Innovation Series https://www.innoseries.com

A Systematic Review of Aesthetic Judgment and Creativity Cognition in AI Visual Art

Haibo Gao
City University of Macau, Zhuhai, Guangdong, China

Abstract: Objective: With generative AI technologies such as diffusion models increasingly shaping the production and circulation of visual art, the question of how AI-generated works acquire identity and value within artistic contexts has become urgent and calls for a systematic response. This study conducts a systematic literature review focusing on (1) how the aesthetic status of AI visual art is conceptualized, (2) evaluative differences between human- and AI-produced works, and (3) the mechanisms underlying aesthetic bias.
Methods: Following the PRISMA 2020 guidelines, we searched Scopus and Web of Science for English-language publications from 2016 to 2026. After merging records, removing duplicates, and screening according to predefined inclusion and exclusion criteria, 55 journal articles were included. We coded and synthesized research orientation, artwork samples and the information provided in their presentation, evaluation dimensions, experimental paradigms, and primary findings using a thematic approach.
Results: First, at the level of value justification, the eligibility of AI visual art as an aesthetic object is largely framed as context-dependent, hinging on the attribution and traceability of responsibility in the creative process, the explainability of agency, and recognition by artistic institutions and cultural frameworks. Second, at the level of empirical comparison, "source information" exerts a relatively stable influence on evaluations and affects perceptions of artistry, creativity, and authenticity more strongly than it affects ratings of beauty. Third, at the level of mechanisms, insufficient attribution of creative agency and intention, essentialist beliefs about creator uniqueness, culturally embedded evaluative frameworks, and the joint effects of data and algorithms on the visible form of artworks lead to systematic downward adjustments in perceived artistry and creativity of AI works, with effect sizes varying by task context and detection methods.
Contributions: Using an "acceptance value–creative agency" analytical framework, this review stratifies normative arguments and empirical evidence, clarifies cross-study differences in evaluation metrics, artwork sampling and presentation, and task design, and proposes testable recommendations to improve cross-study and cross-cultural comparability.

Keywords: AI art; Generative AI; Visual art; Aesthetic judgment; Creativity cognition; Systematic literature review

1. Introduction

Driven by advances in generative modeling, generative AI has rapidly entered the production and circulation of visual art, making the question of "how visual works produced by generative AI acquire identity and value within an art context" a theoretical and empirical issue that requires systematic response. This shift has also brought the relationship among "art–intelligence–creativity" back to the center of scholarly debate, prompting researchers to examine the value structure and evaluative logics that emerge when AI intervenes in creative practice from the perspectives of both art philosophy and techno-cultural critique [1].

Current research largely revolves around intention, authorship, and agency/subjecthood. One stream, grounded in art philosophy and media theory, discusses how the attribution structure of a work changes when generative systems enter the creative workflow, and whether AI is understood as a tool, a co-creator, or a quasi-subject—thereby shaping the aesthetic status of AI art and the justificatory basis for evaluating it as an aesthetic object [2].

Related debates further address how data and models reorganize style and generate "paradigmatic" effects, suggesting that AI art is not a neutral technical output but is produced under the constraints and supports of existing cultural resources and institutional frameworks, and is subsequently interpreted through processes of criticism and reception.

Meanwhile, the critical tradition's distinction regarding "reception value" indicates that evaluation is not merely about audience preference; it also concerns whether a work realizes its intentions and value claims, and how creative subjecthood is presented and recognized in the work. In the AI context, evaluation therefore often co-occurs with a hierarchical framework of creative agency: from human-led creation to human–AI collaboration, and further toward stronger system-led production—different agency configurations invite different ways of attributing artistic value and creativity. Empirical studies based on viewers' evaluations test comparative judgments of AI versus human works across dimensions such as beauty, artistic value, and creativity by manipulating "provenance/creation information" while controlling for work type and presentation format, accumulating relatively consistent findings: source cues can significantly shift evaluations, and their effects on attributing creativity and artistic value are often more robust [8]. At the mechanistic level, research further emphasizes that viewers' attribution of creative agency and intention, the cultural knowledge and evaluative norms on which such attributions rely, and the specific evaluative context jointly shape judgments of creativity and artistic value for the same work—thereby explaining systematic differences under AI versus human authorship labels. In parallel, computational aesthetics and neuroaesthetics, starting from artwork features, processing mechanisms, and evaluation indices, provide quantifiable detection methods and model-analytic approaches for understanding "which visible cues are taken as evidence of skill, intention, and novelty" [3].

