survey evaluation algorithms and virtual reality (VR) haptic applications respectively rather than learning research, which leaves the field without a pedagogically oriented synthesis until now.

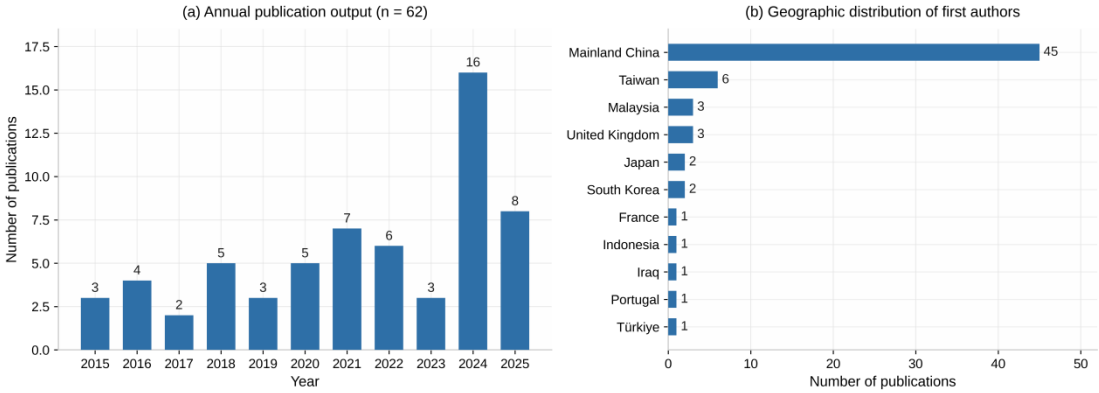
The overlap among the three databases is itself informative. Scopus contributed 104 unique records, yet only 20 survived screening, and the exclusion profile of these records reproduces that of the Web of Science set almost exactly, recognition-style engineering work dominates in both. The Chinese-language arm proved starker still: 357 records yielded eight eligible studies, an inclusion rate of 2.2%, because practitioner experience essays dominate the Chinese magazine literature while rigorous studies cluster in a small number of core journals. The three-source design therefore offers no sign that the boundaries drawn in this review are an artefact of any single index or language, although residual coverage bias cannot be excluded.

3.2. Publication Trends and Geographic Distribution

Annual output was modest and stable between 2015 and 2019, at one to five publications per year, and began to expand in 2020. Growth accelerated from 2023 onward and peaked in 2024 with 16 publications. Over half of the corpus, 35 of 62 studies, appeared in the final four years of the review window, a pattern consistent with an emerging field whose visibility is still rising (see Figure 2).

Geographic concentration is pronounced and, crucially, survives two independent expansions. Under full counting, which assigns studies by all author affiliations so that multi-country studies contribute to every affiliated country, mainland China accounts for 45 of the 62 studies, Taiwan for six, the United Kingdom for four, Malaysia for four, Japan and South Korea for three each, and France, Indonesia, Iraq, Portugal and Türkiye for one each. Country totals therefore sum to 71, exceeding the 62-study corpus as multi-country studies are counted for each affiliation. Of the 20 studies added through Scopus, 14 are China-affiliated, and all eight Chinese-language core-journal studies originate from mainland institutions: two further, independent retrievals replicate the concentration rather than diluting it, which indicates that the pattern is not an artefact of Web of Science coverage alone, although the English-only restriction of the international databases means

Figure 2
Annual publication trend and geographic distribution of included studies



that retrieval design may still shape the geographic distribution. Xiamen University remains the most productive institution with five studies, followed by Soochow University, Northumbria University, Aberystwyth University and Yuan Ze University with four each, while the CNKI arm reveals a second sustained programme at Nanjing Normal University, whose three studies on intelligent handwriting assessment span 2016 to 2018. What the expansion did change is cultural range: the corpus now spans six calligraphic traditions, Chinese (n = 49), Arabic (n = 5), Japanese (n = 3), Korean (n = 2), Persian (n = 1) and general or Western traditions (n = 2), with the Arabic-language tradition in particular growing from a single early tutoring system to a small cluster spanning Indonesia, Iraq and Malaysia, and Western flat-tip calligraphy entering the corpus through a Portuguese self-learning application.

Institutional patterns sharpen the picture. Xiamen University's five studies form the field's most sustained research programme, centred on robotic calligraphy and machine assessment, while the British cluster at Northumbria and Aberystwyth universities has concentrated on immersive environments, reflecting those institutions' strengths in human–computer interaction. Co-authorship across national borders is uncommon in the corpus, and where it occurs it links technical partners more often than it connects education researchers working in different calligraphic traditions. The practical consequence is that systems travel poorly: models built for Chinese calligraphy have seldom been adapted to Korean, Japanese, Arabic or Persian practice, even where the underlying technology would transfer directly.

3.3. Source Outlets and Influential Works

The 62 studies are dispersed across 55 source outlets, and only a handful published more than one study, among them IEEE Access, Multimedia Systems, the ACM International Conference Proceeding Series and Springer's Advances in Intelligent Systems and Computing. No education journal has yet become a natural home for this research; publication venues range across education technology, human–computer interaction, robotics, computer vision and cultural heritage, testimony to a field that is genuinely interdisciplinary but not yet consolidated as a community.

Citation impact was assessed within each study's own indexing database, since the citation pools of the Chinese- and English-language literatures barely overlap. The 54 international-database studies have attracted 193 citations in total (M = 3.6), recorded in the Web of Science Core Collection for the 34 Web of Science studies and in Scopus for the 20 Scopus-added studies at the time of retrieval. The eight Chinese-language studies, indexed in neither international database, have attracted 64 further citations as recorded in CNKI (M = 8.0); that their mean exceeds the international figure signals influence that international databases do not capture. Taken together, the corpus has accumulated 257 citations (M = 4.1 per study), although the two strata rest on different citation pools and are not directly comparable. Table 2 lists the ten most-cited studies among the international-database studies. Their pattern is instructive: the three most influential works correspond to the field's three dominant technical directions, mixed-reality learning, AI-

supported assessment, and calligraphy robotics, suggesting that these directions carry the field's accumulated research capital.

The outlets group into recognizable disciplinary homes. Education and educational-technology journals host the pedagogically oriented minority; human–computer-interaction venues carry the immersive and haptic work; and robotics, vision and computing outlets hold the technical majority. Citation impact is correspondingly uneven. The ten studies in Table 2 account for 156 of the 193 citations recorded for the international-database studies, more than four-fifths of that stratum, and several of them date from the field's early years, when a technical demonstration could still define a direction. Within the Chinese-language stratum the same concentration appears around intelligent assessment: Jiang et al. (2016) and Xu et al. (2016) hold 26 and 15 CNKI-recorded citations respectively, the two highest counts in that stratum. Most included studies, particularly those published after 2022, have yet to accumulate measurable influence, a pattern typical of young fields whose citation economy is dominated by a small vanguard.

Table 2
The ten most-cited studies among the international-database studies

Study	*Technology focus*	*Outlet*	*Citations*
Loup-Escande et al. (2017)	Mixed reality learning	Computers in Human Behavior	27
Li et al. (2022)	VR distance learning	Frontiers in Education	24
Wu et al. (2020b)	Actor–critic calligraphy robot	Neurocomputing	20
Sun et al. (2023)	Aesthetic evaluation (SRAFE)	CAAI Transactions on Intelligence Technology	19
Wu et al. (2020a)	GAN calligraphy robot	Information Sciences	19
Lee and Hsu (2021)	AR learning for older adults	Int. J. Environmental Research and Public Health	15
Wu et al. (2018)	Deep reinforcement learning robot	IEEE ROBIO	14
Zhang et al. (2017)	Style-consistency detection	ICIVC	7
Chao et al. (2019)	Robot with human preference	IEEE SmartWorld	6
Kobayashi and Katsura (2022)	Generative calligraphy model	Precision Engineering	5

Note. Citation counts as recorded in the Web of Science Core Collection (Web of Science studies) and Scopus (Scopus-added studies) at the time of retrieval. The eight Chinese-language studies are not indexed in either database and do not enter this table; their 64 CNKI-recorded citations are reported in the text.

