ISRG Journal of Arts, Humanities and Social Sciences (ISRGJAHSS)

ISRG PUBLISHERS
Abbreviated Key Title: ISRG J Arts Humanit Soc Sci
ISSN: 2583-7672 (Online)
Journal homepage: https://isrgpublishers.com/isrgjahss
Volume – IV Issue - IV (July – August) 2026
Frequency: Bimonthly

The Monet Test: AI Prejudice, Attribution-Conditioned Formalism, and the Future of Aesthetic Judgment

James Hutson

Lindenwood University, https://orcid.org/0000-0002-0578-6052

| **Received:** 28.07.2026 | **Accepted:** 01.08.2026 | **Published:** 04.08.2026

***Corresponding author:** James Hutson

Abstract

A 2026 conceptual intervention by the artist SHL0MS presented a cropped reproduction of the 1915 Water Lilies by Claude Monet as a newly generated artificial intelligence image and invited viewers to explain its inferiority to an authentic Monet. Respondents supplied confident accounts of deficient composition, depth, color, intentionality, and soul, although the image reproduced an established Monet painting. This episode illustrates attribution-conditioned formalism, a process in which provenance organizes perception and formal vocabulary rationalizes the resulting judgment. This critical integrative review places the incident in dialogue with empirical aesthetics, creativity research, artworld theory, historical reattribution, and contemporary cases of mistaken AI detection. Controlled studies show that AI labels can depress ratings of beauty, profundity, creativity, skill, and monetary value even when stimuli are identical or difficult to distinguish. Conversely, blind evaluations have often favored AI-generated poetry, human-AI haiku, machine-produced humor, and AI-assisted stories, while large-scale research preserves human advantages at the highest levels of divergent creativity. Historical cases involving Marie Denise Villers and Judith Leyster demonstrate that gender and reputation have long altered the estimation of unchanged objects. The article argues that craft and quality remain relevant but cannot independently explain artistic value, which is mediated by provenance, institutional authority, presumed effort, social identity, and market prestige. It proposes a four-domain framework separating formal-perceptual, conceptual-interpretive, procedural-ethical, and institutional-social evaluation, together with a two-stage blind and disclosed review process. These reforms can preserve contextual interpretation while preventing provenance prejudice from masquerading as visual analysis.

Keywords: AI art; aesthetic judgment; attribution bias; creativity; provenance

DOI: 10.5281/zenodo.21788635

Introduction

On May 12, 2026, the pseudonymous conceptual artist SHL0MS posted a cropped image from Claude Monet's (1840-1926) *Water Lilies* (1915) **(Figure 1)** and described it as a newly generated artificial intelligence image in Monet's style. The post asked viewers to explain, in detail, what made the image inferior to a real Monet, and an added platform label strengthened the false provenance cue (SHL0MS, 2026). The represented work was not a synthetic imitation but Monet's *Water Lilies*, painted around 1915 and held by the Bayerische Staatsgemäldesammlungen at the Neue Pinakothek in Munich. Many respondents nevertheless identified allegedly obvious defects in spatial depth, reflection, focal organization, color, coherence, emotion, and artistic soul, while others dismissed the image as generic technological debris (Harrison Dupré, 2026). Some knowledgeable painters and art historians challenged the premise and recognized features associated with Monet's late practice, an important qualification that prevents the event from being reduced to a universal failure of connoisseurship. Even with that qualification, the intervention isolated a revealing contradiction: the object remained constant, but its presumed authorship changed the vocabulary through which viewers believed they saw it. Formal properties that had belonged to the recognized practice of the famous Impressionist were redescribed as evidence of machine incompetence. The episode therefore offers a compact diagnostic test for the relationship among provenance, perception, and value in the contemporary artworld.

Figure 1. Claude Monet, *Water Lilis*, 1915. Bayerische Staatsgemäldesammlungen, Neue Pinakothek, Munich. Public Domain.

The most consequential lesson is not that craft, technique, or quality never matter, because visual and material distinctions plainly structure aesthetic experience. The stronger and more defensible conclusion is that craft and quality do not independently determine evaluation or value, and they cannot explain a judgment that changes while the object remains unchanged. When an identical image appears incompetent under an AI label and accomplished under a Monet attribution, the causal difference lies outside the visible artifact. It lies in expectations about agency, effort, intention, authenticity, status, and the perceived threat posed by nonhuman production. This article names the mechanism attribution-conditioned formalism: an evaluator first accepts a category or provenance cue, then converts that cue into apparently object-based claims about composition, expression, finish, or meaning. The resulting criticism sounds formalist, yet its direction has already been set by contextual belief. In this sense, the Monet intervention does not show that viewers lack visual intelligence; it shows that visual intelligence is top-down, socially organized, and vulnerable to post hoc rationalization. The same mechanism appears in controlled studies of AI labels and in older episodes when paintings praised under the names of famous men were re-described and devalued after attribution to lesser-known or female artists.

The argument developed here extends prior scholarship on cultural prejudice against artificial intelligence and on the distributed character of human-machine authorship. Hutson and Plate (2024) traced resistance to AI through technophobia, algorithm aversion, dystopian narratives, and anxieties about human replacement, while cautioning that evidence of machine competence is often discounted when it disturbs assumptions of human exceptionalism. Similarly, the earliest monographs on AI art situated these generative systems within a longer history in which artistic communities initially resist new technologies, then absorb them into revised definitions of practice, skill, and authorship (Hutson et al., 2024). Furthermore, the concept of Absent Presence identifies the human influence that persists within apparently automated outputs through training cultures, prompting, selection, framing, and interpretation (Vosevich & Hutson, 2024). Another multidimensional account of collaborative authorship further rejected a single continuum from fully human to fully automated, replacing it with separate dimensions of generative, structural, creative, and analytical contribution (Hutson, 2025). Applied to visual art, this framework reveals why the binary label "AI-made" is epistemically crude: it compresses a distributed process into a stigmatized category and erases the artist's decisions across ideation, iteration, curation, editing, fabrication, and contextualization. The present article builds on those claims by connecting them to aesthetic valuation and historical attribution. Its central proposition is that the artworld's difficulty reconciling AI with human creativity emerges less from a stable criterion of quality than from a threatened theory of who is permitted to count as a creator.