Although the relevant evidence base has continued to grow, the existing literature still exhibits three salient limitations. First, normative discussions on the aesthetic status of AI visual art have not been systematically aligned with the evaluative dimensions and detection indicators adopted in empirical research, resulting in a weak and unsystematic correspondence between theory and evidence. Second, explanations of aesthetic bias and creativity skepticism largely remain at the level of isolated predictors; a coherent mechanism chain integrating cognition, cultural value orientations, and algorithmic conditions is still underdeveloped and insufficiently tested empirically. Third, substantial heterogeneity across studies—in artwork sample types and composition (including presentation information such as source labels), model types, and experimental paradigms—undermines cross-study comparability and constrains the generalizability of conclusions.

Against this backdrop, this systematic literature review addresses three research questions:

RQ1: How is the aesthetic status of AI art as an aesthetic object conceptualized and defined?

RQ2: How do people evaluate AI-generated art across the dimensions of beauty, artistic value, and creativity, and how do these evaluations differ from those of human-made works?

RQ3: What mechanisms shape aesthetic bias and creativity skepticism toward AI art?

This review contributes by integrating interdisciplinary evidence through an "acceptance value–creative agency" analytical framework, developing a traceable mapping of research paradigms and indicators, and, on this basis, identifying key boundary conditions and promising directions for future empirical tests.

The remainder of the article is organized as follows: Methods (search and screening), Results, Discussion, and Conclusion.

2. Methods

2.1 Search Strategy

This review followed the PRISMA framework to identify, screen, and include eligible studies, and reported the search and exclusion process transparently in accordance with PRISMA guidance (Page et al., 2021). The search period was restricted to 2016–2026, and two databases were selected: Scopus and Web of Science (WoS). The search query comprised four sets of keywords—terms related to (1) AI technologies, (2) visual art, (3) aesthetics and value, and (4) creativity/perception/evaluation—combined using Boolean operators. The full search string was:

(("Artificial Intelligence" OR "AI" OR "Generative AI" OR "GenAI" OR "Deep Learning" OR "Machine Learning" OR "Neural Network*" OR "GAN" OR "Algorithm*") AND ("Art" OR "Arts" OR "Artwork*" OR "Visual Art*" OR "Generative Art") AND ("Aesthetic*" OR "Beauty" OR "Artistic Value") AND ("Creat*" OR "Percept*" OR "Judg*" OR "Evaluat*" OR "Appreciat*" OR "Recept*"))。

To enhance cross-study comparability, after merging and deduplicating records, we further restricted the corpus to English-language publications.

2.2 Inclusion and Exclusion Criteria

The inclusion criteria were as follows:

1) The study addressed AI/generative AI in relation to visual artworks or visual-art contexts;

2) The study directly discussed or assessed aesthetic judgment (e.g., beauty, aesthetic value, artistic value) and/or creativity cognition (e.g., creativity evaluation, originality, attribution of intention/agency);

3) The study was an academic research report with a traceable methodology and evidentiary chain, conducted within a peer-reviewed context (with subsequent analyses focusing primarily on journal articles).

The exclusion criteria were as follows:

1) The study primarily focused on algorithm/model performance or output-quality optimization, without examining aesthetic judgment or creativity cognition in an art context;

2) The target object did not fall within visual art (or only broadly discussed "creative technologies" without visual-art evidence);

3) The study centered on law, copyright, or ethical governance rather than aesthetic and creativity evaluation;

4) The study focused on application scenarios but lacked empirical or theoretical engagement with mechanisms of "artistic value/creativity" judgment;

5) The work was mainly technological narrative and lacked replicable methods and verifiable

evidence.

2.3 PRISMA Screening Process

Screening was organized according to the four PRISMA stages: Identification, Screening, Eligibility, and Inclusion. The procedure was as follows:

1) Identification: Records were retrieved separately from Scopus and WoS using the same search string, and metadata were exported.

2) Deduplication and preprocessing: Records from the two databases were merged and deduplicated; a language filter was then applied to retain English-language publications only.

3) Title–abstract screening: Titles and abstracts were screened for relevance against the criteria in Section 2.2, yielding a candidate set.

