

The CRAFT Framework: A Pedagogical Model for Teaching Expressive Drawing and Painting with Artificial Intelligence—A Critical Integrative Review

Manal Jumah Ahmad Alhomaiddi 

Department of Visual Arts, College of Art and Design, Taibah University, Medina, Kingdom of Saudi Arabia

Email: mhomydy@taibahu.edu.sa

How to cite this paper: Alhomaiddi, M. J. A. (2026). The CRAFT Framework: A Pedagogical Model for Teaching Expressive Drawing and Painting with Artificial Intelligence—A Critical Integrative Review. *Creative Education*, 17, 1187-1202. <https://doi.org/10.4236/ce.2026.177072>

Received: May 5, 2026

Accepted: July 5, 2026

Published: July 8, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper proposes CRAFT (Conceive, Render, Articulate, reFine, Transmit), a five-stage pedagogical framework for teaching expressive drawing and painting with generative artificial intelligence (AI), and evaluates the framework against the recent empirical literature. A critical integrative review of ten primary studies (eleven datasets, as one study reports two experiments) and three systematic reviews from 2023-2026 was conducted under a Design Science Research orientation. Reported findings from each study were mapped onto CRAFT's five stages by a single author. Within the art-and-design education literature, prompt-engineering literacy and AI knowledge predict innovative thinking and creative self-efficacy in design students; structured AI-assisted instruction is associated with higher achievement, motivation, and self-efficacy in design and art courses; and short, unscaffolded interventions are reported to leave students frustrated with AI's execution of nuanced artistic tasks. Adjacent creative-task evidence provides additional, though non-studio, support for co-creative interface design. Direct experimental evidence from drawing and painting studios remains thin; CRAFT is therefore presented as a candidate scaffold that motivates, but does not substitute for, direct comparative evaluation in studio art settings. The paper contributes a structured pedagogical framework, an honest synthesis of recent primary studies, and a research agenda for future direct evaluation.

Keywords

Artificial Intelligence, Generative Art, Painting and Drawing Pedagogy, Studio Practice, Conceptual Framework, Critical Integrative Review

1. Introduction

Both paintings and drawings co-evolved with technology, from the camera obscura to the modern digital tablet (Manovich, 2018). Generative artificial intelligence (AI) is the latest addition to this evolution: latent diffusion architectures (Rombach et al., 2022) and denoising diffusion probabilistic models (Ho et al., 2020) translate text-based prompts into complete paintings in seconds. The adoption of AI by art and design educators has led to numerous studies examining how this technology affects students' creativity, self-efficacy, motivation, and productivity (Manovich & Arielli, 2024; Boden, 1998; Hertzmann, 2020). According to systematic reviews of recent years, the academic community witnessed a surge in interest, from two empirical studies in 2023 to fourteen by 2025 (Jiang et al., 2025). Research efforts have also converged on the conclusion that positive learning effects hinge on intentional pedagogical scaffolding rather than on the simple provision of tools (Jiang et al., 2025; Wu et al., 2025; Ogunleye et al., 2024).

Even so, there are very few pedagogical frameworks developed for organizing the student studio experience at a level of detail useful to a studio instructor. Current curricula have addressed either prompt-engineering skills or general digital literacy without considering the affective and reflective elements of artistic activity (Park & Choo, 2025; Tour & Zadorozhnyy, 2025).

This paper presents a five-stage framework, CRAFT (Conceive, Render, Articulate, reFine, Transmit), and tests it against the available empirical literature through a critical integrative review. Instead of contributing new empirical results, the present paper investigates whether the existing literature supports, complicates, or undermines the mechanisms posited by CRAFT. This is a deliberately conservative approach: because CRAFT has not yet been tested empirically, the paper examines whether the propositions underlying each individual stage are supported by the empirical literature.

The remainder of the paper provides background in Section 2, the methodology in Section 3, the proposed framework in Section 4, the synthesis and implications in Section 5, and the conclusions in Section 6.

2. Related Works

Recent systematic reviews of generative AI in art education converge on three findings. Jiang et al. (2025) synthesized 19 peer-reviewed empirical studies from 2019-2025 and reported that text-to-image models (DALL-E, Midjourney, Stable Diffusion) and conversational AI dominate the literature, with most studies in higher education and East Asian contexts. Wu et al. (2025) synthesized 99 studies on responses, attitudes, and behaviors and found that perceived ease of use and academic utility are the strongest predictors of adoption. Ogunleye et al. (2024) synthesized 355 papers on generative AI for teaching and learning more broadly, identifying gaps around assessment design and curricular integration. Across all three reviews, structured pedagogical scaffolding is identified as the differentiator between studies reporting positive and null effects.

Foundational technical work has produced the architecture on which the contemporary art-education literature rests. Diffusion-based generative models (Ho et al., 2020; Rombach et al., 2022) and recent improvements in their training dynamics (Karras et al., 2024) made high-resolution prompt-conditioned synthesis tractable, while concerns about training-data provenance and fair use remain unresolved at the policy level (Lemley & Casey, 2021).