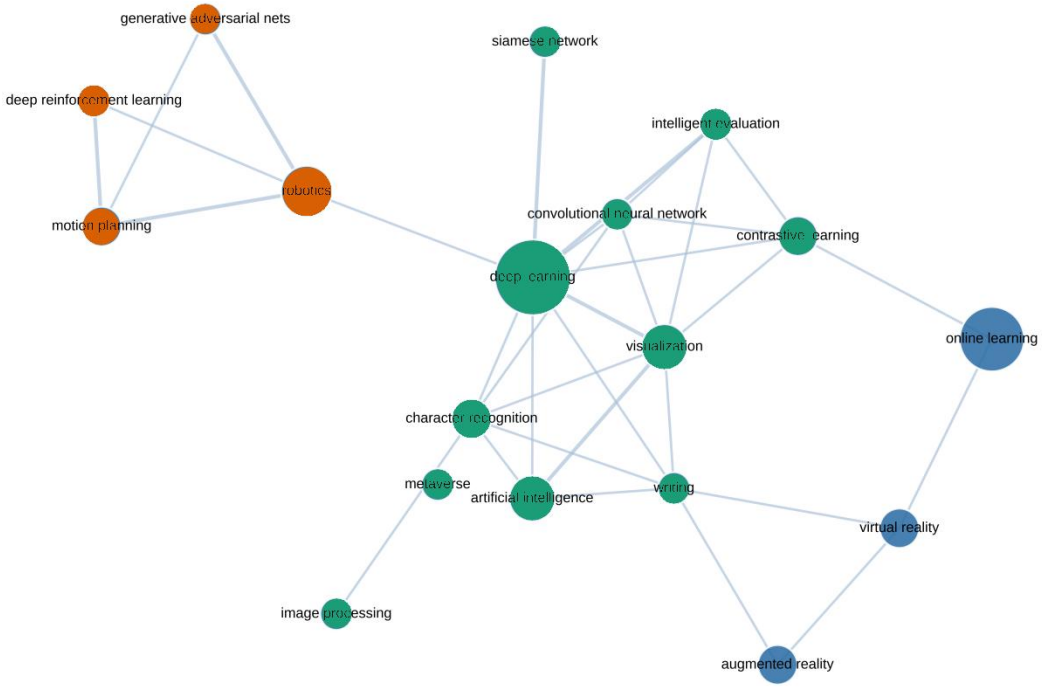
3.4. Keyword Co-occurrence and Thematic Clusters

Standardization of author keywords across the 53 international-database studies that carry them yielded 25 keywords occurring in at least two studies; CNKI records, whose keyword metadata follow different conventions, were excluded from the network analysis for consistency. The two principal hubs are deep learning (frequency 10) and calligraphy-family terms, and modularity-based partitioning resolves two dominant communities (Figure 3): an AI-assessment community gathering deep learning, character recognition, convolutional neural networks, siamese networks, image processing and intelligent or calligraphy evaluation; and a robotics community assembling generative adversarial networks, robot control, motion planning and deep reinforcement learning. Immersive and online terms, virtual reality, metaverse, visualization, writing, bridge the two, marking the field's most education-oriented edge.

Because the network is built partly from the same vocabulary that defined the search, circularity is a live concern, and two sensitivity analyses address it directly. First, removing the calligraphy, education and teaching terms carried over from the search string leaves an 18-keyword network whose two-community structure is preserved intact: the clusters therefore reflect how authors describe their work, not how the search string named the subject. Technology terms shared with

the search string, among them deep learning and virtual reality, remain in this specification because they also serve as the field's own labels for its methods; it is their pattern of co-occurrence, not their presence, that carries the interpretation. Second, rebuilding the network from index keywords alone yields a complementary hub organized around students, e-learning and curricula, confirming that an educational layer exists in the corpus even though it rarely drives author-chosen terminology. The wiring remains as informative as the clusters: educational framing enters the field mainly through studies of learner experience and platform delivery rather than through instructional design.

Figure 3
Keyword co-occurrence network



Note. Keyword co-occurrence network of the included studies after removal of the calligraphy, education and teaching terms carried over from the search string (sensitivity specification); node size is proportional to keyword frequency, and colour distinguishes the two detected communities, with bridging terms left in blue

3.5. Technology Integration in Calligraphy Education: A thematic Synthesis

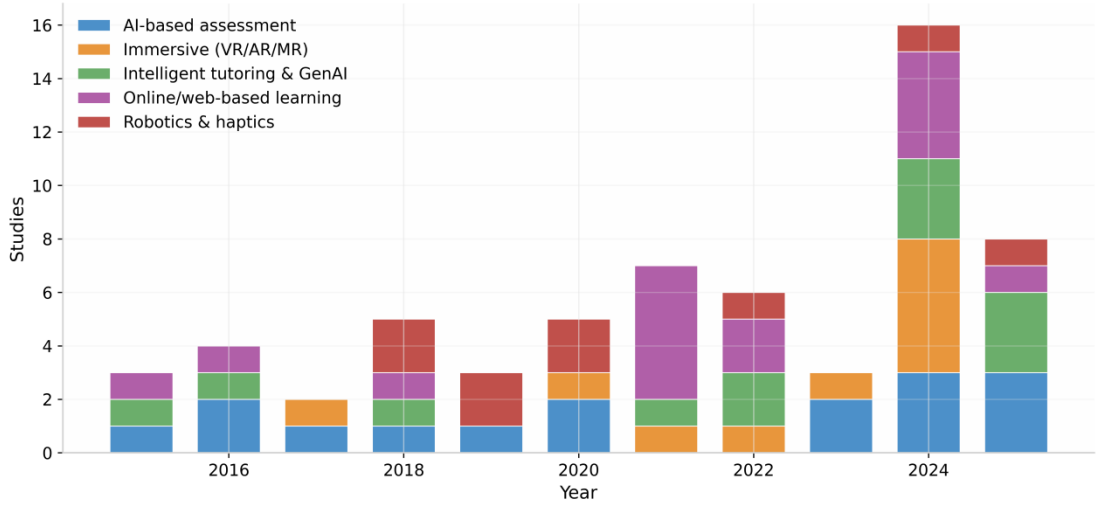
Full-text coding organized the corpus into five themes (see Table 3 and Figure 4). Each study was assigned to a single primary theme according to its dominant technology function, with borderline cases resolved by the study's stated primary contribution; studies whose technologies span categories are therefore counted once, and the counts are analytically simplified rather than exhaustive.

The largest theme, AI-based assessment (n = 16), applies deep architectures to scoring and feedback, and its methodological progression runs from engineered image features toward end-to-end similarity learning. An early system paired deep features with a one-class support vector machine to flag departures from a model style in beginners' copying (Zhang et al., 2017); later work adopted siamese designs that learn similarity directly, culminating in SiamHCC, a siamese network for handwritten-character quality evaluation (Chen et al., 2025), and in a siamese transformer built for intelligent scoring of hard-pen calligraphy (Yan et al., 2024). Convolutional and transformer architectures are likewise used to evaluate writing standardization (Yan et al., 2024), while aesthetic modelling has grown more elaborate, from siamese regression that fuses multiple aesthetic cues for copy evaluation (Sun et al., 2023) to interpretable models that expose which visual features drive an aesthetic score (Sun et al., 2025). Two studies address school-age populations directly, through hard-pen classification for primary and secondary students (Qiang et

Table 3
Thematic distribution of included studies (n = 62)

Theme	*n*	*Representative studies*
AI-based assessment of calligraphy quality	16	Chen et al. (2024); Sun et al. (2023); Sun et al. (2025); Jiang et al. (2016); Liu and Zhang (2025)
Online and web-based learning	15	Wong et al. (2024); Ch'ng et al. (2025); Mi (2018); Alashari et al. (2024); Wu et al. (2024)
Intelligent tutoring, generative AI and pedagogy	12	Gong et al. (2025); Natheer (2024); Mamoun et al. (2015); Sadikoglu and Soroush (2016); Tang et al. (2025)
Immersive learning environments (virtual reality, augmented reality and mixed reality)	10	Loup-Escande et al. (2017); Luo et al. (2023); Li et al. (2022); Yu and Ma (2024); Li, Bao, and Yang (2024)
Robotics and haptic skill training	9	Wu et al. (2018, 2020a, 2020b); Chao et al. (2019); Raza et al. (2019); Tsuda et al. (2017)

Figure 4
Thematic evolution of the corpus, 2015–2025



al., 2019) and image-recognition evaluation linked to children's writing development (Avila Cordova et al., 2020), and one system turns the same machinery on instruction itself, assessing calligraphy teaching quality with a BP neural network optimized by a sparrow search algorithm (Wang, 2023). The 2024 comprehensive survey confirms that character quality evaluation has matured into a distinct methodological line with shared tasks and benchmarks (Chen et al., 2024). The expansion adds two entries that fit the same mould, a ResNet-based scorer for single characters aimed at self-practice (Liu et al., 2020) and a CNN–transformer hybrid with stroke-level segmentation for hard-pen evaluation (Liu & Zhang, 2025), reinforcing rather than altering the theme's centre of gravity. The Chinese-language arm adds a coherent earlier programme that the international indexes had missed: between 2015 and 2018 the Nanjing Normal University group built a digital-capture evaluation of handwriting quality (Xu et al., 2016), established the validity of feature-based intelligent scoring of hard-pen neatness (Jiang et al., 2016), designed an intelligent-assessment-integrated teaching scheme for teacher trainees (Ma et al., 2018), and, in parallel work at Shaanxi Normal University, an ASP.NET platform for assessing blackboard-writing skill (Li & Zhou, 2015). Validation in this stream remains overwhelmingly technical; integration into live classroom assessment is the exception, although the Nanjing programme's validity study marks an early step in that direction.