Scope and Method of the Critical Review

This article is a critical integrative review rather than a systematic review or a report of new experimental data. It synthesizes peer-reviewed research on AI authorship labels, empirical aesthetics, creativity assessment, effort and authenticity heuristics, museum context, and reputation effects, with emphasis on studies available through July 2026. The review also incorporates primary or official collection records for Monet, Marie Denise Villers, and Judith Leyster, together with documented contemporary cases in which human work was rejected as AI-generated or AI work crossed established competition categories. Sources were selected because they permit a comparison between stable artifacts and variable attributions, or because they directly compare human, AI, and human-AI creative performance. The method is therefore argumentative and triangulatory: a viral conceptual intervention supplies the initiating case, controlled studies test the causal role of labels, historical reattributions demonstrate institutional continuity, and current controversies expose operational consequences. No single class of evidence carries the entire argument. Social media reactions cannot establish population-level effects, while laboratory studies may not reproduce the stakes of museums, markets, and professional identity. Read together, however, these

DOI: 10.5281/zenodo.21788635

materials make it possible to distinguish object-level properties from attribution effects and to propose evaluative procedures appropriate to hybrid creative practice.

The Monet Test and Attribution-Conditioned Formalism

The Monet intervention functions as a test because it holds the represented image constant while manipulating its presumed origin. A viewer who first describes the image as lacking depth or coherence and then praises the same image after learning that Monet painted it confronts a direct inconsistency. The inconsistency cannot be repaired by claiming that provenance is irrelevant, since attribution can legitimately alter interpretation, historical importance, legal status, and market value. It can be repaired only by identifying which judgments concern visible organization and which concern contextual meaning. A claim about the reflection of willow branches, for example, presents itself as a formal observation and should remain stable across labels unless new technical evidence changes what is visible. A claim about Monet's late vision, the history of Giverny, or the significance of seriality properly depends on attribution and context. The problem arises when a contextual aversion to AI is redescribed as a visible defect, allowing "AI" to become both the accusation and the proof. Attribution-conditioned formalism therefore names a category error as much as a bias: evaluators treat provenance-based disapproval as though it were direct optical evidence.

The episode also clarifies why references to artistic "soul" carry so much weight in AI debates despite their analytical vagueness. In ordinary criticism, soul often serves as a compressed term for perceived intention, lived experience, vulnerability, embodied labor, or communicative presence. None of these qualities is directly visible in the way hue, edge, scale, or material texture is visible; viewers infer them from the object and from what they know or believe about its maker. Bellaiche et al. (2023) demonstrated this inferential structure by randomly assigning human or AI labels to artworks that were all generated by AI. Human labels increased ratings of liking, beauty, profundity, and worth, while perceived narrative and effort helped explain parts of the effect. The object did not acquire human experience when the label changed, but participants believed that it did, and that belief reorganized their evaluation. The Monet critics who discovered soullessness in an authentic painting enacted the same logic in public. Their responses reveal that soul is often not an aesthetic property found in the artifact; it is an attribution of personhood granted or withheld before the work receives sustained interpretation.

Effort provides a related mechanism because audiences commonly use presumed labor as a shortcut for value. Kruger et al. (2004) called this pattern the effort heuristic: when quality is difficult to assess, people infer value from the amount of work they believe production required. Cho and Schwarz (2008) subsequently showed that the heuristic is flexible and depends on assumptions about the maker's talent, which means that effort does not operate as a neutral quantity. Generative systems disrupt this shortcut because they can produce complex outputs rapidly, encouraging observers to equate reduced execution time with reduced achievement even when conceptual development, iteration, selection, or postproduction remains extensive. Horton et al. (2023) found that AI labels sharply reduced estimated production time and monetary value, even when the visual content was held constant. Such responses do not prove that time is aesthetically irrelevant, because process can matter conceptually and ethically. They show instead that presumed speed becomes a moralized proxy for merit. The phrase "just pressing a button" thus performs the same evaluative work that the word "soul" performs: it converts limited knowledge of a workflow into certainty about the poverty of the resulting artifact.

Attribution-conditioned formalism also helps explain why the criticism appeared detailed rather than merely dismissive. Bias rarely eliminates observation; it selects, weights, and narrates observations so that they support an expected conclusion. Once viewers accepted that the image was AI-generated, ambiguous features became diagnostic symptoms: a loose transition became incoherence, an unusual color relation became statistical imitation, and Monet's submerged reflections became an inability to understand physical space. The same features might have been described under an established attribution as optical daring, late-style compression, chromatic experimentation, or resistance to conventional focal hierarchy. Formal vocabulary is therefore not a transparent instrument applied after perception. It is part of the interpretive machinery through which perception becomes communicable and institutionally credible. This does not make all interpretations equivalent, nor does it abolish standards of close looking. It makes reflexivity essential, especially when a label has supplied the conclusion before formal analysis begins.

What Controlled Studies Show About the AI Label

Controlled experiments provide causal evidence that the Monet episode alone cannot supply. Chiarella et al. (2022) assigned AI and human authorship labels to human-made abstract paintings and found lower aesthetic appreciation for works presented as AI-authored. Bellaiche et al. (2023) used only AI-generated paintings, randomly labeled them as human or AI, and again found a consistent advantage for the human label across sensory and communicative judgments. Horton et al. (2023) extended the pattern across six experiments and 2,965 participants, showing that people devalued art labeled as AI-made even when they regarded it as indistinguishable from human art and even when the work was described as human-AI collaboration. The strongest penalties often concerned skill, creativity, likelihood of qualifying as art, production time, and monetary value. These dimensions are especially revealing because they sit at the boundary between visual judgment and social inference. Participants did not merely report that they disliked a certain surface. They reconstructed the presumed competence and labor of the maker from the label, then used those inferences to price the object.

Grassini and Koivisto (2024) produced an especially striking dissociation between actual source and believed source. In their sample, participants liked the selected AI-generated images more than the selected human-made images and reported more positive emotion in response to them. Participants nevertheless could not reliably distinguish the two sources, and images they *believed* were AI-generated received lower evaluations regardless of their actual origin. This reversal captures the central claim of the present article in empirical form: preference can favor machine output at the perceptual level while attribution suppresses that preference at the evaluative level. The result also demonstrates why disclosure studies must distinguish actual production history from subjective source belief. An AI label does not simply add neutral information; it activates assumptions about authenticity, effort, intention, and cultural threat. When evaluators then report reduced emotion, they may be reporting the consequence of those assumptions rather than

DOI: 10.5281/zenodo.21788635

a deficit detectable in the image. The Monet intervention reproduced the same reversal with an art-historical masterpiece rather than a controlled stimulus set.