Procedural and metamodeling work in adjacent technical domains supplies a useful methodological precedent. Al-Dhaqm and colleagues (Al-Dhaqm et al., 2017, 2020) developed common-process models in database forensic investigation using design-science research and metamodeling: collect candidate processes, extract concepts, reconcile synonyms, and abstract a generalized model. The substantive content of art pedagogy is very different from forensic investigation; only the metamodeling procedure is borrowed here.

The human-AI co-creation literature offers principles for the Articulate and re-Fine stages of CRAFT. Foundational work on mixed-initiative co-creativity (Yannakakis et al., 2014) and participatory sense-making with co-creative agents (Davis et al., 2016) established that creative friction and iterative loops increase user agency. Among recent experimental studies, McGuire, De Cremer, and Van de Cruys (2024) found that interface design-whether the human is positioned as an editor or a co-creator-measurably affects creative outcomes (Studies 1 and 2, total $N \approx 200$, online poetry-writing task, Prolific Academic). The McGuire study is from an adjacent creative domain (text/poetry) rather than studio painting and is therefore cited here as suggestive rather than directly applicable evidence; the genuine art-and-design education base is reviewed in Section 5.2.

Classical foundations from art education complete the theoretical grounding: Eisner (2002) on artistic cognition; Dewey (1934) on art as experience; Schön (1987) on reflective practice; Hetland et al. (2013) on Studio Habits of Mind; Mishra and Koehler (2006) on TPACK for technology integration; and Anderson and Krathwohl (2001) on the revised Bloom taxonomy.

3. Methodology

The present study applies a Design Science Research (DSR) orientation (Anderson & Shattuck, 2012) to construct a conceptual pedagogical framework for the use of generative AI in painting and drawing. To define the problem space and identify design requirements, a critical integrative review of the recent empirical literature was conducted. A critical integrative review is a non-systematic synthesis that combines evidence from heterogeneous study designs in order to appraise, critique, and integrate a body of literature into a conceptual contribution, rather than to estimate a pooled effect (Torraco, 2005; Whittemore & Knafl, 2005). The review reported here is deliberately not a PRISMA-compliant systematic review: no protocol was preregistered, and dual independent screening of all candidate records was not performed (Section 5.4 returns to this limitation). For clarity, the single label “critical integrative review” is used throughout this paper to describe the re-

view design.

3.1. Search Strategy

The search targeted Web of Science, Scopus, ERIC, and PubMed/PMC for English-language articles published from January 2023 to April 2026, using combinations of (“generative AI” OR “text-to-image” OR “diffusion”) AND (“art education” OR “design education” OR “creative learning” OR “painting”) AND (“intervention” OR “experiment” OR “quasi-experimental” OR “pre-test” OR “effect size”). Hand-searches of three recent systematic reviews (Jiang et al., 2025; Wu et al., 2025; Ogunleye et al., 2024) supplemented the database search.

3.2. Inclusion and Exclusion Criteria

Candidate studies were screened against three inclusion criteria: (i) primary empirical data, quantitative or mixed methods; (ii) involvement of generative AI in a creative or art/design context; and (iii) at least one outcome measure relevant to one of the CRAFT stages-ideation, prompting, critique, iteration, or curation/publication. Purely conceptual papers, opinion pieces, and bibliometric studies without extractable data were excluded. Secondary reviews were not eligible as primary studies; three systematic reviews (Jiang et al., 2025; Wu et al., 2025; Ogunleye et al., 2024) were nonetheless retained separately as secondary sources used to frame, but not to populate, the per-stage evidence map. Where a single article reported more than one independent experiment, each experiment was treated as a separate primary dataset. A review of microteaching in teacher education (Iliasova et al., 2025) does not satisfy criteria (i) or (ii) and is therefore excluded from the primary corpus; it is cited only as adjacent, non-studio evidence in Section 5.2.3.

3.3. Study Selection

The database searches returned 287 records, and hand-searching the reference lists of the three systematic reviews added a further 24, giving 311 records in total. After 63 duplicates were removed, 248 unique records were screened on title and abstract, of which 201 were excluded as off-topic (no generative AI, no creative or art/design context, or no CRAFT-relevant outcome). The remaining 47 records were assessed in full text; 37 were excluded, comprising 28 that lacked primary empirical data (conceptual, opinion, or bibliometric pieces), 6 that reported no outcome mappable to a CRAFT stage, and 3 that fell outside a creative or art/design context. Ten primary studies met all inclusion criteria. Because McGuire et al. (2024) reports two independent experiments (Study 1 and Study 2), these are treated as two primary datasets; the corpus therefore comprises eleven datasets drawn from ten articles. The three retained systematic reviews are used as secondary framing sources and are not counted among the primary studies.

3.4. Data Extraction and Synthesis

From each included primary study the following items were extracted: setting and

population, sample size, study design, key reported outcome statistics (means, standard deviations, $t/F/\beta$ values, p -values, and reported effect sizes where available), and the dominant CRAFT stage(s) the study informs. Mapping of studies to CRAFT stages was performed by the sole author and is presented as an interpretive single-coder synthesis; no second coder, codebook, or inter-rater reliability check was used. This is a significant source of bias and is acknowledged here and in Section 5.4. Future evaluations of CRAFT should employ at least two independent coders mapping primary studies against an explicit codebook. A quantitative meta-analysis was not performed because the included studies are too heterogeneous in outcome measures to support a common effect-size metric; instead, the evidence was combined narratively, with a per-stage evidence-mapping table (Table 1 and Table 2).