Immersive learning environments (n = 10) recreate studio conditions for practice, and the theme's experimental core is small but consistent. Loup-Escande et al. (2017) manipulated supplementary visual feedback in a mixed-reality tracing task and found that benefits concentrate

among novices, with expertise moderating both cognitive load and gestural performance; Luo et al. (2023) showed that an inexpensive sponge can stand in for the brush's compliant resistance, letting VR learners rehearse the pressure control that ordinary controllers cannot convey. Augmented reality has served both ends of the age spectrum, underpinning self-learning of Japanese calligraphy in school settings (Ishihara & Makino, 2020) and proving competitive with e-book instruction in a quasi-experiment with older adults during the pandemic (Lee & Hsu, 2021), while a mixed-reality studying environment has been prototyped to situate practice among digitized masterworks (Wang & Gao, 2024). Two survey studies complement these laboratory results by modelling adoption at scale: one examined the factors shaping students' intention to use metaverse technology for calligraphy learning (Li, He, & Yang, 2024), and another combined coolness theory, the value-based adoption model and the technology acceptance model to account for learners' behaviour in spherical video virtual reality courses (Li, Bao, & Yang, 2024), with perceived usefulness, interactivity and novelty emerging as salient predictors. Two further studies push the theme toward the metaverse: a mixed-methods study reported motivational and cultural-engagement gains for international learners in metaverse-based Chinese character and calligraphy education (Yu & Ma, 2024), and an analytical examination catalogued how current VR applications still fail to convey the tactile grain of brush, ink and paper that calligraphy practice depends on (Ma & Yu, 2024).

Robotics and haptic skill training (n = 9) embodies a paradigm of guided doing, every included study having passed the two-part eligibility rule of Section 2.3, and its centre is a single sustained programme. The Xiamen group progressed from a deep reinforcement learning robot that acquires stroke policies (Wu et al., 2018) through inverse reinforcement learning from human preferences (Chao et al., 2019) to GAN-based and actor–critic systems that reproduce master-level brush work with increasing fidelity (Wu et al., 2020a, 2020b), all framed around cultural inheritance and education. A generative model trained on images and human motion extends the same reproduction logic to robotic inheritance in Japan (Kobayashi & Katsura, 2022), and the most recent entry positions an AI-powered calligraphy creation system explicitly as a teaching tool (Hu et al., 2025). Haptic work addresses the learner's body rather than the machine's artistry: stroke-level force feedback has been used to transfer brush skills in Korean calligraphy (Raza et al., 2019), an accessible three-axis plotter scaffolds novices through generated motion at a fraction of conventional hardware costs (Fang et al., 2024), and a pressure-presentation device conveys expert brushwork force directly to apprentices' hands in Japanese calligraphy instruction (Tsuda et al., 2017). Rigorous classroom evaluation of learning effects is scarce here as well; most systems are demonstrated rather than tested with learners.

Intelligent tutoring, generative AI and pedagogy (n = 12) spans the field's longest arc and, after the database expansion, its widest cultural range. The earliest included study proposed educational software for the basic skills of Arabic calligraphy (Mamoun et al., 2015), and an intelligent tutoring system for Persian calligraphy followed (Sadikoglu & Soroush, 2016). Later work examined intelligent calligraphy teaching within digital-media environments (Liu et al., 2022), and generative AI has most recently entered the construction of multimodal resources for Chinese character learning (Yu et al., 2025). The expansion enriches this theme on three fronts. For the Arabic tradition, a qualitative study documents digital calligraphy learning innovation in cyberspace media (Pujiati, 2022), a framework for interactive tutoring in Islamic calligraphy handwriting supplies design guidelines grounded in instructor shortage (Jamaluddin et al., 2024), and the corpus's strongest quasi-experimental evidence outside East Asia shows that a generative-AI tool improved art-education students' design proficiency for Arabic calligraphy ornamentation, with design skills mediating the effect (Natheer, 2024). For Western calligraphy, a Portuguese project delivers a complete self-directed learning stack, application, system and tooling, for flat-tip practice (Brandão, 2018). And the methodologically strongest new entry is CalliSense, an interactive tool for process-based learning developed through formative interviews and reported at CHI, which targets the procedural knowledge that static exemplars cannot transmit (Gong et al., 2025). Conceptual strategy papers from the Chinese higher-education context (Du, 2024; Xia et al.,

2021) round out the theme. This theme also contains the corpus's single explicit STEAM study: Tang et al. (2025) tested a STEAM project-based learning (PBL) intervention with vocational students and reported positive effects on calligraphy vocational skills and artistic creativity, with learning motivation mediating the effect, a result that supplies the field's clearest evidence that an integrated design can move the outcomes teachers care about.

Online and web-based learning (n = 15) delivers calligraphy at distance through platforms that pair deep-learning recognition with personalized feedback. A web application for Chinese learners combines convolutional neural networks with scale-invariant feature transform to assess uploaded practice and return corrective guidance (Wong et al., 2024), and a parallel platform for Korean calligraphy applies deep learning to the same instructional loop of submission, evaluation and feedback (Ch'ng et al., 2025). Alongside these systems, survey-based modelling has examined the factors and mediation effects behind learners' intention to study calligraphy online (Wu et al., 2024), extending the acceptance-research tradition that runs through the corpus. The Scopus expansion makes this the fastest-growing theme: a virtual 3D camera system supports intelligent remote calligraphy instruction (Xu, 2024); big-data analytics and k-means clustering of achievement data inform course design for primary and secondary schools (Li, 2022; Ouyang, 2021); micro-class and interest-cultivation schemes adapt calligraphy teaching to fragmented online time (Wang, 2021; Z. Yu, 2021); multimedia and digital-media pathways frame calligraphy education for public communication (Dai, 2021; Zhou, 2021); and a case study of the Calligrapher platform documents the first large-scale e-learning environment for Arabic calligraphy and Islamic ornament, hosted in Malaysia (Alashari et al., 2024). The Chinese-language literature anticipated much of this development: core-journal studies argued for online courses as a pathway for school calligraphy (Mi, 2018), analysed the distinctive features of calligraphy education in networked environments, proposed a writing-instruction platform for international students, and reconstructed character-writing materials around a smart-writing curriculum for schools (Li, 2022). Much of this work is design-oriented and evaluatively thin, but it establishes that platform-mediated calligraphy learning now operates at scale in both Chinese and Arabic traditions.

Behind the thematic map stands a small empirical core, assembled in Table 4. Ten studies report learner outcome or performance data, and even these are thin: intervention duration is stated in none of them, samples are small, and barely half compare conditions. The practical implications drawn in the discussion are calibrated to this evidence ceiling.

3.6. Research Gaps Observed in the Corpus

Four gaps emerge directly from the coding. Explicit STEAM integration is almost absent: every included study involves technology, yet only one of 62 frames calligraphy education through STEAM as a pedagogical construct, which means the "A" is present in content but rarely in instructional design. Effect evidence is thin: system-development studies account for half of the corpus (n = 31), while experimental, quasi-experimental and empirical classroom studies total 15 (24.2%). The MMAT appraisal sharpens this diagnosis, of the 48 appraisable empirical studies, only three met four of the five design-appropriate criteria, 17 met three, and 28 met no more than two, and no study met all five. The dominant limit was missing evidence of learning rather than defective method. System-development studies frequently satisfied every criterion internal to their own validation designs while reporting no learner outcomes, and this gap is recorded separately from methodological weakness in Supplementary Table S3. Formal schooling is under-represented, with K-12 appearing in nine studies and vocational education in one, while nearly half of the corpus targets unspecified general learners, a mismatch with policy contexts in which calligraphy is compulsory school content. The evidence base is culturally concentrated: 51 of 62 studies (82.3%) carry mainland-China or Taiwan affiliations, although the expanded corpus now contains meaningful Arabic-tradition work and a first Western-tradition entry. These gaps structure the discussion that follows.