4) Full-text eligibility assessment: Candidate articles were read in full to verify whether they met the thematic requirements of "visual-art context + evidence on aesthetic judgment/creativity cognition," and reasons for exclusion were documented (e.g., primarily technical performance, non-visual-art objects, predominantly ethics/copyright).

5) Inclusion: The final set of journal articles was included for synthesis and entered the stages of data extraction and thematic integration.

The PRISMA flow numbers are reported in the Results section (Section 3.1) and presented visually in the manuscript (Figure 1, PRISMA flow diagram). Full-text exclusions were categorized by primary reasons to enhance transparency and reproducibility of the review process.

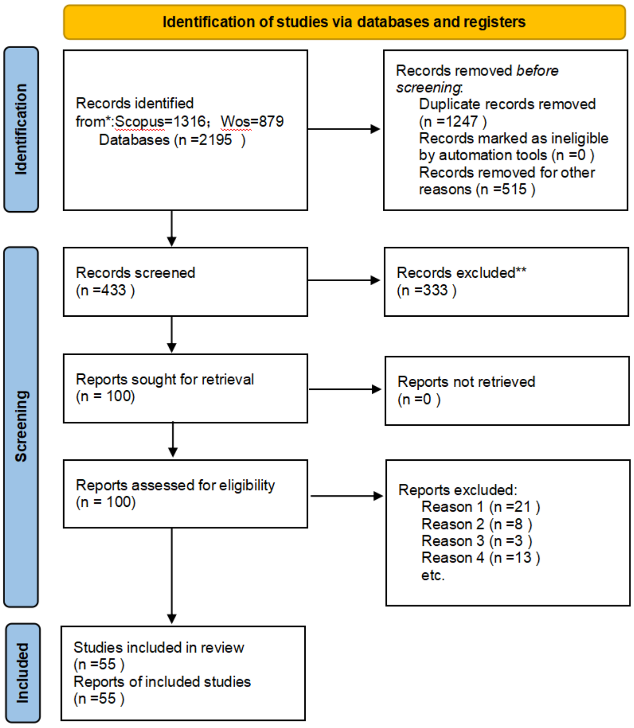

Figure 1: PRISMA Flowchart

2.4 Data Extraction and Evidence Synthesis

Full-text data were extracted from all included studies. Extracted items covered study type and disciplinary orientation; artwork sampling strategies and presentation formats; evaluative dimensions and measurement/detection approaches (e.g., beauty/aesthetic appeal, artistic value, creativity, authenticity, attribution of intention, and attribution of agency/subjecthood); as well as key findings and boundary conditions.

Evidence synthesis combined qualitative synthesis with thematic analysis/inductive theme development. Results were organized around the three research questions: RQ1 focuses on aesthetic and institutional-level conceptualizations; RQ2 examines evaluative differences and their moderating conditions; and RQ3 synthesizes the mechanisms underlying these differences and the associated evidence chains. All references follow APA 7th edition style.

3. Results

3.1 Study Selection and Overview of Included Research

Following the search strategy described in Section 2.1, we searched Scopus and WoS. The search yielded 1,316 records from Scopus and 879 records from WoS. After merging the two databases and removing duplicates, 949 unique records remained; after further restricting to English-language publications, 891 records were retained. By document type, the 891 English records comprised 433 journal articles (journalArticle), 346 conference papers (conferencePaper), 90 book chapters (bookSection), and 22 books (book).

Given that this review focuses on comparable evidence regarding "AI visual art in relation to aesthetic judgment and creativity cognition," subsequent screening concentrated primarily on journal articles (n = 434). We first conducted a title–abstract relevance screening to form a candidate set, followed by full-text eligibility assessment. Ultimately, 56 journal articles highly aligned with the topic were included. The main reasons for exclusion at the full-text stage were: (1) a primary focus on model/algorithm performance with little discussion of aesthetic judgment or creativity cognition; (2) non-visual-art targets or unclear art context; (3) a predominant emphasis on law/copyright/ethical frameworks; (4) strong application orientation but lacking evidence on mechanisms of "artistic value/creativity" judgment; and (5) absence of traceable methods and verifiable evidence chains.