3.5. Framework Construction Procedure

CRAFT was constructed using a metamodeling procedure adapted from Al-Dhaqm et al. (2017, 2020): a) collection of candidate process descriptions from published syllabi and the included primary studies; b) concept extraction; c) reconciliation of synonymous concepts; and d) abstraction into a generalized model. This procedure was chosen rather than closer education-research equivalents (framework synthesis or thematic synthesis) because its concept-reconciliation step is unusually explicit, which is helpful when working with heterogeneous instructor and researcher vocabulary. Five canonical stages emerged: Conceive, Render, Articulate, reFine, Transmit. The pipeline is illustrated in Figure 1, including the iteration loop from reFine back to Render or Conceive.

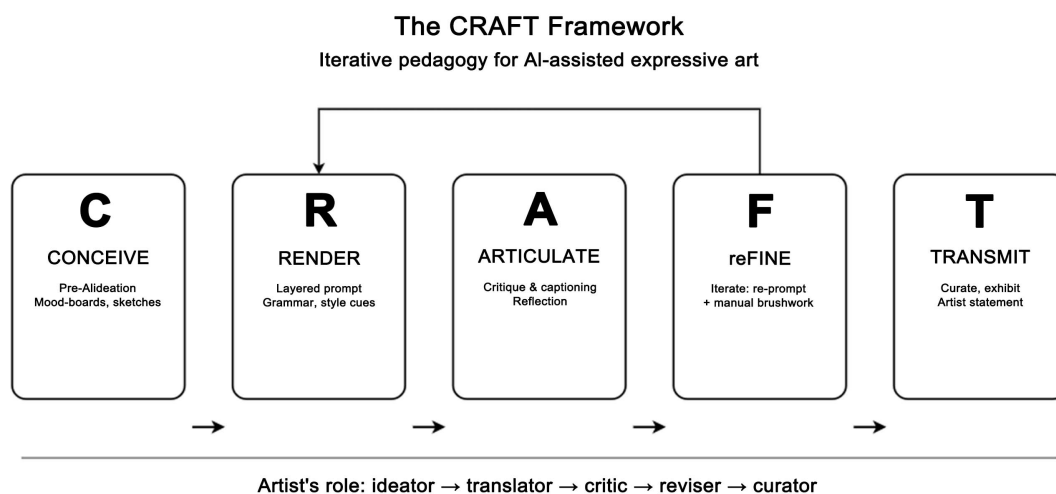


Figure 1. The CRAFT pipeline.

4. The Proposed CRAFT Framework

The process of creating an art project through CRAFT consists of five interrelated steps. All steps are shown in Table 1, which provides information about the ped-

agogical goal, student action, and AI role at each step of the process.

Table 1. The five CRAFT stages with pedagogical focus, student activity, and AI role.

CRAFT Stage	Pedagogical Focus	Student Activity	AI Role
Conceive	Concept ideation and emotional intent	Mood-boards, journaling, sketching	Brainstorm partner via prompts
Render	Translation of concept into visual prompt	Prompt engineering, style cues	Image generator (diffusion)
Articulate	Critical reflection on output	Group critique, written reflection	Captioning, semantic feedback
reFine	Iterative revision and over-painting	Re-prompting, manual brushwork	In-painting, image-to-image
Transmit	Curation, exhibition, sharing	Digital portfolio, gallery, talk	Documentation assistant

4.1. Conceive-Pre-AI Ideation

This stage sets the emotional and intellectual tone for what follows. Students engage in mood-boarding, journaling, and thumbnail sketching without using any AI tools. The working pedagogical assumption behind this stage-treated here as a hypothesis rather than an established finding-is that protecting students' initial creative intent before AI tools are introduced may support more authentic ideation; the indirect evidence bearing on this assumption is discussed in Section 5.2.1.

4.2. Render-Layered Prompt Engineering

The output of Conceive is then converted into a stratified visual prompt that conveys mood, object, and style reference. Students learn prompt grammar (role, context, instruction, format) explicitly, as modelled by the teacher. AI operates as the picture creator. The Render stage rests on a small but growing corpus of scholarship on the teaching of prompt engineering (Park & Choo, 2025; Liu & Chilton, 2022).

4.3. Articulate-Critical Reflection on AI Outputs

The generated outputs are subjected to an organized process of critique that includes feedback from peers, the instructor, and the AI acting as captioner. Reflection is the dynamic part of the process that prevents superficial acceptance or rejection of AI-produced outputs. The concept draws on the practice of reflection in both design education and wider practice theory.