More recent studies indicate that the label effect extends beyond explicit self-report. Zhang et al. (2025) conducted two preregistered experiments, with a combined sample of 125 participants, in which objectively identical paintings received human or AI labels. Participants rated the human-labeled works as more creative, aesthetically valuable, delicate, and preferable, and neural measures suggested differences in attention and semantic-emotional processing during evaluation. Wang, Leng, He, and Lin (2026) likewise found, across three studies using behavioral and eye-tracking measures, that lay viewers preferred paintings labeled as human-created even when they could not reliably identify the works' actual sources. Eye-movement results were mixed, which is theoretically useful because it cautions against imagining a single perceptual signature of prejudice. Labels may shape attention, interpretation, response strategy, or all three, depending on task and viewer. The consistent behavioral outcome remains difficult to explain by craft or quality because the stimuli were unchanged or indiscriminable. What changes is the viewer's model of agency.

The strongest synthesis currently available also requires restraint. De Rooij's (2025) meta-analysis aggregated 35 experiments conducted between 2017 and 2024 and found negative effects of AI attribution across sensory-motor, emotion-valuation, and knowledge-meaning systems, with the largest pooled effect in knowledge and meaning. The effects were heterogeneous and moderated by factors such as age, art style, image source, and situatedness, so bias against AI should not be treated as a universal or structurally identical response. That nuance strengthens rather than weakens the present argument. The artworld does not apply one invariant penalty to all AI-associated work; it applies a variable penalty where authorship, intention, and human significance become salient. The meta-analysis also suggests that the label's greatest influence appears not necessarily in low-level perception but in judgments about what the work means and what kind of agency it represents. Attribution-conditioned formalism becomes especially powerful at precisely this junction because evaluators translate meaning-level aversion into claims about visible quality. A rigorous account must therefore reject both extremes: AI labels are not irrelevant, but neither do they reveal intrinsic inferiority.

When AI Wins Blind: Poetry and Creative Performance

Research on poetry makes the contradiction between blind preference and disclosed devaluation particularly difficult to ignore. Porter and Machery (2024) compared poems generated in the styles of well-known poets with poems written by those poets. In their first study, 1,634 participants performed below chance when identifying AI poems, and AI-generated poems were more often judged to be human than the human-authored poems. In a second study with 696 participants, the AI poems received higher overall evaluations and scored higher on nearly every assessed dimension, yet the belief that a poem was AI-generated reduced its rating. The result does not prove that language models are superior poets in every genre, period, or evaluative context. It demonstrates that readers can prefer machine-generated verbal artifacts while maintaining a theory that machine authorship should lower value. The label therefore produces a judgment contrary to the preference elicited by the text itself. This is the literary equivalent of finding defects in Monet only after being told to look for AI.

Table 1. Blind Evaluation of Human, AI, and Human-AI Creative Performance

Study	Creative domain	Conditions compared	Blind-evaluation result	Disclosure or limitation	Interpretive significance
Porter and Machery (2024)	Poetry	Human versus AI	AI poetry rated more favorably	AI belief reduced ratings	Preference and provenance judgment diverged
Hitsuwari et al. (2023)	Haiku	Human, AI, collaboration	Collaboration rated highest	Conditions difficult to identify	Distributed creativity can outperform binary categories
Gorenz and Schwarz (2024)	Humor	Human versus ChatGPT	AI jokes equaled or exceeded human responses	Task and model specific	Machine involvement can satisfy human quality criteria
Lee and Chung (2024)	Ideation	Unassisted, web search, ChatGPT	ChatGPT increased creativity	Strongest for incremental novelty	AI can support combinatorial creativity
Doshi and Hauser (2024)	Short fiction	Human versus AI-assisted	Assistance improved quality and enjoyment	Collective similarity increased	Individual gains may produce cultural homogenization
Wang et al. (2026)	Divergent thinking	Humans versus language models	Humans stronger at the exceptional tail	Task-specific comparison	Average fluency does not equal exceptional innovation
Wan and Kalman (2026)	Collaborative storytelling	Human-only and diverse AI personas	Diverse personas mitigated convergence	Structured intervention required	Homogenization is design-sensitive

Haiku research reveals a complementary advantage for collaboration rather than autonomous generation. Hitsuwari et al. (2023) asked participants to evaluate human-written, AI-generated, and human-AI collaborative haiku. The collaborative poems received the highest aesthetic ratings, while participants could not reliably distinguish among production conditions. Human-only and AI-only poems were evaluated similarly, which complicates narratives of either total machine triumph or obvious machine

DOI: 10.5281/zenodo.21788635

deficiency. The important result is that the most valued output emerged from an interaction whose contributions could not be cleanly read from the text. This finding aligns with Hutson's (2025) multidimensional framework, in which creative agency is distributed across functions rather than assigned through a binary origin label. It also shows why mandatory disclosure, although ethically important in some contexts, can alter the reception of work that evaluators otherwise find compelling. Disclosure should inform judgment, but institutions must prevent it from becoming a substitute for judgment.

Comparable findings appear in humor, ideation, and narrative writing. Gorenz and Schwarz (2024) reported that blind evaluators rated jokes produced by ChatGPT 3.5 as equally funny or funnier than human responses across several comedic tasks, including comparison with professional comedy writing. Lee and Chung (2024), across five experiments, found that ChatGPT increased the creativity of ideas relative to both unassisted work and conventional web search, particularly for incrementally novel solutions that combined remote concepts into coherent proposals. Doshi and Hauser (2024) found that access to generative ideas improved the creativity, quality, and enjoyment of short stories, with especially strong benefits for less creative writers. These studies do not establish that output quality is independent of prompting, model version, genre, or evaluation protocol. They establish that machine involvement can generate aesthetic and creative gains that human judges recognize when source identity does not preempt evaluation. The repeated pattern matters: AI is often capable of passing or winning the quality test, then losing the provenance test. The artworld's challenge is to decide whether it intends to evaluate artifacts, processes, moral economies, or all three, and to stop collapsing those questions into a single adjective such as "good."

The evidence also imposes limits on any triumphalist account of generative creativity. Doshi and Hauser (2024) found that AI-assisted stories became more similar to one another, revealing a tradeoff between individual improvement and collective diversity. Wang, Huang, Shen, and Uzzi (2026), in a large-scale comparison involving 9,198 humans and more than 215,000 model observations, found that humans were slightly more creative on average in a divergent-thinking task and substantially stronger in the far-right tail of the distribution. These results preserve an important human advantage in exceptional, high-variance production and caution against equating fluent average performance with cultural innovation. They also shift the practical question from whether AI is creative in the abstract to how creative systems are designed and situated. Wan and Kalman (2026) showed that structured access to diverse AI personas could preserve story diversity relative to a human-only baseline and in some conditions enhance it, suggesting that homogenization is not an unavoidable property of collaboration. Human creativity remains vital through problem definition, counterexample generation, aesthetic risk, selection, responsibility, and the construction of diverse inputs. The point is not that machines replace creators, but that the binary opposition between machine output and human creativity no longer describes the actual ecology of production.