In terms of publication timing, the topic has intensified markedly over the past three years: 13 studies were published between 2017 and 2022, whereas 42 studies appeared between 2023 and 2026 (approximately 75%), with a peak in 2025 (20 studies). In research orientation, the included studies can be broadly grouped into four categories: (a) philosophy/aesthetics and critical theory, focusing on authorship, intentionality, autonomy, and art institutions; (b) empirical studies on audience evaluation and bias, focusing on author-label effects, affective connection, and creativity attribution; (c) computational–psychological–neural mechanism studies explaining aesthetic judgment via image attributes, machine learning, or neural mechanisms; and (d) evidence synthesis and methodological frameworks, including meta-analyses and evaluation systems.

3.2 How is the Aesthetic Status of AI art as an Aesthetic Object Conceptualized

Across the included studies, conceptualizations of AI visual art "as an aesthetic object" largely revolve around three core questions: how authorship is defined, how artistic intention is interpreted, and how works are recognized as "art" within institutions and specific contexts.

First, regarding authorship and the creative process, research commonly shifts the analytical focus from the "single author" to the "creative configuration." AI art is typically described as a product co-produced through training data, model mechanisms, and human selection and intervention; accordingly, its authorship more closely resembles a collaborative arrangement than sole responsibility borne by a single subject. From this perspective, training data and their selection are not merely technical prerequisites; they also accumulate and transmit particular aesthetic orientations, gradually forming stylistic boundaries with "paradigmatic" tendencies in practice, which in turn shape the legitimacy of AI works being recognized as "art" in art contexts. Debates over whether an AI system can be an artist further suggest that the core controversy is not whether AI can generate works that "look like art," but whether the requirements of artist identity—responsibility, intentional expression, and social recognition—can be clearly and reasonably allocated within this creative structure.

Second, regarding the attribution of intention and creative autonomy, philosophical and art-theoretical research broadly forms two relatively diverging lines of argument. One position holds that under conditions of automation and generation, works may still participate in the evolution of the art system in new ways; therefore, AI works should be granted more open recognition at the levels of art institutions and aesthetic experience [9]. The other position argues that "artificial art" contains a difficult conceptual tension: works are expected to manifest creative agency and intentional expression, yet are often understood as products driven by programs—leaving their artistic status persistently contested in axiological and normative debates [10].

Third, with respect to art institutions and practice settings, studies show that acceptance of "AI art" is not determined solely by a work's formal features, but also by how viewers, critics, and practitioners understand the creative process and to whom value should be attributed. Audience-oriented research indicates that information about "how it was made" can significantly alter the ways and degrees to which a work is recognized within art contexts [1]. In addition, discussions of how specific cases enter art institutions and contemporary art contexts are often used to test the conditions under which the claim "AI-generated works can be regarded as contemporary artworks" can be sustained. Practitioner-focused interview studies further show that digital artists frequently connect AI works to concerns about insufficient expressivity and subjecthood, as well as related ethical anxieties; consequently, they tend to adopt more cautious or conditional stances toward recognizing their artistic status. In cultural and critical research, whether AI works possess "artistic eligibility" is not treated as a neutral conclusion with universal applicability; rather, it is deeply shaped by specific cultural experiences and knowledge frameworks.

Overall, the aesthetic status of AI visual art is a form of aesthetic objecthood that is highly context-dependent and must be "established" through social relations and institutional practices. Put differently, whether an AI work can be recognized as an "art object" primarily depends on three factors: whether the creative process has a clear attribution structure, whether intention and subjecthood are sufficiently explainable, and whether art institutions and cultural frameworks are willing to recognize and support the attribution of value. This definitional framework also provides a direct reference for subsequent empirical studies: when viewers are cued about "who created" and "how it was created," their judgments of artistic value, creativity, and authenticity often shift accordingly. Therefore, the next section (3.3) turns to evaluative differences under AI-labeled versus human-labeled authorship conditions and further synthesizes their boundary conditions.

3.3 Evaluative Differences Between AI-generated art and Human-made Art Across Beauty, Artistic Value, and Creativity

Across the included empirical studies, a relatively stable conclusion has emerged regarding "evaluative differences": judgments of beauty are more readily driven by a work's formal features and individual preference, whereas judgments of artistic value, creativity, and authenticity are more sensitive to authorship information and the associated attributions of intention and agency/subjecthood.