4.4. reFine-Iterative Re-Prompting and Manual Brushwork

Insights gained from articulation inform either a fresh start in Render (with changed parameters) or in Conceive (changing the emotion to convey). The re-Fine stage combines re-prompting with brushwork, positioning the human as an equal partner in the iterative process rather than as an editor of the resulting AI

product. Within CRAFT, reFine forms the framework's core commitment; however, evidence from drawing/painting labs is scarce, and the relevant data are mostly borrowed from neighbouring creativity tasks (Section 5.2).

4.5. Transmit-Curation, Statement-Writing, and Public Sharing

For closure, the process has students select certain works to curate, write artist statements, and distribute their outputs through a public exhibit or portfolio. The role of AI in this stage is documentation. As currently formulated, Transmit receives the least attention in the literature.

5. Results and Discussion

5.1. Overview of the Included Evidence Base

The ten included primary studies (eleven datasets) are summarized in **Table 2**. Among the studies that report a single participant count, sample sizes range from $N = 10$ (Anson et al., 2025) to $N = 405$ (Zhuang & Li, 2025); several studies use cohort-, class-, or workshop-level designs for which **Table 2** does not list a single headline N (Hwang & Wu, 2025b; Sáez-Velasco et al., 2024; Chiu & Hwang, 2025; Chen et al., 2025). Settings span K-12 (Vartiainen et al., 2025), higher-education studios and design courses (Hwang & Wu, 2025a, 2025b; Hiçiyilmaz, 2025; Anson et al., 2025; Sáez-Velasco et al., 2024; Chiu & Hwang, 2025; Chen et al., 2025), and an online experimental platform (McGuire et al., 2024). Designs include two-condition experiments (McGuire et al., 2024, Studies 1 and 2), pre/post quasi-experiments (Vartiainen et al., 2025; Chen et al., 2025), cross-sectional surveys with mediation analysis (Hwang & Wu, 2025a), structural equation models (Zhuang & Li, 2025), and small-sample qualitative studies (Hiçiyilmaz, 2025; Anson et al., 2025). Reported effect-size families differ across studies (Cohen's d , Hedges' g , β coefficients, mean differences); this heterogeneity, together with the single-coder mapping noted in Section 3.4, means the synthesis is interpretive and precludes pooled meta-analysis.

Table 2. Included primary studies with reported sample sizes, designs, and key findings. References are listed in full in the bibliography.

Study	Setting	N	Design	Key reported finding
McGuire et al. (2024), Study 1	Online; poetry writing	101	Solo vs. AI-edit	Solo writers rated more creative than AI-editors by professional poets.
McGuire et al. (2024), Study 2	Online; poetry writing	101	Two-condition (editor vs. co-creator)	Co-creator higher on creative self-efficacy ($M = 4.62$ vs. 3.74 , $p = .003$); expert creativity ratings 14.70 vs. 12.53 , $p = .026$.
Hwang & Wu (2025a)	Design students, S. China	121	Cross-sectional survey + mediation	AI use $\rightarrow$ innovative thinking $\beta = .610$, $p < .001$; mediated by self-efficacy and anxiety reduction.

Continued

Hwang & Wu (2025b)	Graphic design course, Korea	Workshop cohort	Case-study workshop (TPACK-aligned)	Comprehensive understanding of text/content/context shaped image outcomes.
Zhuang & Li (2025)	Music collaboration, China	405	Mixed (SEM + experimental)	GAI collaboration → creative interest $\beta = .616$; self-efficacy $\beta = .557$; perceived competence $\beta = .357$.
Vartiainen et al. (2025)	Grade 4 & 7, Finland	209	Pre/post written-reasoning + lesson	Significant gains in data-driven explanations of algorithmic bias.
Hiçyilmaz (2025)	Pre-service art teachers, Türkiye	26	Qualitative case study	AI broadened compositional ideation; participants reported execution frustration.
Anson et al. (2025)	Bachelor art students, Latvia	10	Observation + questionnaire	Students valued AI for ideation, found it limited for nuanced execution.
Sáez-Velasco et al. (2024)	Higher-ed arts, Spain	Mixed cohort	Cross-disciplinary survey	AI raised participation and motivation; depended on instructor framing.
Chiu & Hwang (2025)	Higher ed, Taiwan	Two-class quasi-exp.	Mind-mapping × GAI vs. control	Higher creative and critical thinking scores in mind-mapping + GAI condition.
Chen et al. (2025)	Design & art course, China	Two-cohort	Quasi-experimental	Higher achievement, motivation, self-efficacy in GAI condition.

5.2. Stage-Level Findings

5.2.1. Conceive

Direct studio-painting evidence for this stage is limited and largely qualitative; the strongest quantitative evidence comes from design-education surveys, and the most rigorous experimental evidence comes from an adjacent (online text) task. These sources should not be read as equivalent to painting-studio evidence. Within art-and-design education, the strongest direct evidence supporting Conceive comes from Hwang and Wu (2025a): in $N = 121$ design students at universities in southern China, AI knowledge predicted innovative thinking ($\beta = .610$, $p < .001$), with creative self-efficacy as a significant mediator (standardized $\beta = .256$, 95% CI [.140, .418]) and anxiety reduction as a serial second mediator. A companion case-study workshop by the same authors (Hwang & Wu, 2025b) showed that students who developed a layered understanding of text, content, and context produced higher-quality posters using DALL-E and Midjourney. Chen et al. (2025) extend this in a quasi-experimental design-and-art course: the GAI condition outperformed the control on student achievement, motivation, and self-efficacy. Counter-evidence comes from Hiçyilmaz (2025) ($N = 26$ pre-service art teachers in Türkiye) and Anson et al. (2025) ($N = 10$ bachelor art students in Latvia): in both, students valued AI for ideation but reported frustration when