(Insert Table 4)

Table 4
Empirical studies reporting learner outcomes or learner performance data

Study	*Participants*	*Design*	*Duration*	*Measured outcomes*
Loup-Escande et al. (2017)	Novices and experts (adults)	Between-subjects experiment (MR tracing, supplementary visual feedback)	Not reported	Gestural performance; cognitive load
Qiang et al. (2019)	Primary and secondary students	Validation study on student writing	Not reported	Hard-pen writing-quality classification (K-12 samples)
Raza et al. (2019)	Korean calligraphy learners	Experiment (haptic guidance)	Not reported	Brush-skill transfer
Avila Cordova et al. (2020)	Children	Evaluation of image-recognition scoring	Not reported	Writing-development evaluation
Lee and Hsu (2021)	Older adults	Quasi-experiment (AR vs. e-book)	Not reported	Learning effectiveness; motivation; technology acceptance; cognitive load
Luo et al. (2023)	VR learners	Experiment (compliant proxy controller)	Not reported	Pressure-control performance
Yu and Ma (2024)	International learners	Mixed methods (metaverse course)	Not reported	Motivation; cultural engagement
Fang et al. (2024)	Novices	Plotter-guided practice study	Not reported	Guided motion practice
Natheer (2024)	Art-education students	Quasi-experiment (generative-AI tool; two groups, N = 60)	Not reported	Ornamentation design proficiency
Tang et al. (2025)	Vocational students	STEAM-PBL intervention (quasi-experimental)	Not reported	Vocational skills; artistic creativity

Note. Studies reporting learner outcome or learner performance data are listed; system-validation studies without learner data and perception-only surveys are excluded. Duration is rarely reported in the source studies, which itself illustrates the reporting limits discussed in the text.

4. Discussion

4.1. Principal Findings

This review set out to map a field that many educators sense exists but few could delineate. The mapping confirms that technology-enhanced calligraphy education is real, young and accelerating: annual output tripled after 2020 and reached its highest level in 2024, a trajectory that coincided with the pandemic-era normalization of distance learning, the maturation of deep-learning methods for character analysis, and policy attention to calligraphy in national curricula. The bibliometric analysis establishes the temporal pattern rather than its causes, and these accounts remain contextual interpretations rather than tested explanations. The field's intellectual centre of gravity sits in AI-based assessment, closely trailed by online delivery and intelligent tutoring, with immersive learning and robotics as substantial technical flanks. Production is concentrated geographically in mainland China and Taiwan, a concentration that recurs across three databases in two languages rather than arising from any single index, institutionally in a handful of universities, and methodologically in system development, which accounts for half of all included studies. Quality appraisal adds a sobering layer: the studies with the strongest designs, such as the CHI-reported CalliSense and the Arabic-tradition quasi-experiment, are recent and few, while much of the older technical core rests on validation without learners. Against this background, the solitary explicit STEAM study stands out less as an outlier than as a proof of concept: Tang et al. (2025) demonstrated that a STEAM-PBL design can move both vocational calligraphy skills and artistic creativity, with motivation mediating the effect.

Read against the research questions, the picture is coherent. Publication trends (RQ1) describe an emerging field on the steep part of its growth curve, concentrated in the Chinese-speaking world but with footholds in Britain, Japan, Korea and Malaysia. Influence (RQ2) accrues to the three technical flagships, mixed-reality learning, AI assessment and robotics, whose leading works anchor the citation counts. The keyword structure (RQ3) resolves into two communities that mirror the field's technical organization, with immersive and online terms bridging them rather than standing as a community of their own. Technology integration (RQ4) is dominated by system building across five themes and spans every education level from primary pupils to older adults, yet formal K-12 schooling, the context in which calligraphy is compulsory, remains thinly covered. The gaps (RQ5) converge on a single diagnosis: the field has invested heavily in what technology can do and sparingly in what learners gain.

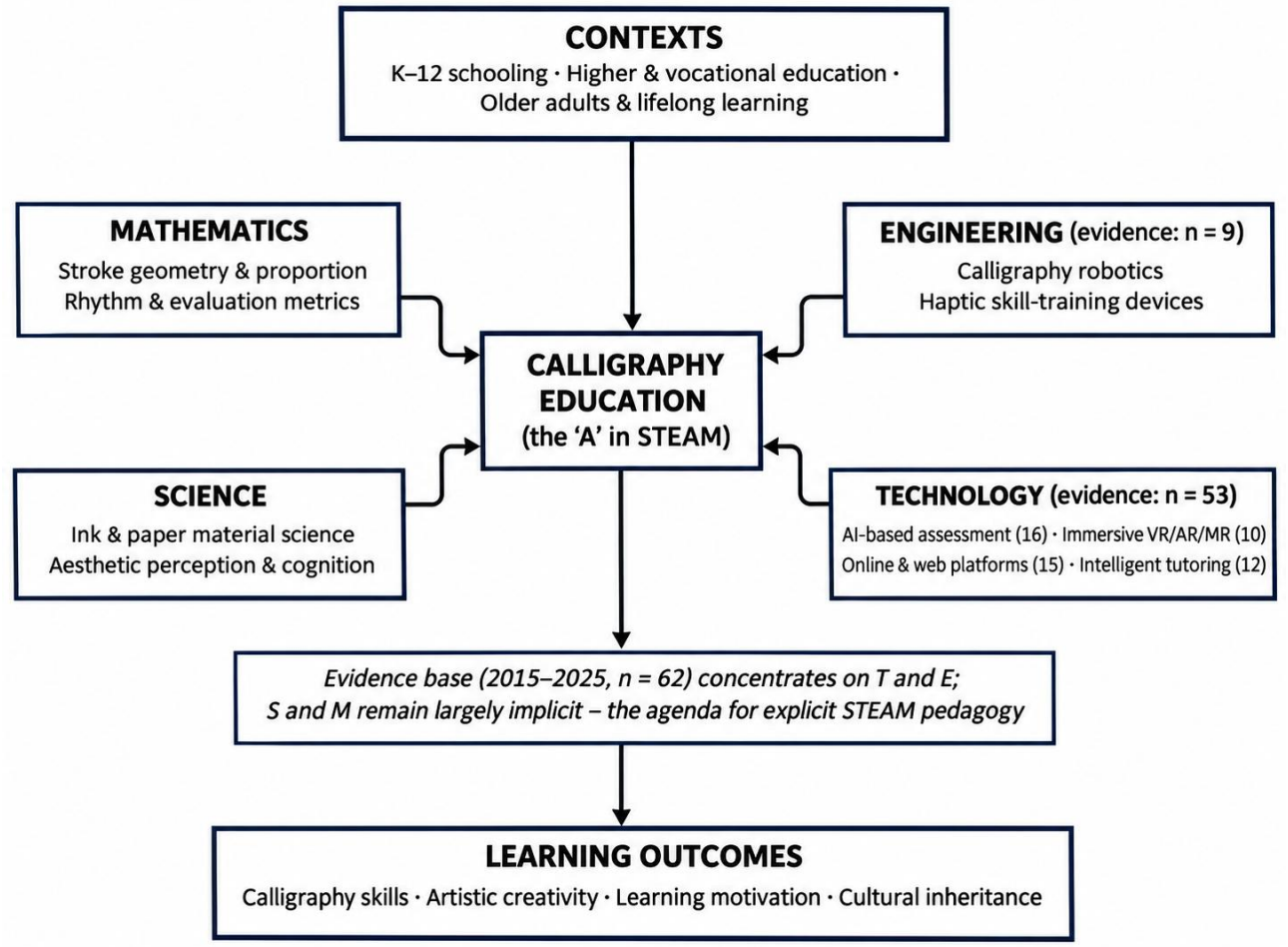
4.2. From Technology Use to STEAM Pedagogy: Reframing Calligraphy Education

The central finding of this review is also the least obvious from publication counts alone. Every included study deploys technology in the service of calligraphy education, yet only one of 62 organizes that deployment as STEAM pedagogy. The framework advanced in this section should therefore be read from the outset as a research agenda derived from the corpus's gaps, not as a description of established practice, which the evidence does not yet support. The distinction matters. Perignat and Katz-Buonincontro (2019) showed that much of what travels under the STEAM label amounts to arts-flavoured activity without interdisciplinary substance, and Mejías et al. (2021) documented how loosely the framework is often applied. Judged against that caution, the present corpus is technology-rich but integration-poor: systems are built on or for calligraphy, while the science and mathematics internal to the art form remain implicit in instructional design.