The Artworld Has Always Priced the Name

AI bias appears novel because the category is novel, but the institutional mechanism is familiar. Danto (1964) argued that artworks become legible within an artworld of theories, histories, and interpretive possibilities, while Walton (1970) showed that aesthetic properties depend partly on the categories under which a work is perceived. Becker (1982) located production within cooperative networks, and Bourdieu (1993) explained how fields of cultural production convert recognition, position, and symbolic capital into value. These accounts undermine the fantasy that artistic judgment is ever purely retinal. Museum context can increase interest, liking, arousal, and memory, even when the displayed works are otherwise comparable to laboratory presentations (Brieber et al., 2015). Artist names also alter ranking and monetary estimates: Cleeremans et al. (2016) found that the presence of names raised rankings among attentive nonexpert viewers, while Hernando and Campo (2017) showed that artist brand influenced general appraisal, perceived economic value, and willingness to pay. The name is not an accidental supplement to the work; within the market and museum, it is part of the product being evaluated.

Contextualism does not itself constitute prejudice, because historical origin can be relevant to what a work is and what it means. An original, a forgery, a quotation, and an authorized replica may be visually identical yet differ in historical agency, legal status, risk, and conceptual achievement. Newman and Bloom (2012) showed that judgments of authenticity and value depend on beliefs about an object's creative history, not only on sensory properties. The problem begins when institutions deny that they are judging history and instead claim that reputation has revealed quality in the surface. A famous name may prompt longer viewing, richer associations, and more charitable interpretation, while an AI label may prompt rapid dismissal and diagnostic defect hunting. Both responses are context-sensitive, but only one is often presented as disinterested connoisseurship. The proper response is not to eliminate context, which would make much art history unintelligible. It is to separate context-dependent interpretation from claims about formal success and to make the evaluative transition explicit.

This distinction changes how one should state the relationship between quality and value. Formal quality can contribute to value, but value is a relational achievement produced through institutions, provenance, scarcity, narrative, identity, and expectations about labor. Quality itself is not immune to those relations because evaluators learn what to notice and how to name it within communities of practice. A rough surface may signal incompetence in one frame and radical materiality in another; repetition may indicate laziness, serial inquiry, ritual, or algorithmic dependence depending on context. The Monet episode made that ordinary artworld process visible by switching only one cue and allowing the critical vocabulary to reverse. AI did not create the susceptibility; AI activated it. For that reason, the controversy should not be treated as a narrow argument about new software. It is a stress test for the artworld's longstanding claim that its hierarchies reflect quality rather than socially authorized attribution.

Gendered Reattribution: Villers and Leyster

The reattribution history of Marie Denise Villers' (1774-1821) *Marie Joséphine Charlotte du Val d'Ognes* (1801) (Figure 2) offers a powerful precedent. When the Metropolitan Museum of Art received the painting in 1917, it was attributed to Jacques-Louis David (1748-1825), promoted as the "New York David," and became a popular work in the collection (Metropolitan Museum of Art, 2023). In 1951, Charles Sterling demonstrated that

DOI: 10.5281/zenodo.21788635

David could not have painted it and proposed Constance Marie Charpentier, another woman artist, before Margaret Oppenheimer established the attribution to Villers in 1996 (Oppenheimer, 1996; Sterling, 1951). The canvas did not change across these revisions. What changed was the status attached to its author and, with that status, the language available to describe the painting. Sterling's reattribution famously recoded features once admired under David's name through a gendered vocabulary of charm, femininity, and concealed weakness. The formal evidence did not suddenly become softer or less structurally ambitious. Institutional expectations about what a woman artist could have achieved supplied a new interpretive template.

Figure 2. Marie Denise Villers, Young Woman Drawing, 1801. Metropolitan Museum of Art. Public Domain.

Villers's case is especially instructive because the work depicts a young woman artist, making the history of attribution part of its contemporary meaning. Under David's name, the painting could stand as an accomplished product of Neoclassical mastery; under a woman's name, the same achievement became vulnerable to diminutive description. Nochlin's (1988) institutional critique explains why such reversals cannot be reduced to isolated bad judgment. Women's historical exclusion from academies, training, commissions, collections, and canonical writing produced a circular evidentiary system: women were assumed to lack greatness because few were recognized as great, and their unrecognized work then appeared to confirm the assumption. AI enters a different moral and historical category, and prejudice against machines must not be equated with the lived oppression of women. The structural analogy is narrower and nevertheless important. In both cases, evaluators begin with a hierarchy of authorized creators and then discover aesthetic reasons for the hierarchy in objects that had previously supported another conclusion. Attribution becomes an engine for seeing what the institution already expects.

The rediscovery of Judith Leyster (1609-1660) makes the economic consequence even more explicit. Leyster was respected in seventeenth-century Haarlem, admitted as a master to the Guild of Saint Luke, and ran a workshop, yet much of her oeuvre disappeared into attributions to Frans Hals or her husband after her death (Hofrichter, 1989). In 1893, a work sold as Hals was found to bear Leyster's distinctive monogram beneath a false Hals signature, and the publicity helped Cornelis Hofstede de Groot identify additional paintings by her **(Figure 3).** Contemporary accounts of the litigation report that the settlement reduced the retained price from £4,500 to £3,500, commonly summarized as a roughly 25% loss after the work ceased to be a Hals (Benzine, 2024). Nothing in the brushwork deteriorated when the monogram emerged. The price changed because the market had purchased Hals's symbolic capital, not only the painted scene. Leyster's recovery eventually transformed the scholarly assessment of her work, but the initial response treated correct attribution as financial damage rather than as the discovery of an artist capable of matching a celebrated man. The episode demonstrates that market value can register prestige bias with numerical clarity.

Figure 3. Judith Leyster, Carousing Couple, 1630. Public Domain.

Villers and Leyster also reveal why reattribution should not be discussed only as a moral failure of past connoisseurs. Their histories expose an evaluative architecture that remains active whenever institutions use reputation as a quality signal. Famous names reduce uncertainty, attract scholarship, encourage conservation, and generate comparison with an established oeuvre; unknown names receive less interpretive investment and therefore appear less complex. AI labels intensify the asymmetry because they can imply not merely an unknown maker but the absence of a maker altogether. The machine category arrives burdened by assumptions of theft, automation, infinite reproducibility, and ontological emptiness, some connected to legitimate ethical

DOI: 10.5281/zenodo.21788635

concerns and others functioning as generalized stigma. Once that category is activated, evaluators may withdraw the patient attention through which complexity becomes perceptible. The historical lesson is therefore procedural: where status strongly predicts attention, attention will later appear to validate status. Blind or staged evaluation becomes necessary not because context is meaningless, but because institutions need a way to measure how much their context has preselected the outcome.