First, concerning authorship-information effects and their directional bias, multiple experimental studies consistently show that when the artwork itself is held constant, labeling the same work as "AI-generated" tends to lower viewers' overall evaluations, with the decline more concentrated in value-laden dimensions such as artistic value, creativity, and authenticity. This pattern is also supported at the meta-analytic level in reviews focused on visual art: relative to human authorship labels, an "AI author" label is generally associated with lower evaluations, though the magnitude of the effect is moderated by contextual factors such as task design, metric selection, and sample composition [13]. At the same time, public attitudes toward "whether AI can be creative" are not homogeneous: across different groups, at least three response profiles can be observed—avoidance, appreciation, and no clear difference—suggesting that the standards underlying creativity judgments are not uniform across populations.

Second, regarding the key conditions under which "beauty does not necessarily decrease," studies indicate that when a work's formal features align more strongly with viewers' aesthetic preferences, the beauty gap between AI and human works can shrink substantially. For example, in abstract painting and style-transfer outputs, statistical image properties can predict aesthetic ratings relatively well, implying that formal features and the perceptual processing they elicit play a comparatively stable role in beauty judgments. In addition, self-relevance has been shown to significantly increase the aesthetic attractiveness of both authentic and synthetic works, suggesting that aesthetic judgment is not determined solely by authorship information. Relatedly, on the question of "whether people can detect AI works," evidence suggests that viewers cannot always reliably discriminate AI from human works; more importantly, even when detection is inaccurate, merely believing that a work comes from AI can still shape judgments of value and authenticity.

Third, with respect to emotional investment and aesthetic experience, studies find that when evaluators place greater emphasis on the premise that "there should be human emotion and creative intention behind the work," their emotional engagement with AI works tends to be weaker, thereby reducing the intensity and quality of the overall aesthetic experience [14]. Meanwhile, aesthetic judgments exhibit pronounced individual differences and remain relatively stable over time, indicating that it is insufficient to infer the structure and regularities of aesthetic value solely from group-level mean differences [5].

Fourth, regarding measurement approaches and cross-study comparability, included studies increasingly move beyond reliance on single "subjective ratings" toward evaluation frameworks that combine subjective judgments with objective indicators. On the one hand, machine-learning models can extract quantifiable cues related to symbolism, affect, and imaginativeness from artworks to predict creativity evaluations. On the other hand, research on quality assessment for AI-generated images has proposed operational schemes that integrate subjective ratings with objective metrics to improve the feasibility of comparison and replication across studies [7]. For widely used quantitative image attributes in aesthetic research, "toolbox-style" methods and resources have also been developed, providing a more standardized computational basis for subsequent work. In addition, studies on "image memorability" suggest that aesthetic processing may generate predictable downstream effects, offering a new entry point for comparing the enduring impact of works from different sources [11-12].

In sum, compared with human works, AI works are not necessarily disadvantaged on beauty; differences are more contingent on formal features, individual preferences, and self-relevance. However, on dimensions such as artistic value and creativity, authorship information—and the resulting attributions of intention, agency/subjecthood, and authenticity—more readily triggers systematic evaluative discounting. This pattern is broadly supported at the meta-analytic level, while exhibiting clear contextual and measurement-related boundary conditions. These differences indicate that evaluation is not determined by form alone, but is tightly coupled with how viewers interpret "who is creating" and "for what purpose." This leads directly to the next section: why authorship information persistently elicits bias and skepticism, and how cognitive, cultural, and algorithmic mechanisms jointly produce these effects.

3.4 Mechanisms Underlying Aesthetic Bias and Creativity Skepticism

Addressing the core question of "where aesthetic bias and creativity skepticism come from," the included studies propose interlocking and mutually reinforcing explanatory lines. Overall, such bias and skepticism are rarely caused by a single factor in isolation; rather, they emerge and are sustained through the interplay of attribution processes at the cognitive level, evaluative frameworks at the cultural level, and algorithmic generation together with salient formal cues in artworks.

First, at the level of cognitive attribution, the most prominent explanatory pathway is intentionality attribution. When evaluators struggle to attribute experience, emotion, and agency to AI, the work becomes less likely to be understood as "creation with expressive intention," and is therefore more prone to devaluation in artistic value and creativity; conversely, when individuals are more willing to attribute subjecthood and creative intention to AI, their appreciation and value judgments tend to be more positive.