This limited integration is particularly noteworthy because calligraphy itself offers a natural context for STEAM education. The behaviour of ink on sized paper is material science; the perception of balance and rhythm in a composition is cognition; the structure of a character is geometry governed by proportion and symmetry; the evaluation of a copy is measurement against quantifiable criteria; and the engineering of writing machines forces an explicit account of the physics of the brush. Building on this observation, Figure 5 proposes a framework in which calligraphy education occupies the arts position, with the four surrounding components made curricularly explicit rather than assumed. Science enters through ink and paper behaviour and

through the cognitive science of aesthetic perception. Technology enters through the field's best-developed assets: AI-based assessment, immersive environments and online platforms. Engineering enters through robotics and haptic devices that render brush mechanics inspectable. Mathematics enters through stroke geometry, proportional analysis and the metrics underlying automated evaluation. The framework is deliberately symmetrical with the discipline it serves: each letter corresponds to content a calligraphy teacher can actually teach, which is precisely the condition under which STEAM stops being decoration and starts being curriculum. The proposed framework follows the broader principle that interdisciplinary design becomes educationally meaningful when disciplinary knowledge is made explicit within a shared design task, rather than merely co-located within an activity (Fikri et al., 2026; Tang et al., 2025).

Figure 5
A STEAM integration framework for calligraphy education



The evidence base currently populates the framework unevenly. Technology and engineering are well resourced by the 62 studies in those themes, whereas the science and mathematics components appear mainly as tacit features inside systems, for instance, geometric stroke models embedded in robots or aesthetic metrics embedded in evaluators, rather than as taught content. Tang et al. (2025) offers the empirical precedent that an integrated design is workable; the framework generalizes that precedent and gives the field a shared vocabulary for designing the next generation of interventions.

What the framework asks of teachers is concrete rather than rhetorical. A science inquiry might have learners measure how ink spread varies with paper sizing and brush moisture, connecting material behaviour to the visual qualities they are trained to control. A mathematics sequence could treat the character as a geometric object, with pupils measuring stroke angles, testing the symmetry claims of classical copybooks and expressing proportion as ratio, so that the rules they memorize become quantifiable. Engineering design finds a natural home in building and

calibrating a simple writing machine, where the gap between planned and produced strokes makes brush mechanics tangible. Technology supplies the feedback infrastructure the corpus has already built, so that automated assessment informs rather than terminates the learning conversation. Each activity is recognizably calligraphy and recognizably STEAM at once, which is the framework's point: integration draws on content the discipline already contains instead of importing alien content for the sake of a label.

The framework should be read as a design heuristic and a research agenda rather than a validated model. Its components are grounded in the reviewed evidence, but no study has yet tested a curriculum built explicitly on it, and the balance among components will rightly vary by tradition and education level: the mathematics of proportion means something different to a primary pupil tracing her first characters than to a vocational student preparing for the cultural industries. Empirical validation of the framework is therefore itself part of the research agenda, most plausibly through design-based studies in which teachers co-construct integrated units and learning is measured on both artistic and interdisciplinary outcomes.

4.3. Implications for Practice

For school systems, the most immediately transferable asset is the assessment literature. Hard-pen evaluation systems built for primary and secondary learners and for children's writing development suggest that automated scoring is a promising candidate for carefully monitored formative piloting rather than a mature solution ready for routine implementation, a judgement consistent with broader evidence that AI-supported tools should be weighed by pedagogical fit, learner outcomes and cognitive demands as much as by technical performance (Danişman & Evran Acar, 2026; Kavitha & Joshith, 2024), particularly in contexts like mainland China's, where the 2022 curriculum standard mandates calligraphy but qualified teachers are unevenly distributed. An assessment tool that delivers consistent, criterion-referenced feedback does not replace the teacher so much as extend the teacher's capacity across the one-to-many classroom.

Vocational education holds the corpus's only direct effect evidence, and it points the same way with the caution that a single study affords: one STEAM-PBL intervention improved both skills and creativity among vocational students, a preliminary proof of concept that warrants replication across institutions and traditions before integrated designs are adopted beyond pilot programmes. At the other end of the age spectrum, a single quasi-experiment found augmented reality competitive with e-book instruction for older adults during the pandemic, and a low-cost plotter has lowered the hardware barrier to guided practice, giving community and lifelong-learning providers low-cost options to pilot rather than a proven infrastructure path.

Teacher preparation emerges from the reviewed studies as one major condition shaping how far technology becomes pedagogically integrated. Where calligraphy remains a stand-alone subject, technology in the corpus typically enters as remediation for instructor scarcity; where programmes reframe it as STEAM-anchored content, the same tools tend to support interdisciplinary units that also serve science and mathematics objectives. Current programmes rarely equip calligraphy teachers to orchestrate data-rich tools or to read automated feedback critically, a capacity gap consistent with wider evidence that teachers' digital and AI-related competencies shape how far new tools become integrated practice (Çakır & Güner, 2026; Çengel & Acar, 2026). Museums and cultural institutions occupy an underused position in this ecology: immersive installations and robotic demonstrations are already familiar to public audiences, and converting them from spectacle into structured learning experiences would extend the field's reach well beyond formal schooling.

Two cautions temper these implications. The first is that technical validation is not pedagogical validation: most systems were evaluated for accuracy rather than for learning, and classrooms will surface requirements that laboratories do not. The second concerns the teacher's role. The acceptance studies in the corpus indicate that perceived usefulness and interactivity drive learners' intention (Davis, 1989), which implies that tools must be woven into instruction by an orchestrating teacher rather than handed to learners as self-contained solutions.

4.4. Implications for Research

The clearest methodological imperative is to convert system development into effect research. Quasi-experimental and randomized designs with measured learning outcomes, skill gain, retention, transfer, motivation, are needed across all five themes, and the assessment literature in particular awaits classroom-based validation against expert judgement. Intention models have served their purpose in opening the field; the next stage requires behavioural and achievement data. Reviews of AI in adjacent subject areas report the same imbalance, with classroom-based evaluation lagging behind the rapid diversification of tools (Kavitha & Joshith, 2024).

A second imperative is cultural breadth. Arabic and Persian calligraphy appear in the corpus only through early tutoring systems, despite their deep pedagogical traditions, and cross-cultural comparative designs would test whether the patterns observed in East Asian contexts generalize. Researchers might also exploit the data exhaust of deployed systems: usage logs from online platforms and assessment tools could support learning-analytics studies of how practice behaviour predicts progress, an approach none of the included studies attempted. Such work will depend on interdisciplinary teams in which education researchers frame the questions and computer scientists frame the systems, a division of labour the current publication pattern suggests is still exceptional.

Instrument development is a precondition for the effect research just described. Validated measures of calligraphic skill that go beyond expert impression would make classroom trials comparable across studies, and the assessment literature is the obvious source from which to build them. Longitudinal designs are equally necessary, since brush skill develops over years while every experiment in the corpus is short-term. Generative AI, finally, deserves close watching as the newest integration pathway: its first appearance in the corpus concerns resource construction rather than interaction with learners, and whether generated exemplars and feedback help or hinder the acquisition of a disciplined hand is an open question with immediate classroom stakes.

4.5. Limitations

Several limitations bound these findings. The language and database design is regionally asymmetric: Web of Science and Scopus were restricted to English-language records, and only Chinese-language scholarship received a dedicated vernacular database search. The review is therefore better grounded for Chinese calligraphy than for Japanese, Korean, Arabic or Persian traditions, and the geographic concentration reported in Section 3.3, though replicated across three databases, cannot be assumed independent of retrieval design. The review drew on three databases in two languages; disciplinary and regional indexes beyond Web of Science, Scopus and CNKI were not searched, and relevant studies in venues they do not cover, including other calligraphic traditions' vernacular literatures such as Japanese-, Korean- and Arabic-language scholarship, may have been missed. The CNKI arm was restricted to core-journal sources to approximate the selectivity of the international databases, a threshold that trades completeness for comparability and may discard sound work published in practitioner venues. The corpus remains small, at 62 studies, so keyword clusters are interpreted conservatively as indicative themes rather than definitive boundaries, although the sensitivity analyses show that their structure does not depend on query vocabulary. Dual independent screening achieved substantial reliability ($\kappa = .79$), and thematic coding was independently verified on a one-third sample with complete agreement; even so, full double coding of every study was not undertaken, and quality appraisal was conducted by one analyst, so residual single-coder bias cannot be excluded. The MMAT ratings for a subset of technically oriented studies were inferred from published reports of uneven detail and should be read as indicative. A source-reliability screen (retraction status and journal indexing) was applied uniformly to every candidate record; the screen was not prespecified, and although four excluded reports came from a journal later delisted amid paper-mill concerns, all four had independently failed the eligibility criteria, so no substantive conclusion rests on this rule. Finally, early-access studies dated 2026 were excluded by design, and the field's rapid growth means the picture presented here will need periodic updating.