False Positives and Category Reversals in the Generative Era

The stigma attached to AI has begun to harm human artists whose work is merely suspected of machine involvement. In late 2022, digital artist Ben Moran posted *A Muse in Warzone* to Reddit's large r/Art community and was banned under a prohibition on AI art, although the image was produced through a conventional digital illustration workflow. Moran offered process evidence, yet a moderator rejected the explanation and reportedly advised the artist to change style (Cole, 2023). In 2023, judges praised an iPhone photograph by Suzi Dougherty before excluding it from a Sydney competition because they suspected it might be AI-generated, despite the absence of reliable evidence that it was synthetic (Shepherd, 2023). These cases show that categorical suspicion can become aesthetically self-confirming. The more polished, uncanny, saturated, or stylistically hybrid a human work appears, the more likely some viewers are to read those qualities as proof of AI. Artists then face pressure to document process, avoid certain visual conventions, or perform visible labor so that their work remains legible as human. A prejudice ostensibly designed to protect human creativity can thus narrow human experimentation.

Competition controversies reveal a related but distinct issue: a work may succeed aesthetically while failing category rules. Boris Eldagsen's AI-assisted *Pseudomnesia: The Electrician* won the creative open category of the 2023 Sony World Photography Awards before the artist declined the prize and argued that AI images and photography required clearer separation (Lau, 2023). In 2024, Miles Astray reversed the experiment by entering a camera-made photograph, *F L A M I N G O N E*, in an AI category; the image won both a jury placement and a public vote before being disqualified when its human origin was disclosed (Katz, 2024). The two episodes should not be narrated simply as judges being fooled. In both, the selected artifact met the evaluative expectations of the category well enough to win, while the disclosed production process changed its eligibility. That outcome is legitimate when rules define a medium, just as an oil painting would be ineligible for a watercolor prize regardless of quality. The analytical error occurs when disqualification is treated as evidence that the work was aesthetically weak. Category compliance, process ethics, and aesthetic merit answer different questions and should remain separately reported.

Together, false positives and category reversals create a double bind. Institutions demand disclosure because provenance affects fairness, copyright, consent, and public trust, yet the AI label itself can trigger a penalty unrelated to the artifact's perceptual or conceptual accomplishment. Attempts to solve the problem through automated detection remain unstable because contemporary systems and editing workflows blur source boundaries, while human judgment often mistakes stylistic novelty for synthetic origin. The solution cannot be secrecy, since deception may invalidate competitions and undermine accountability. Nor can it be a binary badge that collapses one automated fill, one generated texture, extensive iterative co-creation, and autonomous batch production into the same category. What is required is contribution-sensitive provenance. Evaluators need to know what human agents and computational systems did, when they did it, what materials informed the process, and who assumed responsibility for selection and publication. Only then can institutions judge formal achievement, conceptual stakes, ethical process, and category compliance without using one dimension as a proxy for all the others.

Reconciling Artificial Intelligence With Human Creativity

The artworld's most persistent conceptual error is to treat AI creativity and human creativity as mutually exclusive quantities. Generative models do not possess biographies, bodies, social vulnerability, or moral accountability in the human sense, but their outputs are produced within human-designed infrastructures and human cultural archives. Artists formulate problems, choose systems, construct prompts or training sets, vary parameters, reject outputs, edit fragments, combine media, fabricate objects, and situate results within discourses. Vosevich and Hutson (2024) described this condition as an absent presence because human influence remains distributed through content that may appear automatically generated. The phrase is particularly apt for AI art, where the visible artifact often hides the labor of selection and the invisible histories embedded in model development. Recognizing distributed agency does not require pretending that the model is a person. It requires abandoning the fiction that creative causation must be singular to be meaningful.

Hutson's (2025) multidimensional authorship model provides a practical alternative to binary classification. In a visual-art adaptation, contribution can be documented across at least six functions: problem formation, generative production, structural organization, aesthetic direction, analytical evaluation, and final accountability. A model may perform a large share of image generation while the human artist retains control over the concept, selection criterion, serial logic, editing, material realization, and public claim. In another project, a human may provide only a short prompt and publish the first output, leaving little evidence of sustained authorship despite nominal human initiation. Both works might receive the same "made with AI" label, although their creative structures differ substantially. Contribution analysis therefore protects human agency more effectively than categorical exclusion. It also permits criticism to identify where the work's achievement actually lies, whether in image synthesis, system design, conceptual framing, curation, fabrication, or social intervention.

The Monet test further suggests that human creativity is not secured by denying machine competence. When defenders of human uniqueness insist that AI output must be inferior, any successful output becomes an existential threat and any failure becomes comforting proof. This produces motivated criticism rather than discriminating criticism. A more stable account locates human creativity in capacities that remain important even when machines generate persuasive forms: choosing consequential problems, entering lived relationships, making commitments, taking responsibility, sustaining practices over time, and transforming outputs into culturally situated acts. Recent research showing a human advantage at the highest tail of divergent creativity supports this account without returning to mystical exceptionalism (Wang, Huang, Shen, & Uzzi, 2026). Human creativity can remain distinctive because it is embodied, accountable, developmental, and social, not because every human

DOI: 10.5281/zenodo.21788635

artifact must outperform every model artifact. The artworld can then recognize machine competence without surrendering the human conditions through which art acquires stakes.

Such reconciliation also demands a historical perspective on tools. Photography, mechanical reproduction, video, digital compositing, and algorithmic media repeatedly altered the location of skill rather than abolishing skill. Hutson et al. (2024) argued that emerging technologies shift artistic emphasis among ideation, conceptualization, process, and execution, requiring education to move beyond technique-centered definitions of mastery. Generative AI accelerates this shift because execution can be rapid and massively variable, increasing the importance of selection, direction, provenance, and conceptual necessity. The result does not make craft obsolete; it pluralizes craft. Prompt design, dataset construction, model fine-tuning, image compositing, animation, physical translation, and critical sequencing can each involve deep expertise, although none should be romanticized merely because it is new. The proper question is not whether the hand touched every mark. It is whether the artist's decisions produce a coherent, consequential, and accountable practice.