asked to execute nuanced artistic tasks without explicit pre-prompt scaffolding. Adjacent-domain evidence (McGuire et al., 2024)—an online text-based co-creation experiment—is consistent with the same general pattern but is not from a studio art setting and is therefore cited as suggestive rather than as direct studio support. The convergent reading is that protecting authentic creative intent before tool use may be associated with better outcomes, with the strongest evidence coming from design education and the most rigorous experimental design from outside the studio. **Figure 2** illustrates the kind of expressive subject treatment the Conceive and Render stages aim to support.



Figure 2. The kind of expressive subject treatment the Conceive and Render stages aim to support.

5.2.2. Render

The evidence for Render is drawn from design and graphic-design coursework rather than from drawing/painting studios specifically; it is adjacent but remains within art and design. Hwang and Wu (2025b) found that students who developed a layered understanding of text, content, and context produced higher-quality posters when using DALL-E and Midjourney for graphic-design coursework. Park and Choo (2025) document that explicit teaching of prompt structure (role, context, instruction, format) reliably improves output alignment in educational settings. However, Ansone et al. (2025) (N = 10) and Hiçyilmaz (2025) (N = 26) both report that students who used AI without prompt scaffolding found the tools frustrating for executing nuanced artistic tasks. The combined evidence supports CRAFT's Render stage but emphasizes that prompt instruction must be explicit and modelled.

5.2.3. Articulate

Direct studio-painting evidence for reflective critique is limited; the evidence be-

low is drawn mainly from adjacent domains (K-12 lessons on AI, higher-education design discussion, and teacher education) and should not be equated with painting-studio critique. [Vartiainen et al. \(2025\)](#) (N = 209, Grades 4 and 7 in Finland) showed that a single workshop on algorithmic bias produced significant gains in students' written reasoning about generative-model outputs. A companion study with K-9 learners co-creating digital art ([Vartiainen et al., 2023](#)) reached similar conclusions about the importance of structured reflection. At the higher-education level, [Sáez-Velasco et al. \(2024\)](#) reported increases in participation and motivation when AI integration was paired with structured discussion, and [Fleischmann \(2024\)](#) documented how online studio-critique protocols externalize tacit knowledge in design education. Adjacent teacher-education evidence ([Iliasova et al., 2025](#), a review of microteaching outside studio art) converges on the same point: reflective scaffolding is the active ingredient in skill-development interventions. The most informative counter-evidence is summarized in [Jiang et al. \(2025\)](#): across the reviewed corpus, short AI-supported interventions can produce gains in objective performance not matched by gains in self-efficacy. CRAFT's Articulate stage is designed to close this gap.

5.2.4. reFine

This is the stage with the least direct studio-painting evidence; most support comes from adjacent creative domains (poetry, music) and should be read as suggestive only. Direct studio-art evidence on iterative reFine is sparse but instructive: [Hiçyilmaz \(2025\)](#) and [Ansone et al. \(2025\)](#) both report that art students value AI for ideation but find iteration without scaffolding frustrating, with the execution of nuanced artistic effects remaining stubbornly hard. [Chen et al. \(2025\)](#) report converging quasi-experimental gains in achievement, motivation, and self-efficacy when iterative GAI use was structured into a design-and-art course. Experimental data from adjacent domains (poetry: [McGuire et al., 2024](#); music co-creation: [Zhuang & Li, 2025](#)) support the general claim that human-in-the-loop iteration tends to be rated more favourably than one-shot generation, yet neither case represents a painting-studio test. The reFine phase of CRAFT implements the facilitative condition-re-prompting combined with manual painting-as the type of iteration most likely to transfer to studio work. This suggested transfer is a conjecture produced by the model and not a proven result at this point. [Figure 3](#) shows the reFine cycle in diagrammatic form: a rough sketch on the left is iteratively transformed into an expressive painting by hand on the right. [Figure 4](#) illustrates the kind of memory-laden content that pilot lessons in this phase can aim at.

5.2.5. Transmit

No direct studio-painting study specifically tests this stage; the two supporting studies are situated in higher-education art/design courses more broadly. Direct empirical evidence for the Transmit stage is the thinnest in the corpus, reflecting the broader assessment-design gap noted across the systematic-review literature

(Jiang et al., 2025; Ogunleye et al., 2024). Sáez-Velasco et al. (2024) document motivation gains tied to public sharing of AI-assisted work, and Chiu and Hwang (2025) report critical-thinking gains in a mind-mapping plus GAI condition culminating in a presentation phase. The Transmit stage, as currently formulated, is therefore the least directly supported and is the priority target for future research.



Figure 3. The reFine loop, shown schematically.