Second, closely connected to this is a human-centered defensive mechanism. Research emphasizes the role of beliefs in human uniqueness: when individuals more strongly believe that creativity is an exclusively human capacity, they are more likely to maintain symbolic human superiority by derogating AI art, thereby intensifying aesthetic discounting and creativity skepticism toward AI works.

Third, at the cultural-mechanism level, cross-cultural comparative studies show that different cultural groups assign markedly different weights to cues used in judging "what counts as beauty." This suggests that aesthetic judgment is not fully determined by universal perceptual–cognitive processing, but is deeply shaped by culturally embedded evaluative habits and knowledge frameworks. Moreover, critical scholarship argues that when prevailing evaluative frameworks presuppose particular cultural and institutional conditions, AI works are more easily categorized under interpretations such as "lacking a soul" or "lacking subjecthood," thereby weakening both the likelihood and justificatory grounds for recognizing them as legitimate art objects.

Fourth, at the level of algorithmic and formal mechanisms, relevant studies build arguments across two stages: the "generation side" and the "presentation side." On the generation side, research indicates that training data and the ways they are curated can implicitly stabilize certain aesthetic preferences and shape stylistic boundaries in the outputs, thus affecting the types of art viewers are ultimately exposed to and the "paradigmatic/canonical" feel they may perceive [6]. On the presentation side, studies show that artworks' formal features can be quantified and mapped onto viewers' aesthetic responses in testable ways; corresponding tools and methods make these "feature–response" relations more verifiable and comparable across studies. In addition, neuro-mechanistic research suggests that aesthetic value may be formed through the integration of multi-level visual features and, during processing, extend into higher-order cognitive and decision systems. At the level of methodological integration, interdisciplinary reviews and framework papers on "aesthetic and creativity evaluation" argue that future work must build clearer correspondences among psychological measures, controlled choices in artwork sampling and presentation, and algorithmic indicators in order to improve comparability and cumulative knowledge-building [4].

Overall, aesthetic discounting and creativity skepticism toward AI art are primarily driven by limited attribution of creative intention and agency/subjecthood, beliefs in uniquely human subjectivity, culturally embedded evaluative frameworks, and the combined effects of data–algorithm shaping and formal cues in artworks. These effects vary in strength and expression across cultural groups, task contexts, and measurement approaches.

4. Discussion

This review synthesizes research on AI and visual art with respect to aesthetic judgment and creativity cognition, organized around three research questions. For RQ1, we find that there is still no scholarly consensus on the aesthetic status of AI art. On the one hand, some theoretical work advocates expanding definitions of art, treating AI-generated outputs as an emerging artistic form and emphasizing that human–AI collaboration may introduce new artistic paradigms. On the other hand, many scholars and practitioners maintain that art requires human intention and emotion; because AI works are perceived as lacking this "soul," they are regarded as deficient in essential ways. Empirical evidence further indicates that audiences intuitively associate outstanding art with human creators, which renders the artistic status of AI works comparatively fragile in public perception.

For RQ2, a large body of psychological experiments demonstrates a substantial impact of authorship information on aesthetic evaluation: when the author is unknown, beauty ratings for AI works can be comparable to those for human works; however, when the work is explicitly identified as AI-created, its artistic value and creativity are often underestimated, reflecting a robust negative bias. Recent studies also suggest that this bias is most pronounced in competitive comparative contexts, and that its underlying structure may be better characterized as an "added-value premium" granted to human-made works—namely, a positive favoritism toward human art.

For RQ3, we synthesize mechanisms that account for these biases: cognitively, people more readily acknowledge creators with agency and emotional expressiveness, so AI—perceived as "less mind-like"—is more easily devalued; culturally, art is frequently construed as an expression of humanity, which reinforces stereotypes of AI art as "soulless"; and algorithmically, AI's data-driven mode of production and its lack of autonomous intention invite doubts about originality and authenticity. Taken together, aesthetic prejudice and creativity skepticism toward AI art arise from the interaction of multiple factors rather than any single cause.