5. Conclusion

Between 2015 and 2025, technology-enhanced calligraphy education grew from a scattering of demonstrations into a recognizable research field with a clear centre in AI-based assessment and strong technical flanks in immersive learning and robotics. Its output is concentrated in the Chinese-speaking world even when indexed independently across three databases and two languages, its evidence is concentrated in system development of modest methodological quality, and its pedagogy is, with a single exception among 62 studies, not yet STEAM pedagogy in any explicit sense. The synthesis presented here shows that the raw materials for integration already exist inside the discipline: calligraphy contains physics, geometry, cognition and measurement as naturally as it contains art. The framework proposed in this study organizes those materials into a curriculum-level design in which the arts component is not an ornament but the anchor. For a subject that sits on UNESCO's heritage list while facing a shortage of qualified teachers, the stakes of getting this integration right are unusually concrete. If the field's next decade pairs its technical fluency with effect evidence and explicit interdisciplinary design, calligraphy education could become one of the most instructive cases of heritage arts meeting twenty-first-century pedagogy, advancing, in a small but tangible way, the inclusive-quality-education ambition of the United Nations' Sustainable Development Goal 4 (United Nations, 2015).

The study's contributions are threefold. Empirically, it supplies the first systematic map, identified through these searches, of technology-enhanced calligraphy education spanning international and Chinese-language scholarship, documenting a 62-study, three-database corpus whose contours were previously visible only in fragments. Conceptually, it distinguishes technology use from STEAM pedagogy and shows that the field has produced abundant instances of the former and only one of the latter, a gap that reframes what integration should mean for heritage arts. Practically, it offers a framework through which teachers, curriculum designers and researchers can make the science and mathematics of calligraphy explicit, together with an agenda that privileges effect evidence, cultural breadth and formal schooling. These contributions matter beyond calligraphy: heritage art forms worldwide face the same pairing of cultural obligation and instructional scarcity, and the trajectory traced here, technology first and pedagogy catching up, is likely to be theirs as well.

Author contributions: Enqi Tang: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization, Writing - review & editing. Siti Nur Diyana Mahmud: Supervision, Writing - review & editing. Nur Atiqah Jalaludin: Validation, Writing - review & editing. Mohamad Sattar Rasul: Writing - review & editing. All authors have read and agreed to the published version of the manuscript.

Data availability: The bibliographic dataset analysed in this study was retrieved from the Web of Science Core Collection, Scopus and the China National Knowledge Infrastructure (CNKI); the screening records, the full coding table of included studies and the MMAT quality appraisal are provided as Supplementary Material (Supplementary Tables S1–S3).

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

Ethical statement: As this study is based exclusively on a systematic review and analysis of previously published literature, it did not involve human participants or the collection of primary data and therefore did not require ethical approval.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Ahdhianto, E., Barus, Y. K., & Thohir, M. A. (2025). Augmented reality as a game changer in experiential learning: Exploring its role cultural education for elementary schools. *Journal of Pedagogical Research, 9*(1), 296–313. https://doi.org/10.33902/JPR.202533573

Alashari, D., Berghout, A., & Sohsan, I. (2024). The role of electronic platforms in the educational process: Calligrapher platform as a model. In *2024 9th International Conference on Business and Industrial Research (ICBIR)* (pp. 193-198). IEEE. https://doi.org/10.1109/ICBIR61386.2024.10875046

Brandão, J. (2018). Calligraphy practice app, system and tool for learning calligraphy. In A. G. Ho (Ed.), *Advances in communication of design: AHFE 2017* (Advances in Intelligent Systems and Computing, Vol. 609, pp.97-103). Springer. https://doi.org/10.1007/978-3-319-60477-0_11

Çakır, D., & Güner, P. (2026). The current landscape of teachers' artificial intelligence acceptance: Relationships with TPACK and technostress. *International Journal of Didactical Studies, 7*(3), Article e43911. https://doi.org/10.33902/ijods.202643911

Çengel, E., & Acar, F. (2027). Teachers' online teaching competencies for sustainable digital education: Evidence from a cross-national survey. *Sustainable Education and Digital Transformation, 2*(1), Article e44144. https://doi.org/10.33902/SEDT.202644144

Chao, F., Lyu, J.-T., Wu, R.-Q., Gao, X.-G., Zhou, C.-L., Yang, L.-Z., Lin, C.-M., & Shang, C.-J. (2019). Robotic Chinese calligraphy with human preference. In *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing* (pp. 360-366). IEEE. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00105

Chen, W.-R., Su, J.-Q., Song, W.-T., Xu, J.-L., Zhu, G.-Q., Li, Y., Ji, Y., & Liu, C.-P. (2024). Quality evaluation methods of handwritten Chinese characters: A comprehensive survey. *Multimedia Systems, 30*(4), Article 194. https://doi.org/10.1007/s00530-024-01396-8

Chen, W.-R., Zhu, G.-Q., Li, Y., Ji, Y., & Liu, C.-P. (2025). SiamHCC: A novel siamese network for quality evaluation of handwritten Chinese characters. *Multimedia Systems, 31*(4), Article 303. https://doi.org/10.1007/s00530-025-01886-3

Ch'ng, G. Y., Osman, M. A., & Sarvghadi, M. A. (2025). A web-based platform for Korean calligraphy learners using deep learning. *IEEE Access, 13*, 165970–165988. https://doi.org/10.1109/ACCESS.2025.3612578

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E, 70*(6), Article 066111. https://doi.org/10.1103/PhysRevE.70.066111

Avila Cordova, A. W., Flores Choque, A., & Paucar Nuñez, J. C. (2020). Evaluating the quality of a person's calligraphy using image recognition. *International Journal of Advanced Computer Science and Applications, 11*(8), 739–744. https://doi.org/10.14569/IJACSA.2020.0110890

Dai, Y. (2021). Research on the application of computer multimedia in calligraphy education. *Journal of Physics: Conference Series, 1915*(3), Article 032019. https://doi.org/10.1088/1742-6596/1915/3/032019

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

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Dima, A., Kotsidis, K., & Lafazanis, I. S. (2026). Drama education and artificial intelligence in sixth-grade geography lessons: An approach to cultivating critical thinking. *International Journal of Didactical Studies, 7*(2), e37844. https://doi.org/10.33902/ijods.202637844

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

Du, Y. (2024). The innovative path of calligraphy education in universities in the era of artificial intelligence. *Revista Ibérica de Sistemas e Tecnologias de Informação*, *67*, 560–569. https://risti.xyz/issues/ristie67.pdf

Elmali, S., & Balkan Kiyici, F. (2025). Pre-service primary teachers' STEM design experiences. *Journal of Pedagogical Research, 9*(3), 25–42. https://doi.org/10.33902/JPR.202528478

Fang, C.-M., Huang, L.-D., Kuang, Q., Lieberman, Z., Maes, P., & Ishii, H. (2024). An accessible, three-axis plotter for enhancing calligraphy learning through generated motion. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp.1-11). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642792

Fikri, R. A., Suwono, H., Susilo, H., & Kusairi, S. (2026). Fostering students' creative thinking through inquiry-design-based STEM water purification project. *Journal of Pedagogical Research, 10*(4), 81–104. https://doi.org/10.33902/JPR.202638877

Gong, X., Tao, W., & Ma, Y. (2025). CalliSense: An interactive educational tool for process-based learning in Chinese calligraphy. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Article 1058, pp. 1-20). Association for Computing Machinery. https://doi.org/10.1145/3706598.3714176

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function

using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference* (pp. 11–15). SciPy. https://doi.org/10.25080/TCWV9851

Hong, Q. N., Pluye, P., 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. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018*. Canadian Intellectual Property Office, Industry Canada.

Hu, S.-M., Wang, N., Zhao, K.-X., Jing, Y., Zhang, Y., Wang, Z., Guo, B., & Yu, Z.-W. (2025). Masterpiece creation: An AI-powered robotic calligraphy creation system. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 9*(3), Article 87. https://doi.org/10.1145/3749957

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Ishihara, Y., & Makino, M. (2020). AR based self-learning system of Japanese calligraphy skills. In P. Y. Lau, & M. Shobri (Eds.), *International Workshop on Advanced Imaging Technology (IWAIT) 2020* (Article 1151525). SPIE. https://doi.org/10.1117/12.2566958

Jamaluddin, N., Mat Zain, N. H., & Diah, N. M. (2024). Blending tradition with innovation: An interactive tutoring in handwriting Islamic calligraphy. In *2024 IEEE 22nd Student Conference on Research and Development (SCOReD)* (pp. 652-657). IEEE. https://doi.org/10.1109/SCOReD64708.2024.10872672

Jiang, J., Xu, H., Fan, S.-Q., Qian, X.-F., & Li, Y. (2016). Jiyu tezheng biaoji de yingbi hanzi shuxie gongzhengxing zhineng ceping xiaodu yanjiu [Validity of feature-marking-based intelligent assessment of hard-pen Chinese character writing neatness]. *Dianhua Jiaoyu Yanjiu, 37*(6), 84–89.