Figure 4. The kind of memory-laden subject matter that pilot teaching exercises in this stage might target.

5.3. Comparison with Existing Frameworks

Despite the importance of pedagogical structure, much of the available literature on generative AI in art education is either descriptive of the technology or prescriptive about general digital literacy, with comparatively little operational specification for the studio. CRAFT proposes integrating these strands into a holistic concept that takes an explicitly stage-by-stage, evidence-mapped path. **Table 3** compares CRAFT with the dominant existing accounts.

Table 3. CRAFT compared with existing accounts of generative AI in art and design education.

Aspect of Comparison	Existing Accounts	CRAFT Framework
Primary orientation	Either descriptive (technical mechanics) or prescriptive (general digital literacy); rarely both. Limited operational specification.	Integrative and operational. Five stages, each with an explicit pedagogical focus, student activity, and AI role.
Methodology	Often implicit. Position pieces and ad-hoc theoretical contributions without unified review or design methodology.	Combines a critical integrative review of recent empirical primary studies with a metamodeling-inspired framework construction.
Scope of evidence	Typically a single pedagogical context or a small number of model architectures.	Synthesizes ten primary studies (eleven datasets) and three systematic reviews across K-12 and higher education, 2023-2026.
Treatment of artistic process	Mostly absent or treated as a single “AI use” block.	Decomposed into five stages (Conceive, Render, Articulate, reFine, Transmit) with stage-specific evidence mapping.
Educational implications	Discussed only in a small subset of the literature, usually for general AI literacy.	Integrated throughout, with stage-level guidance for studio instructors and an explicit Articulate stage for reflection.
Empirical claims	Variable; some accounts report perception studies, others remain purely conceptual.	Honest. The framework does not introduce new experimental data; it synthesizes existing evidence and identifies concrete next empirical steps.

5.4. Practical Implementation Challenges and Limitations

CRAFT and its synthesis face several limitations. First, this is a critical integrative review rather than a full PRISMA-compliant systematic review; no protocol was preregistered, and dual independent screening of all candidate records was not performed. Second, the ten primary studies use heterogeneous outcome measures, precluding pooled meta-analysis. Third, the studies are concentrated in higher education and East Asian contexts (Jiang et al., 2025), limiting generalizability, and equity considerations (Holstein & Doroudi, 2022) are not addressed uniformly. Fourth, the strongest experimental evidence cited (McGuire et al., 2024) is from an online text/poetry co-creation task with non-artist Prolific participants, not a studio drawing or painting study; the genuine art-and-design education evidence base, while real, is dominated by smaller and more qualitative studies (Hiçyilmaz, 2025; Ansone et al., 2025). The synthesis is therefore stronger on whether CRAFT’s component mechanisms have any empirical support than on whether they have been demonstrated specifically in studio painting and drawing. Fifth, none of the included studies tests CRAFT directly; the synthesis evaluates the framework’s component mechanisms, not the framework as a whole. Sixth, AI tools evolve rapidly: improvements in diffusion-model training (Karras et al., 2024) and unresolved fair-use questions around training data (Lemley & Casey, 2021) mean that 2023-2024 model-generation findings may not transfer cleanly

to 2025-2026 generations. Seventh, the framework presumes equitable access to GPUs, stable internet, and licensed software-conditions not met in many art programs (Selwyn, 2024).

5.5. Clear Steps for Future Empirical Work

To validate CRAFT and move it from concept toward field use, six concrete next steps are proposed. (A) A direct quasi-experimental evaluation of CRAFT against an active control, with sufficient power for medium effects ($n \geq 25$ per group), external blind raters scoring artifacts on a rubric that excludes intervention-confounded constructs, fidelity-of-implementation measurement, and pre-registration of the analysis plan. (B) Multi-site replication across at least two institutional contexts, to address the East-Asia and higher-education skew documented in Jiang et al. (2025). (C) Targeted Render-stage research on prompt-grammar instruction grounded in the design guidelines of Liu and Chilton (2022). (D) Articulate-stage research on reflective scaffolds adapted from Chaseley and Abercrombie (2025). (E) Transmit-stage research developing curation rubrics for AI-assisted artworks, drawing on the wider AI-art review literature (Cetinic & She, 2022). (F) Face-validation of the framework by domain experts and cognitive-taxonomy alignment of CRAFT outcomes with revised Bloom levels (Anderson & Krathwohl, 2001).