4.1 Theoretical and Practical Implications

Theoretical implications. This review prompts aesthetics and philosophy of art to reconsider conventional categories. The emergence of AI art challenges a long-standing Enlightenment-derived proposition in Western thought concerning art and humanity: if something affords us an experience analogous to art in both sensory and meaning-making terms, yet has no human creator, should we still call it art? Nannicelli and colleagues offer a conciliatory position: we may appreciate the formal beauty of AI works while acknowledging that they lack certain dimensions required for "art," such as creator intention. This resembles our appreciation of the grandeur of natural landscapes—we may find a sunset aesthetically moving, yet we would not call it a painting. By contrast, Coeckelbergh and others caution against overly rigid classifications and argue for conceptual openness, including the possibility of "hybrid creativity," in which humans and AI are jointly treated as creative agents. Such theoretical exploration enriches definitions of creativity and art, allowing them to accommodate technological development.

At the psychological level, the human aesthetic bias synthesized in this review resonates with broader patterns of "interpersonal preference": people tend to favor entities perceived as similar to themselves and emotionally endowed, while remaining cautious toward cold, machine-like intelligence. This reflects anthropocentric bias and parallels patterns observed in human–computer interaction and related domains. In this sense, AI art provides a distinctive window into how humans perceive non-human intelligence: it carries both aesthetic and social meanings, and both dimensions shape attitudes and judgments.

Practical implications. Understanding these mechanisms matters for improving public acceptance of AI art and supporting its applications. First, artists and curators may benefit from providing richer narratives and contextual information when presenting AI art. For example, highlighting the human artist's role in dataset curation, algorithm training, selection, and post-editing can foreground human creative guidance and help narrow the perceived gap of "mindless creation." Second, in educational settings, strengthening public understanding of AI creative processes may correct the misconception that "AI merely copies mechanically." In practice, many AI works are produced through complex probabilistic generative principles rather than simple replication. Explaining, for instance, how generative models learn from large-scale art images and then generate novel images in a dreamlike manner may shift audiences' perceptions of uniqueness and intentionality. More fundamentally, cultivating an open aesthetic stance may be essential: just as photography was initially viewed as a threat by painters but later became recognized as an art form, society may similarly need time to accept AI as a new artistic medium or collaborator. Finally, for the development of AI art itself, creators may experiment with "humanizing" elements to enhance resonance—for example, producing artworks based on human physiological data or generating works through interaction with viewers. Such approaches may reduce psychological distance and mitigate bias.

4.2 Limitations and Future Directions

This review also highlights several limitations of the current evidence base. First, existing empirical studies are concentrated largely on Western audiences; comparatively little is known about how people from other cultural backgrounds perceive and evaluate AI art. Cultural variation may yield substantively different patterns and warrants deeper investigation. Second, many studies rely on short-term experiments and questionnaires that capture immediate reactions, whereas artistic evaluation often evolves over time. It is plausible that the value of AI artworks may be reinterpreted after temporal "settling." Longitudinal tracking of public attitudes and shifts in the art world's perspectives would therefore be a meaningful direction. Third, this review primarily addresses two-dimensional visual art; aesthetic evaluation of AI in music, literature, and other art domains also merits comparative work to examine whether similar or domain-specific bias patterns emerge. Finally, from the AI side, ongoing technological progress may enable future generative models to simulate intention or exhibit stronger forms of autonomy (e.g., acquiring sustained creative direction through interactive learning). Such advances may partially satisfy public expectations of "creative agency," potentially reducing bias. This also points to future research on how more anthropomorphic AI creation affects human evaluations and whether a "tipping point" exists—namely, whether perceptions of AI's artistic status shift fundamentally once AI creativity surpasses a certain threshold. Overall, AI art remains an emerging interdisciplinary field with many unresolved questions awaiting deeper inquiry.

5. Conclusion

Generative AI is intervening in visual-art creation in unprecedented ways, posing challenges to both aesthetic theory and audience psychology. Through a systematic review of the past decade of literature, this study arrives at three main conclusions. (1) The aesthetic status of AI art remains contested. Some accounts treat it as an extension of artistic innovation, whereas others maintain that genuine art is inseparable from human intention and emotion. In broad terms, AI art compels a renewed reconsideration of the definitional boundaries of "art." (2) In aesthetic evaluation, public judgments of AI-generated art are ambivalent. When authorship is unknown, AI works can receive beauty ratings comparable to human works, indicating that their aesthetic potential is real; however, when the work is identified as AI-generated, people often lower their evaluations of artistic value and creativity. This discrepancy appears to stem primarily from psychological bias rather than from the work's formal qualities alone. (3) The origins of such bias are multi-layered, including humans' tendency to attribute agency and emotional expression to creators, cultural traditions that link art to humanity, and suspicions that AI's data-driven mode of production lacks original intention. Together, these factors shape aesthetic prejudice toward AI art and skepticism toward AI creativity. Importantly, such bias is not immutable: as technologies evolve and public understanding deepens, attitudes toward AI art continue to change.