Kavitha, K., & Joshith, V. P. (2024). Pedagogical incorporation of artificial intelligence in K-12 science education: A decadal bibliometric mapping and systematic literature review (2013–2023). *Journal of Pedagogical Research, 8*(4), 437–465. https://doi.org/10.33902/JPR.202429218

Kobayashi, R., & Katsura, S. (2022). A generative model of calligraphy based on image and human motion. *Precision Engineering, 77*, 340-348. https://doi.org/10.1016/j.precisioneng.2022.06.006

Kaşkaya, A., & Ateş, Ş. (2025). The effect of artificial intelligence-supported visualization applications on students' writing disposition in the creative writing process. *Journal of Pedagogical Research, 9*(2), 5–21. https://doi.org/10.33902/JPR.202531916

Lee, C.-J., & Hsu, Y. (2021). Promoting the quality of life of elderly during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health, 18*(13), Article 6813. https://doi.org/10.3390/ijerph18136813

Liao, C. (2016). From interdisciplinary to transdisciplinary: An arts-integrated approach to STEAM education. *Art Education, 69*(6), 44-49. https://doi.org/10.1080/00043125.2016.1224873

Li, J., & Zhou, Y.-W. (2015). Banshu shuxie nengli zhineng ceping pingtai de yanjiu yu sheji [Research and design of an intelligent assessment platform for blackboard-writing ability]. *Jisuanji Jishu yu Fazhan, 25*(3), 154–157.

Li, J., Luo, H., Zhao, L., Zhu, M., Ma, L., & Liao, X. (2022). Promoting STEAM education in primary school through cooperative teaching: A design-based research study. *Sustainability, 14*(16), Article 10333. https://doi.org/10.3390/su141610333

Li, J.-J. (2022). Shuzihua zhuanxing beijing xia hanzi shuxie xuecai de zaigou yu shishi celüe: Yi "'Shuangyou Yunqiao · Shoushi Youzi' zhihui shuxie" kecheng wei li [Reconstructing and implementing Chinese character-writing learning materials under digital transformation: The "Shuangyou Yunqiao · Shoushi Youzi" smart-writing course]. *Zhongxue Yuwen Jiaoxue*, *12*, 23–26.

Li, P.-F., Fang, Z.-L., & Jiang, T. (2022). Research into improved distance learning using VR technology. *Frontiers in Education, 7*, Article 757874. https://doi.org/10.3389/feduc.2022.757874

Li, Y. (2022). Innovative strategies of primary school calligraphy education model under the background of big data. In J. C. Hung, J. W. Chang, Y. Pei, & W. C. Wu (Eds.), *Innovative computing* (pp. 529-536). Springer. https://doi.org/10.1007/978-981-16-4258-6_66

Li, Y., Bao, S.-L., & Yang, Y. (2024b). Research on college students' behavior in calligraphy courses within a virtual learning environment based on SVVR: A convergent model approach of coolness theory, VAM and TAM. *IEEE Access, 12*, 28538-28553. https://doi.org/10.1109/ACCESS.2024.3365556

Li, Y., He, Z.-N., & Yang, Y. (2024a). Bridging tradition and innovation: Exploring the factors influencing students' intention to use metaverse technology for Chinese calligraphy learning. In J. Y. C. Chen, & G. Fragomeni (Eds.), *Virtual, Augmented and Mixed Reality: 16th International Conference, VAMR 2024* (pp. 22-43). Springer. https://doi.org/10.1007/978-3-031-61041-7_3

Lin, C.-L., & Tsai, C.-Y. (2021). The effect of a pedagogical STEAM model on students' project competence and learning motivation. *Journal of Science Education and Technology, 30*(1), 112-124. https://doi.org/10.1007/s10956-020-09885-x

Liu, H., Wang, Y.-X., Ma, J.-W., Yu, J.-Y., & Ma, Y.-F. (2022). Study on the application of intelligent calligraphy teaching in the context of digital media. In M.M. Soares, E. Rosenzweig, & A. Marcus (Eds.) *Design, user experience, and usability: design thinking and practice in contemporary and emerging technologies* (pp. 365-383). Springer. https://doi.org/10.1007/978-3-031-05906-3_27

Liu, L., & Zhang, L. (2025). Research on intelligent evaluation algorithm for hard-pen calligraphy based on deep learning. In *2025 8th International Conference on Algorithms, Computing and Artificial Intelligence (ACAI)* (pp. 1-5). IEEE. https://doi.org/10.1109/ACAI68217.2025.11406637

Liu, S., Wang, Z., Wang, C., Zheng, J., & Zhang, F. (2020). Preliminary design and application prospect of single Chinese character calligraphy image scoring algorithm. In J. S. Pan, J. Li, P. W. Tsai, & L. Jain (Eds.), *Advances in intelligent information hiding and multimedia signal processing* (pp. 443-450). Springer. https://doi.org/10.1007/978-981-13-9710-3_46

Loup-Escande, E., Frenoy, R., Poplimont, G., Thouvenin, I., Gapenne, O., & Megalakaki, O. (2017). Contributions of mixed reality in a calligraphy learning task: Effects of supplementary visual feedback and expertise on cognitive load, user experience and gestural performance. *Computers in Human Behavior, 75*, 42-49. https://doi.org/10.1016/j.chb.2017.05.006

Luo, G.-L., Lu, T.-S., Xia, H.-B., Hu, S.-C., & Guo, S.-H. (2023). Learning Chinese calligraphy in VR with sponge-enabled haptic feedback. *Interacting with Computers, 35*(4), 530-542. https://doi.org/10.1093/iwc/iwad041

Ma, N., & Yu, D. (2024). How do VR applications address the challenges of conveying the tactile feedback crucial to traditional calligraphy practice. In *Proceedings of the 16th International Conference on Computer Supported Education (CSEDU 2024)* (Vol. 1, pp. 744-750). SciTePress. https://doi.org/10.5220/0012753600003693

Ma, Y.-Q., Song, C.-X., Xu, M.-Y., Li, Y., & Qiu, H.-B. (2018). Jiehe zhineng ceping xitong yingyong de shifansheng yingbi shufa jiaoxue fangan sheji [Design of a hard-pen calligraphy teaching programme for pre-service teachers integrated with an intelligent assessment system]. *Xiandai Jiaoyu Jishu, 28*(12), 101–107.

Mamoun, S., Rahman, A., Eldin, S., & Elhamid, A. (2015). A proposal to build and design an educational software program to teach the basic skills of Arabic calligraphy. In *Proceedings 2015 Fifth International Conference on e-Learning (ECONF)* (pp. 534-538). IEEE. https://doi.org/10.1109/ECONF.2015.49

Mejías, S., Thompson, N., Sedas, R. M., Rosin, M., Soep, E., Peppler, K., Roche, J., Wong, J., Hurley, M., Bell, P., & Bevan, B. (2021). The trouble with STEAM and why we use it anyway. *Science Education, 105*(2), 209-231. https://doi.org/10.1002/sce.21605

Mi, W.-Z. (2018). Wangluo kecheng kaifa: Dangqian zhongxiaoxue shufa jiaoxue de xin chulu [Developing online courses: A new pathway for calligraphy teaching in primary and secondary schools]. *Kecheng, Jiaocai, Jiaofa, 38*(7), 87–91.

Ministry of Education of the People's Republic of China. (2013). *Zhongxiaoxue shufa jiaoyu zhidao gangyao* [Guidelines for calligraphy education in primary and secondary schools]. https://www.moe.gov.cn/srcsite/A26/s8001/201301/t20130125_147389.html

Ministry of Education of the People's Republic of China. (2022). *Yiwu jiaoyu yishu kecheng biaozhun (2022 nian ban)* [Compulsory education arts curriculum standards (2022 edition)]. Beijing Normal University Press.

Natheer, W. I. (2024). The impact of Dream Artificial Intelligence relevance and design skills proficiency on art education students in designing innovative decorative units. *Educational Sciences: Theory & Practice, 24*(2), 107-121. https://jestp.com/menuscript/index.php/estp/article/view/1905

Ouyang, X. (2021). Design of calligraphy online course in primary and secondary schools based on k-means algorithm. In V. Sugumaran, Z. Xu, & H. Zhou (Eds.), *Application of intelligent systems in multi-modal information analytics: MMIA 2020* (pp. 630–635). Springer. https://doi.org/10.1007/978-3-030-51556-0_92

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*, Article n71. https://doi.org/10.1136/bmj.n71

Perignat, E., & Katz-Buonincontro, J. (2019). STEAM in practice and research: An integrative literature review. *Thinking Skills and Creativity, 31*, 31-43. https://doi.org/10.1016/j.tsc.2018.10.002

Pujiati. (2022). Innovation of learning Arabic calligraphy digital Quran in the cryan world. *Res Militaris, 12*(3), 734-761.