6. Conclusion

This paper presented CRAFT, a five-stage pedagogical framework for teaching expressive drawing and painting with generative AI, and evaluated it against ten primary studies (eleven datasets) and three systematic reviews from 2023-2026. The framework's stage-level mechanisms are partly supported by the recent art-and-design education literature: prompt-engineering literacy and AI knowledge predict innovative thinking and creative self-efficacy in design students; structured AI-assisted instruction is associated with higher achievement and motivation in design and art courses; and structured reflection may help mitigate the competence-confidence dissociation documented across the broader review literature. Throughout, direct studio painting-and-drawing evidence should be distinguished from adjacent-domain evidence (design surveys, online poetry, music co-creation, K-12, and teacher education); the latter motivates CRAFT's mechanisms but does not establish them for the painting studio. Direct studio-art evidence on iterative reFine remains thin, with most experimental support coming from adjacent creative tasks. The Transmit stage is the least empirically supported and is the priority target for future work. CRAFT has not been directly evaluated; the synthesis presented here motivates such an evaluation but does not substitute for it. Replicating at additional studio sites, with adequately powered samples, no-AI control conditions, and current model generations, is the most pressing next step. CRAFT is recommended as a candidate scaffold for studio instructors integrating generative AI into their teaching, with the modest empirical claims this synthesis supports.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Al-Dhaqm, A., Razak, S., Othman, S. H., Ali, A., Ghaleb, F. A., Rosman, A. S., & Marni, N. (2020). Categorization and Organization of Database Forensic Investigation Processes. *IEEE Access*, 8, 112846-112858. <https://doi.org/10.1109/access.2020.3000747>
- Al-Dhaqm, A., Razak, S., Othman, S. H., Choo, K. R., Glisson, W. B., Ali, A. et al. (2017). CDBFIP: Common Database Forensic Investigation Processes for Internet of Things. *IEEE Access*, 5, 24401-24416. <https://doi.org/10.1109/access.2017.2762693>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman.
- Anderson, T., & Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, 41, 16-25. <https://doi.org/10.3102/0013189x11428813>
- Ansone, A., Zālīte-Supe, Z., & Daniela, L. (2025). Generative Artificial Intelligence as a Catalyst for Change in Higher Education Art Study Programs. *Computers*, 14, Article 154. <https://doi.org/10.3390/computers14040154>
- Boden, M.A. (1998) Creativity and Artificial Intelligence. *Artificial Intelligence*, 103, 347-356. [https://doi.org/10.1016/s0004-3702\(98\)00055-1](https://doi.org/10.1016/s0004-3702(98)00055-1)
- Cetinic, E., & She, J. (2022). Understanding and Creating Art with AI: Review and Outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 18, 1-22. <https://doi.org/10.1145/3475799>
- Chaseley, T. L., & Abercrombie, S. (2025). Using Scaffolds to Support Preservice Teachers' Reflective Practice. *Frontiers in Education*, 10, Article ID: 1621269. <https://doi.org/10.3389/feduc.2025.1621269>
- Chen, J., Mokmin, N. A. M., & Su, H. (2025). Integrating Generative Artificial Intelligence into Design and Art Course: Effects on Student Achievement, Motivation, and Self-Efficacy. *Innovations in Education and Teaching International*, 62, 1431-1446. <https://doi.org/10.1080/14703297.2025.2503857>
- Chiu, M.-C., & Hwang, G.-J. (2025). Enhancing Student Creative and Critical Thinking in Generative Ai-Empowered Creation: A Mind-Mapping Approach. *Interactive Learning Environments*, 34, 869-890. <https://doi.org/10.1080/10494820.2025.2511244>
- Davis, N., Hsiao, C., Yashraj Singh, K., Li, L., & Magerko, B. (2016). Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. In J. Nichols, J. Mahmud, J. O'Donovan, C. Conati, & M. Zancanaro (Eds.), *Proceedings of the 21st International Conference on Intelligent User Interfaces* (pp. 196-207). ACM. <https://doi.org/10.1145/2856767.2856795>
- Dewey, J. (1934). *Art as Experience*. Minton, Balch & Company.
- Eisner, E. W. (2002). *The Arts and the Creation of Mind*. Yale University Press.
- Fleischmann, K. (2024). Making Tacit Knowledge Explicit: The Case for Online Peer Feedback in the Studio Critique. *International Journal of Technology and Design Education*, 35, 699-721. <https://doi.org/10.1007/s10798-024-09911-8>
- Hertzmann, A. (2020). Computers Do Not Make Art, People Do. *Communications of the ACM*, 63, 45-48. <https://doi.org/10.1145/3347092>
- Hetland, L., Winner, E., Veenema, S., & Sheridan, K. M. (2013). *Studio Thinking 2: The*