A Four-Domain Model for Aesthetic Judgment

The evidence reviewed here supports a four-domain model that separates formal-perceptual, conceptual-interpretive, procedural-ethical, and institutional-social evaluation **(Table 2).** The domains interact, but they should not be collapsed into one global score because a work can excel in one and fail in another. A formally compelling image may rely on an exploitative dataset; an ethically careful process may produce an aesthetically weak artifact; a conceptually incisive project may intentionally use banal forms; a market success may owe more to scarcity and reputation than to perceptual distinction. Traditional criticism often moves among these domains without announcing the transition. AI intensifies the confusion because provenance is treated simultaneously as a formal defect, an ethical violation, a conceptual limitation, and a market signal. Explicit domains allow evaluators to state what they are judging and what evidence supports the judgment. They also make disagreement more productive, since critics can dispute a work's ethical process without pretending that its composition is therefore incoherent.

Table 2. Four-Domain Evaluation Rubric

Domain	Core question	Appropriate evidence	Common attribution error	Recommended procedure
Formal-perceptual	What is perceptually present and how is it organized?	Artifact, display conditions, material examination	Inferring machine origin from stylistic features	Source-blind initial review
Conceptual-interpretive	What does the work mean within its relevant histories and discourses?	Artist statement, attribution, title, context, serial relationships	Assuming that a model's lack of biography eliminates human framing	Disclosed contextual review
Procedural-ethical	How was the work produced, and what harms or obligations follow?	Workflow records, model and data information, consent and licensing	Treating any AI use as equivalent	Contribution-sensitive documentation
Institutional-social	How do reputation, scarcity, markets, and institutions shape reception?	Exhibition, market, collection, and reception histories	Treating price or prestige as direct evidence of quality	Report institutional value separately

a. Formal-Perceptual Evaluation

The formal-perceptual domain concerns the artifact as encountered: composition, color, rhythm, scale, material articulation, spatial organization, complexity, coherence, surprise, affective force, and sensory resolution. Initial evaluation in this domain should occur blind to authorship whenever institutional conditions permit. Blindness does not create a context-free eye, because viewers still bring education, culture, and category knowledge, but it prevents one powerful provenance cue from preloading the result. Reviewers should record observations in falsifiable or at least inspectable terms, distinguishing what is present from what they infer about the maker. A statement that a reflected branch merges with a lily pad can be checked against the image; a statement that the merger proves a machine does not understand water cannot. The Monet test shows why this discipline matters. Formal criticism should survive a change in attribution unless the new attribution supplies genuinely relevant perceptual or material evidence.

b. Conceptual-Interpretive Evaluation

The conceptual-interpretive domain asks what the work does within histories, discourses, genres, and communities. Here, attribution is indispensable because Monet's late serial practice, an AI artist's system design, and SHL0MS's act of deceptive framing create different objects of interpretation even when they involve the same image. Intentionality should be treated as documented or argued rather than mystically presumed. Human selection, juxtaposition, title, exhibition context, and public intervention may generate meaning that no single pixel contains. AI can participate in this domain as material, medium, collaborator, subject, or institutional provocation. The question is not whether the model privately intended the result, since models do not provide accountable intentions in the human sense. The question is how agents organized the process and what interpretive possibilities the resulting work sustains.

c. Procedural-Ethical Evaluation

The procedural-ethical domain addresses training-data provenance, consent, copyright, labor displacement, environmental cost, deception, privacy, bias, and compliance with competition rules. These concerns are substantial and should not be dismissed as mere technophobia. A work can be formally strong and ethically unacceptable, just as a historically important object can be implicated in exploitation or illicit acquisition. Ethical criticism

DOI: 10.5281/zenodo.21788635

becomes more precise when it identifies a process and harm rather than treating the AI label as sufficient evidence. Different systems, datasets, licenses, and workflows create different obligations. Institutions should require enough process documentation to assess those obligations, while recognizing that complete technical transparency may be impossible. Separating ethics from formal evaluation prevents two opposite errors: aesthetic success does not excuse unethical production, and ethical concern does not prove aesthetic failure.

d. Institutional-Social Evaluation

The institutional-social domain concerns reputation, scarcity, market position, community recognition, exhibition context, conservation, and cultural impact. These factors are not embarrassing contaminants that can simply be removed from art. They help determine which works receive attention, resources, and historical continuity. The Villers and Leyster cases demonstrate, however, that institutional signals can reproduce hierarchy and then present the hierarchy as evidence of quality. AI-generated and AI-assisted art enters a field in which prestige is already unequally distributed among artists, galleries, media, regions, and identities. Institutions should therefore report social and market judgments as such rather than naturalizing them. A high auction estimate measures anticipated demand under particular provenance conditions; it does not directly measure formal excellence. The clearer this distinction becomes, the less easily price and reputation can impersonate aesthetic truth.

Institutional Reforms for a Post-Binary Artworld

A two-stage evaluation procedure offers the most direct reform. In the first stage, jurors or reviewers assess formal-perceptual qualities without creator names or AI disclosure, using stable criteria appropriate to the medium. In the second stage, they receive provenance, process documentation, conceptual framing, and ethical information, then evaluate the remaining domains. The difference between stages becomes informative rather than embarrassing. A substantial decline after disclosure may be justified by rule violations or ethical concerns, but it may also reveal a provenance penalty that requires explanation. Museums, competitions, publishers, and classrooms can preserve both scores instead of replacing the first with the second. Such a system does not demand that final decisions be blind; it makes visible how context changed them.

Binary labels should be replaced by contribution records that describe the workflow at a meaningful level of granularity. A concise provenance statement might identify the model and version, the source or licensing status of custom data, the number or structure of iterations, major human edits, compositing or fabrication, and the person responsible for final selection. The goal is not bureaucratic surveillance of every brushstroke or prompt. It is to distinguish processes that currently collapse into the same phrase. Contribution records also protect artists from false accusation because they provide affirmative evidence of a workflow rather than requiring stylistic innocence. They permit institutions to create categories around actual practices, such as generative image synthesis, human-AI compositing, algorithmic installation, or camera-based photography with machine editing. Clear categories reduce the incentive to treat "AI" as a contagious aesthetic condition.

Criticism and education require parallel reforms in vocabulary. Students and reviewers should be trained to label the evidentiary status of their claims: perceptual observation, contextual inference, ethical judgment, or market prediction. Terms such as soul, authenticity, and humanity can remain valuable, but critics should explain what those terms denote and how the work supports the inference. Reverse-image search, provenance verification, process interviews, and uncertainty statements should become ordinary components of digital connoisseurship. The Monet incident showed that expertise was not useless; several experts recognized the work's plausible late-Monet characteristics and resisted the prompt's framing. The task is to institutionalize that resistance to framing rather than celebrate occasional individual success. Critical seeing in the generative era must include critical awareness of labels.