Qiang, X.-X., Wu, M.-H., & Luo, L.-M. (2019). Research on hard-tipped calligraphy classification based on deep learning method. In *Proceedings of 2019 8th International Conference on Educational and Information*

Technology (pp. 275-279). Association for Computing Machinery. https://doi.org/10.1145/3318396.3318417

Raza, A., Abdullah, M., Hassan, W., Abdulali, A., Talhan, A., & Jeon, S. (2019). Painting skill transfer through haptic channel. In H. Kajimoto, D. Lee, S. Y. Kim, M. Konyo, & K. U. Kyung (Eds.), *Haptic Interaction: Perception, Devices and Algorithms* (pp. 66-68). Springer. https://doi.org/10.1007/978-981-13-3194-7_14

Sadikoglu, F., & Soroush, B. (2016). Intelligent system for Persian calligraphy learning. *Procedia Computer Science, 102*, 555–561. https://doi.org/10.1016/j.procs.2016.09.442

Sun, M.-W., Gong, X.-Y., Nie, H.-T., Iqbal, M.-M., & Xie, B. (2023). SRAFE: Siamese regression aesthetic fusion evaluation for Chinese calligraphic copy. *CAAI Transactions on Intelligence Technology, 8*(3), 1077-1086. https://doi.org/10.1049/cit2.12095

Sun, W., Cheng, Y.-L., & Lyu, R.-M. (2025). Interpretable modeling of aesthetic evaluation in calligraphy imitation. In *2025 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-10). IEEE. https://doi.org/10.1109/IJCNN64981.2025.11227615

Tang, E.-Q., Rasul, M.-S., & Jalaludin, N.-A. (2025). Effects of STEAM-PBL on calligraphy vocational skills and artistic creativity: Mediating role of learning motivation. *Foro de Educacion, 23*(2), 338-361. https://forodeeducacion.com/ojs/index.php/fde/article/view/805/612

Tsuda, N., Morikawa, A., Nomura, Y., & Kato, N. (2017). Pressure presentation strength for calligraphy brushwork instruction. In *Proceedings of the 2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)* (pp. 198-202). IEEE. https://doi.org/10.1109/IRIS.2017.8250121

United Nations Educational, Scientific and Cultural Organization. (n.d.). *Chinese calligraphy*. https://ich.unesco.org/en/RL/chinese-calligraphy-00216

United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. https://sdgs.un.org/2030agenda

Wang, C. (2021). Application of offline micro class in calligraphy teaching in the information age. In *Proceedings of the 2021 4th International Conference on Information Systems and Computer Aided Education* (pp. 1374-1376). ACM. https://doi.org/10.1145/3482632.3483154

Wang, Y., & Gao, Z. (2024). Embracing tradition through technology: The mixed reality calligraphy studying environment. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 711-712). IEEE. https://doi.org/10.1109/VRW62533.2024.00152

Wang, Y. (2023). The application of intelligent evaluation method with deep learning in calligraphy teaching. *International Journal of Advanced Computer Science and Applications, 14*(6), 1317-1324. https://doi.org/10.14569/IJACSA.2023.01406139

Wong, A., So, J., & Ng, Z.-T.-B. (2024). Developing a web application for Chinese calligraphy learners using convolutional neural network and scale invariant feature transform. *Computers and Education: Artificial Intelligence, 6*, Article 100200. https://doi.org/10.1016/j.caeai.2024.100200

Wu, C.-H., Liu, C.-H., & Huang, Y.-M. (2022). The exploration of continuous learning intention in STEAM education through attitude, motivation, and cognitive load. *International Journal of STEM Education, 9*(1), Article 35. https://doi.org/10.1186/s40594-022-00346-y

Wu, C.-H., You, A.-H., Dong, T.-P., & Liu, C.-H. (2024). Analysis of factors influencing intention to engage in online learning in Chinese calligraphy and their mediation effects. *Asia-Pacific Education Researcher, 33*(6), 1405-1419. https://doi.org/10.1007/s40299-024-00822-9

Wu, R.-Q., Fang, W.-B., Chao, F., Gao, X.-G., Zhou, C.-L., Yang, L.-Z., Lin, C.-M., & Shang, C.-J. (2018). Towards deep reinforcement learning based Chinese calligraphy robot. In *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (pp. 507-512). IEEE. https://doi.org/10.1109/ROBIO.2018.8664813

Wu, R.-Q., Zhou, C.-L., Chao, F., Yang, L.-Z., Lin, C.-M., & Shang, C.-J. (2020a). GANCCRobot: Generative adversarial nets based Chinese calligraphy robot. *Information Sciences, 516*, 474-490. https://doi.org/10.1016/j.ins.2019.12.079

Wu, R.-Q., Zhou, C.-L., Chao, F., Yang, L.-Z., Lin, C.-M., & Shang, C.-J. (2020b). Integration of an actor-critic model and generative adversarial networks for a Chinese calligraphy robot. *Neurocomputing, 388*, 12-23. https://doi.org/10.1016/j.neucom.2020.01.043

Xia, W., Lyu, R., & Chen, W. (2021). Research on the visualization schemes of writing speed and pressure. In *2021 2nd International Conference on Artificial Intelligence and Information Systems* (Article 257, pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3469213.3470690

Xu, B. (2024). Intelligent remote teaching of calligraphy based on virtual 3D camera technology. In V. Palade, M. Favorskaya, S. Patnaik, M. Simic, & S. Belciug (Eds.), *Artificial intelligence for future society: AIS 2024* (pp. 490-499). Springer. https://doi.org/10.1007/978-3-031-69457-8_45

Xu, H., Qiu, H.-B., Li, Y., & Jiang, J. (2016). Jiyu shuzi caiji shebei de hanzi shuxie zhiliang pingjia yanjiu [Research on evaluating Chinese character handwriting quality using digital capture devices]. *Xiandai Jiaoyu Jishu, 26*(12), 38–43.

Yakman, G., & Lee, H. (2012). Exploring the exemplary STEAM education in the U.S. as a practical educational framework for Korea. *Journal of the Korean Association for Science Education, 32*(6), 1072-1086. https://doi.org/10.14697/jkase.2012.32.6.1072

Yan, F., Lan, X.-P., Zhang, H., & Li, L.-J. (2024). Intelligent evaluation of Chinese hard-pen calligraphy using a siamese transformer network. *Applied Sciences, 14*(5), Article 2051. https://doi.org/10.3390/app14052051

Yan, Z.-H., Lai, S.-R., Zhang, R.-B., & Chen, L.-Y. (2024). Deep learning-based evaluation system for handwritten Chinese character writing standardization. In *2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)* (pp. 388-392). IEEE. https://doi.org/10.1109/MLISE62164.2024.10674339

Yu, D., & Ma, N. (2024). Enhancing Chinese character education and calligraphy through metaverse intelligent technology: Integration and impact. In W. Hong & G. Kanaparan (Eds.), *Computer science and education. Educational digitalization: ICCSE 2023* (pp. 181-193). Springer. https://doi.org/10.1007/978-981-97-0737-9_18

Yu, J.-L., Song, J.-C., & Lu, Y. (2025). Harnessing generative artificial intelligence to construct multimodal resources for Chinese character learning. *Systems, 13*(8), Article 692. https://doi.org/10.3390/systems13080692

Yu, Z. (2021). Online calligraphy teaching interest cultivation scheme based on information technology in the internet era. In *Proceedings of the 2021 4th International Conference on Information Systems and Computer Aided Education* (pp. 1377-1380). ACM. https://doi.org/10.1145/3482632.3483155

Zhang, J.-L., Guo, L.-M., Yang, S., Sun, X.-D., & Li, X.-S. (2017). Detecting Chinese calligraphy style consistency by deep learning and one-class SVM. In *2017 2nd International Conference on Image, Vision and Computing (ICIVC)* (pp. 83-86). IEEE. https://doi.org/10.1109/ICIVC.2017.7984523

Zhou, X. (2021). Exploring new ways of calligraphy education communication in the age of digital media. In C. H. Wu, S. Patnaik, F. Popentiu Vlădicescu, & K. Nakamatsu (Eds.), *Recent developments in intelligent computing, communication and devices: ICCD 2019* (pp. 203-208). Springer. https://doi.org/10.1007/978-981-15-5887-0_30