In sum, the convergence of AI and art is a central issue for 21st-century artistic practice and theory. Current research both reveals psychological boundaries in how humans respond to machine creation and demonstrates the aesthetic value—and limitations—of AI-based production. Looking ahead, more mature forms of AI art and corresponding shifts in public aesthetic sensibilities are likely to emerge. This process will continue to test and refine our understanding of art and creativity, deepening the dialogue between technology and the humanities. In this evolving landscape, maintaining openness alongside rational reflection will help societies evaluate—and potentially embrace—the new artistic vistas introduced by AI.

References

[1] Audry, S., & Ippolito, J. (2019). Can artificial intelligence make art without artists? ask the viewer. Arts, 8(1). https://doi.org/10.3390/arts8010035

[2] Berryman, J. (2024). Creativity and style in GAN and AI art: Some art-historical reflections. Philosophy and Technology, 37(2). https://doi.org/10.1007/s13347-024-00746-8

[3] Bianchi, I., Branchini, E., Uricchio, T., & Bongelli, R. (2025). Creativity and aesthetic evaluation of AI-generated artworks: Bridging problems and methods from psychology to AI. Frontiers in Psychology, 16. https://doi.org/10.3389/fpsyg.2025.1648480

[4] Bortolotti, A., Candeloro, G., Palumbo, R., & Sacco, P. L. (2025). Supercomplexity: Bridging the gap between aesthetics and cognition. Frontiers in Neuroscience, 19. https://doi.org/10.3389/fnins.2025.1552363

[5] Brielmann, A. A., Berentelg, M., & Dayan, P. (2024). Modelling individual aesthetic judgements over time.

Philosophical Transactions of the Royal Society B: Biological Sciences, 379(1895). https://doi.org/10.1098/rstb.2022.0414

[6] Carpo, M. (2024). Every dataset is a canon. Architectural Design, 94(3), 14-19. https://doi.org/10.1002/ad.3050

[7] Chen, J., Shao, F., Chen, H., Wang, X., Guo, H., & Jiang, Q. (2025). Quality evaluation of AI-generated images: Subjective study and objective methodology. IEEE Transactions on Multimedia. https://doi.org/10.1109/tmm.2025.3618562

[8] Chiarella, S. G., Torromino, G., Gagliardi, D. M., Rossi, D., Babiloni, F., & Cartocci, G. (2022). Investigating the negative bias towards artificial intelligence: Effects of prior assignment of AI-authorship on the aesthetic appreciation of abstract paintings. Computers in Human Behavior, 137. https://doi.org/10.1016/j.chb.2022.107406

[9] Coeckelbergh, M. (2017). Can machines create art?. Philosophy and Technology, 30(3), 285-303. https://doi.org/10.1007/s13347-016-0231-5

[10] Cusmariu, A. (2023). Turing algorithms in art. Symposion, 10(1), 31-80. https://doi.org/10.5840/symposion20231012

[11] Davis, T. M., & Bainbridge, W. A. (2023). Memory for artwork is predictable. Proceedings of the National Academy of Sciences of the United States of America, 120(28). https://doi.org/10.1073/pnas.2302389120

[12] Davis, T. M., & Bainbridge, W. A. (2026). The temporal and spatial properties of memorability reveal insights into the art creation process. Cognition, 266. https://doi.org/10.1016/j.cognition.2025.106322

[13] De Rooij, A. (2025). Bias against artificial intelligence in visual art: A meta-analysis. Psychology of Aesthetics Creativity and the Arts. https://doi.org/10.1037/aca0000833

[14] Demmer, T. R., Kühnapfel, C., Fingerhut, J., & Pelowski, M. (2023). Does an emotional connection to art really require a human artist? emotion and intentionality responses to AI- versus human-created art and impact on aesthetic experience. Computers in Human Behavior, 148. https://doi.org/10.1016/j.chb.2023.107875