Creative institutions should also design against homogenization without turning that risk into a reason for exclusion. The diversity costs observed by Doshi and Hauser (2024) suggest that early-stage reliance on a single model can narrow a collective idea space, while Wan and Kalman (2026) indicate that diversified model personas and structured variation can mitigate convergence. Programs can require artists to document sources of variation, compare multiple systems, introduce nonmodel research, delay AI use until after independent ideation, or use AI for critique and selection rather than initial generation. Such strategies treat AI as a design variable rather than an all-or-nothing identity. They also preserve a role for human difference, local knowledge, embodied research, and idiosyncratic risk. The artworld's future depends less on banning generative systems than on cultivating practices that resist their defaults. A medium becomes artistically mature when practitioners can criticize and transform its conventions, not merely adopt or reject it.

Limits, Counterarguments, and Ethical Distinctions

The SHL0MS intervention must be interpreted within its limitations. It was a social media performance, not a preregistered experiment, and respondents self-selected into a public conversation structured by provocation, platform incentives, cropping, uncertain display conditions, and the possibility of ironic participation. Viral summaries preserve the most dramatic errors, while deleted comments and algorithmic amplification make the denominator impossible to establish. Some experts identified the image as credible Monet, which confirms that careful looking and domain knowledge can overcome the label. The incident therefore cannot establish the prevalence or magnitude of AI prejudice. Its value is diagnostic and illustrative: it demonstrates the logical possibility of label-driven formal reversal in a culturally consequential setting. Controlled studies supply the causal and population-level support that the anecdote lacks.

A second limitation concerns the word *prejudice*. In this article, the term denotes a categorical evaluative penalty imposed before or beyond artifact-specific evidence. It does not imply that artificial systems possess human rights, suffer discrimination, or occupy a moral position equivalent to women and other historically marginalized groups. The comparison with gendered reattribution is structural, not experiential. Both cases show how status beliefs can direct attention and valuation, but the harms, histories, and ethical subjects differ profoundly. Human artists, workers, and communities can be harmed by AI systems, by extractive training practices, by market concentration, and by false accusations of machine use. A defensible critique must therefore oppose both ungrounded AI stigma and uncritical technological adoption.

DOI: 10.5281/zenodo.21788635

Precision about the target of criticism is the condition of that balance.

Finally, context sometimes should change value. A forged signature, deceptive submission, nonconsensual dataset, or violation of a medium-specific rule supplies relevant information about trust, originality, and institutional fairness. The claim advanced here is not that an object must retain the same total value after every reattribution. It is that the reasons for changed value should be stated in the correct domain. When a real photograph is disqualified from an AI category, the reason is eligibility, not aesthetic failure. When an AI-assisted image is criticized for unauthorized training, the reason is procedural ethics, not automatically composition. When a Monet is called incoherent only under an AI label, the reason is likely attribution-conditioned perception, not altered craft. The artworld will become more rigorous when it stops asking one undifferentiated question, "Is it good?", and begins asking which kind of good, under which evidence, for which purpose, and at whose cost.

Conclusion

The Monet test exposes a fault line in contemporary aesthetic judgment. Viewers who believed they were confronting AI converted an authorship label into detailed formal defects, even though the represented painting belonged to Monet's established oeuvre. Controlled research shows that this was not merely an online curiosity: AI labels repeatedly reduce ratings of creativity, profundity, skill, beauty, and value when artifacts are identical or difficult to distinguish. Blind evaluation, meanwhile, has often favored AI poetry, machine humor, and AI-assisted stories, although humans retain important advantages in exceptional divergent creativity and in accountable cultural agency. Historical reattributions involving Villers and Leyster demonstrate that the same artworld has long allowed fame and gender to reorganize the perception and price of stable objects. Craft and quality matter, but they do not independently govern value. Attribution, effort, identity, institutional position, and presumed humanity participate in the judgment from the beginning.

The appropriate future is neither provenance blindness nor provenance captivity. Art requires context, and ethical creative culture requires transparent process, yet neither requirement justifies allowing labels to impersonate visual evidence. A four-domain model, combined with staged blind and disclosed evaluation, can preserve formal judgment, historical interpretation, procedural accountability, and institutional analysis without collapsing them. Contribution-sensitive provenance can also reconcile AI with human creativity by showing where human agency resides across direction, selection, editing, fabrication, framing, and responsibility. The artworld does not need to decide whether machines are human in order to evaluate work made with them. It needs criteria capable of recognizing distributed creativity while naming exploitation, deception, and weak form precisely. The central question moving forward is therefore not whether AI has a soul. It is whether art institutions can develop the self-awareness to distinguish what they see from what they expected to see.

References

1. Becker, H. S. (1982). *Art worlds*. University of California Press.
2. Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., Christensen, A., & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications, 8*, Article 42. https://doi.org/10.1186/s41235-023-00499-6
3. Benzine, V. (2024, April 3). Art bites: How the art world rediscovered Judith Leyster. *Artnet News*. https://news.artnet.com/art-world/art-bites-judith-leyster-rediscovery-2457671
4. Bourdieu, P. (1993). *The field of cultural production: Essays on art and literature*. Columbia University Press.
5. Brieber, D., Nadal, M., & Leder, H. (2015). In the white cube: Museum context enhances the valuation and memory of art. *Acta Psychologica, 154*, 36–42. https://doi.org/10.1016/j.actpsy.2014.11.004
6. 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*, Article 107406. https://doi.org/10.1016/j.chb.2022.107406
7. Cho, H., & Schwarz, N. (2008). Of great art and untalented artists: Effort information and the flexible construction of judgmental heuristics. *Journal of Consumer Psychology, 18*(3), 205–211. https://doi.org/10.1016/j.jcps.2008.04.009
8. Cleeremans, A., Ginsburgh, V., Klein, O., & Noury, A. (2016). What's in a name? The effect of an artist's name on aesthetic judgments. *Empirical Studies of the Arts, 34*(1), 126–139. https://doi.org/10.1177/0276237415621197
9. Cole, S. (2023, January 6). "I don't believe you": Artist banned from r/Art because mods thought they used AI. *Vice*. https://www.vice.com/en/article/artist-banned-from-art-reddit/
10. Danto, A. C. (1964). The artworld. *The Journal of Philosophy, 61*(19), 571–584. https://doi.org/10.2307/2022937
11. de Rooij, A. (2025). Bias against artificial intelligence in visual art: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000833
12. Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28), Article eadn5290. https://doi.org/10.1126/sciadv.adn5290
13. Gorenz, D., & Schwarz, N. (2024). How funny is ChatGPT? A comparison of human- and AI-produced jokes. *PLOS ONE, 19*(7), Article e0305364. https://doi.org/10.1371/journal.pone.0305364
14. Grassini, S., & Koivisto, M. (2024). Understanding how personality traits, experiences, and attitudes shape negative bias toward AI-generated artworks. *Scientific*