- Real Benefits of Visual Arts Education* (2nd ed.). Teachers College Press.
- Hiçyılmaz, Y. (2025). An Innovative Approach in Arts Education: Student Experiences of Abstract Art Practices Supported by Generative Artificial Intelligence. *SAGE Open*, 15, 1-19.
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*, 33, 6840-6851.
- Holstein, K. and Doroudi, S. (2022) Equity and Artificial Intelligence in Education. In *The Ethics of Artificial Intelligence in Education*, Routledge, 151-173.
<https://doi.org/10.4324/9780429329067-9>
- Hwang, Y., & Wu, Y. (2025a). The Influence of Generative Artificial Intelligence on Creative Cognition of Design Students: A Chain Mediation Model of Self-Efficacy and Anxiety. *Frontiers in Psychology*, 15, Article ID: 1455015.
<https://doi.org/10.3389/fpsyg.2024.1455015>
- Hwang, Y., & Wu, Y. (2025b). Graphic Design Education in the Era of Text-to-Image Generation: Transitioning to Contents Creator. *International Journal of Art & Design Education*, 44, 239-253. <https://doi.org/10.1111/jade.12558>
- Iliasova, L., Nekrasova, I., Mena, J., & Estrada-Molina, O. (2025). Microteaching on Pre-Service Teachers' Education: Literature Review. *Frontiers in Education*, 10, Article ID: 1562975. <https://doi.org/10.3389/educ.2025.1562975>
- Jiang, Y., Fan, Y., & Liu, Z. (2025). Generative AI in Art Education: A Systematic Review of Research Trends, Tool Applications, and Outcomes (2019-2025). *Education Sciences*, 16, Article 47. <https://doi.org/10.3390/educsci16010047>
- Karras, T., Aittala, M., Lehtinen, J., Hellsten, J., Aila, T., & Laine, S. (2024). Analyzing and Improving the Training Dynamics of Diffusion Models. In *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 24174-24184). IEEE. <https://doi.org/10.1109/cvpr52733.2024.02282>
- Lemley, M. A., & Casey, B. (2021). Fair Learning. *Texas Law Review*, 99, 743-785.
- Liu, V., & Chilton, L. B. (2022). Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In S. Barbosa, P. Lampe, P. Appert et al. (Eds.), *CHI Conference on Human Factors in Computing Systems* (pp. 1-23). ACM.
<https://doi.org/10.1145/3491102.3501825>
- Manovich, L. (2018). *AI Aesthetics*. Strelka Press.
- Manovich, L., & Arielli, E. (2024). *Artificial Aesthetics: Generative AI, Art and Visual Media*. <https://manovich.net/index.php/projects/artificial-aesthetics>
- McGuire, J., De Cremer, D., & Van de Cruys, T. (2024). Establishing the Importance of Co-Creation and Self-Efficacy in Creative Collaboration with Artificial Intelligence. *Scientific Reports*, 14, Article No. 18525. <https://doi.org/10.1038/s41598-024-69423-2>
- Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record: The Voice of Scholarship in Education*, 108, 1017-1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A Systematic Review of Generative AI for Teaching and Learning Practice. *Education Sciences*, 14, Article 636. <https://doi.org/10.3390/educsci14060636>
- Park, J., & Choo, S. (2025). Generative AI Prompt Engineering for Educators: Practical Strategies. *Journal of Special Education Technology*, 40, 411-417.
<https://doi.org/10.1177/01626434241298954>
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution

- Image Synthesis with Latent Diffusion Models. In *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 10674-10685). IEEE. <https://doi.org/10.1109/cvpr52688.2022.01042>
- Sáez-Velasco, S., Alaguero-Rodríguez, M., Delgado-Benito, V., & Rodríguez-Cano, S. (2024). Analysing the Impact of Generative AI in Arts Education: A Cross-Disciplinary Perspective of Educators and Students in Higher Education. *Informatics, 11*, Article 37. <https://doi.org/10.3390/informatics11020037>
- Schön, D. A. (1987). *Educating the Reflective Practitioner*. Jossey-Bass.
- Selwyn, N. (2024). On the Limits of Artificial Intelligence (AI) in Education. *Nordic Journal of Studies in Educational Policy, 10*, 3-14. <https://doi.org/10.23865/ntp.v10.6062>
- Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review, 4*, 356-367. <https://doi.org/10.1177/1534484305278283>
- Tour, E., & Zadorozhnyy, A. (2025). Conceptualizing and Operationalizing Prompt Literacy for English Language Learners. *Journal of Adolescent & Adult Literacy, 69*, e70020. <https://doi.org/10.1002/jaal.70020>
- Vartiainen, H., Kahila, J., Tedre, M., López-Pernas, S., & Pope, N. (2025). Enhancing Children's Understanding of Algorithmic Biases in and with Text-to-Image Generative Ai. *New Media & Society, 27*, 5342-5368. <https://doi.org/10.1177/14614448241252820>
- Vartiainen, H., Tedre, M., & Jormanainen, I. (2023). Co-Creating Digital Art with Generative AI in K-9 Education: Socio-Material Insights. *International Journal of Education Through Art, 19*, 405-423. https://doi.org/10.1386/eta_00143_1
- Whittemore, R., & Knafl, K. (2005). The Integrative Review: Updated Methodology. *Journal of Advanced Nursing, 52*, 546-553. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>
- Wu, F., Dang, Y., & Li, M. (2025). A Systematic Review of Responses, Attitudes, and Utilization Behaviors on Generative AI for Teaching and Learning in Higher Education. *Behavioral Sciences, 15*, Article 467. <https://doi.org/10.3390/bs15040467>
- Yannakakis, G. N., Liapis, A., & Alexopoulos, C. (2014). Mixed-Initiative Co-Creativity. In M. Mateas, T. Barnes, & I. Bogost (Eds.), *Proceedings of the 9th International Conference on the Foundations of Digital Games (FDG 2014)*. Society for the Advancement of the Science of Digital Games.
- Zhuang, Z., & Li, X. (2025). The Influence of Collaborative Music Creation Supported by Generative Artificial Intelligence on Students' Creativity. *Frontiers in Psychology, 16*, Article ID: 1709513. <https://doi.org/10.3389/fpsyg.2025.1709513>