DOI: 10.5281/zenodo.21788635

Reports, 14, Article 4113. https://doi.org/10.1038/s41598-024-54294-4

15. Harrison Dupré, M. (2026, May 15). Devious prankster posts real Monet painting, tells people it's AI-generated, and watches the chaos unfold. *Futurism*. https://futurism.com/artificial-intelligence/real-monet-ai-chaos

16. Hernando, E., & Campo, S. (2017). Does the artist's name influence the perceived value of an art work? *International Journal of Arts Management, 19*(2), 46–58.

17. Hitsuwari, J., Ueda, Y., Yun, W., & Nomura, M. (2023). Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry. *Computers in Human Behavior, 139*, Article 107502. https://doi.org/10.1016/j.chb.2022.107502

18. Hofrichter, F. F. (1989). *Judith Leyster: A woman painter in Holland's Golden Age*. Davaco.

19. Horton, C. B., Jr., White, M. W., & Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports, 13*, Article 19001. https://doi.org/10.1038/s41598-023-45202-3

20. Hutson, J. (2025). Human-AI collaboration in writing: A multidimensional framework for creative and intellectual authorship. *International Journal of Changes in Education*. Advance online publication. https://doi.org/10.47852/bonviewIJCE52024908

21. Hutson, J., Lively, J., Robertson, B., Cotroneo, P., & Lang, M. (2024). *Creative convergence: The AI renaissance in art and design*. Springer. https://doi.org/10.1007/978-3-031-45127-0

22. Hutson, J., & Plate, D. (2024). The algorithm of fear: Unpacking prejudice against AI and the mistrust of technology. *Journal of Innovation and Technology, 2024*, Article 38. https://doi.org/10.61453/joit.v2024no38

23. Katz, L. (2024, June 13). AI photo contest winner disqualified because it's real. *Forbes*. https://www.forbes.com/sites/lesliekatz/2024/06/13/real-photo-wins-ai-photography-contest/

24. Kruger, J., Wirtz, D., Van Boven, L., & Altermatt, T. W. (2004). The effort heuristic. *Journal of Experimental Social Psychology, 40*(1), 91–98. https://doi.org/10.1016/S0022-1031(03)00065-9

25. Lau, E. (2023, April 18). Artist reveals winning image was created by AI for Sony World Photography Awards. *The National*. https://www.thenationalnews.com/arts-culture/art-design/2023/04/18/artist-reveals-winning-image-was-created-by-ai/

26. Lee, B. C., & Chung, J. (2024). An empirical investigation of the impact of ChatGPT on creativity. *Nature Human Behaviour, 8*, 1906–1914. https://doi.org/10.1038/s41562-024-01953-1

27. Metropolitan Museum of Art. (2023). *Museums without men: Marie Denise Villers* [Audio tour transcript]. https://www.metmuseum.org/perspectives/katy-hessel-audio-tour-marie-denise-villers-transcript

28. Newman, G. E., & Bloom, P. (2012). Art and authenticity: The importance of originals in judgments of value. *Journal of Experimental Psychology: General, 141*(3), 558–569. https://doi.org/10.1037/a0026035

29. Nochlin, L. (1988). Why have there been no great women artists? In *Women, art, and power and other essays* (pp. 145–178). Harper & Row. (Original work published 1971)

30. Oppenheimer, M. A. (1996). Nisa Villers, née Lemoine (1774–1821). *Gazette des Beaux-Arts, 127*(1527), 167–180.

31. Porter, B., & Machery, E. (2024). AI-generated poetry is indistinguishable from human-written poetry and is rated more favorably. *Scientific Reports, 14*, Article 26133. https://doi.org/10.1038/s41598-024-76900-1

32. SHL0MS. (2026, May 12). *I just generated an image in the style of a Monet painting using AI* [Post]. X. https://x.com/SHL0MS/status/2054280631807316329

33. Shepherd, T. (2023, July 11). Woman's iPhone photo of son rejected from Sydney competition after judges ruled it could be AI. *The Guardian*. https://www.theguardian.com/australia-news/2023/jul/11/mothers-iphone-photo-of-son-rejected-from-sydney-competition-after-judges-ruled-it-could-be-ai

34. Sterling, C. (1951). A fine "David" reattributed. *The Metropolitan Museum of Art Bulletin, 9*(5), 121–132. https://doi.org/10.2307/3257483

35. Vosevich, K., & Hutson, J. (2024). Absent presence: The human influence in AI-generated content in the age of technoculture. *International Journal of Emerging and Disruptive Innovation in Education: VISIONARIUM, 2*(1), Article 3. https://doi.org/10.62608/2831-3550.1025

36. Walton, K. L. (1970). Categories of art. *The Philosophical Review, 79*(3), 334–367. https://doi.org/10.2307/2183933

37. Wan, Y., & Kalman, Y. M. (2026). Diverse AI personas can mitigate the homogenization effect in human-AI collaborative ideation. *Computers in Human Behavior: Artificial Humans, 8*, Article 100289. https://doi.org/10.1016/j.chbah.2026.100289

38. Wang, D., Huang, D., Shen, H., & Uzzi, B. (2026). A large-scale comparison of divergent creativity in humans and large language models. *Nature Human Behaviour, 10*, 531–540. https://doi.org/10.1038/s41562-025-02331-1

39. Wang, Y., Leng, X., He, G., & Lin, H. (2026). When labels matter more: Behavioral and eye-tracking evidence on aesthetic judgment of AI-generated and human-created paintings by lay viewers. *Empirical Studies of the Arts*. Advance online publication. https://doi.org/10.1177/02762374261441560

DOI: 10.5281/zenodo.21788635

40. Zhang, W., Xie, C., Jiang, L., Yang, L., Hu, Z., & Hao, N. (2025). Neural correlates of evaluative bias against artificial intelligence-labeled versus human-labeled artworks. *Social Cognitive and Affective Neuroscience, 20*(1), Article nsaf071. https://doi.org/10.1093/scan/nsaf071

DOI: 10.5281/zenodo.21788635