

Progress in Aesthetic Models: From Formal Regularities to Situated Multilevel Valuation

Qi Li (李琪)*, Chaoyuan Zhang, Qi Li (李琦)

School of Fine Arts, Shanxi University, Taiyuan, Shanxi, 030006, China

*Corresponding author: Qi Li (李琪), liam@sxu.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Aesthetic models have expanded from accounts centered on formal order, prototypicality, complexity, and processing fluency to multilevel explanations involving predictive dynamics, affective valuation, self-related meaning, expertise, cultural context, and computational representation. This structured integrative review synthesizes foundational and recent evidence from empirical aesthetics, cognitive psychology, neuroaesthetics, and computational image assessment. The evidence shows that no single tradition supersedes the others: formal variables constrain perception; recurrent cognitive operations transform perception into appraisal; distributed neural systems support sensation, valuation, emotion, memory, and meaning; and context and individual differences determine which criteria are applied. Computational models have improved large-scale prediction, distributional scoring, and personalization, but predictive performance alone does not identify psychological mechanisms. To connect these levels, the review proposes a Situated Multilevel Aesthetic Model comprising object representation, predictive processing, affective valuation, self-meaning integration, and social-institutional context. The model treats aesthetic outcomes and their uncertainty as time-dependent functions of stimulus, viewer, and situation. Methodological priorities include construct-specific outcomes, factorial context manipulation, time-resolved measurement, multilevel and cross-cultural sampling, preservation of disagreement in datasets, and intervention-based validation of explanations. The resulting framework positions aesthetic modeling as an evidence-integrated enterprise in which psychological theory defines constructs, neuroscience constrains process, computation tests generalization, and humanistic inquiry identifies the institutional conditions of value.

Keywords: Aesthetic Experience; Empirical Aesthetics; Neuroaesthetics; Processing Fluency; Predictive Processing; Computational Aesthetics; Aesthetic Valuation; Contextual Aesthetics

Published: Sep 22, 2026

DOI: <https://doi.org/10.62177/chst.v3i3.1701>

1. Introduction

Aesthetic experience has moved from the margins of philosophical speculation to a productive meeting point for psychology, neuroscience, design research, and computer vision. Yet the apparent unity of the term “aesthetic model” conceals several different explanatory projects. Some models ask which formal properties make an object pleasing. Others describe the temporal stages through which perception becomes evaluation. Neuroaesthetic accounts identify interacting sensory, valuation, and meaning systems, while computational models predict ratings from large collections of images. Progress therefore cannot be measured by predictive accuracy alone. It must also be evaluated by whether a model distinguishes object

properties from processing dynamics, individual differences, cultural context, affective meaning, and uncertainty.

Early empirical traditions treated aesthetic preference as a lawful response to regularity, prototypicality, novelty, complexity, or ease of processing. Work on prototypicality and musical syntax showed that preference is linked to how stimuli instantiate learned structures^[1], while processing-fluency theory explained pleasure partly through the positively valenced experience of easy perception^[2]. Leder and colleagues then proposed a staged model in which perceptual analysis, implicit memory integration, explicit classification, cognitive mastering, and evaluation jointly produce aesthetic judgment and emotion^[3]. The “most advanced, yet acceptable” principle likewise made preference a relation between novelty and typicality rather than an intrinsic feature of an object^[4]. Prediction-error accounts subsequently argued that reward can arise from the active reduction of uncertainty when a viewer discovers structure in art^[5]. Across these theories, the central unit of explanation gradually shifted from static beauty to a transaction between an organized stimulus and a situated perceiver.

This shift became more explicit in neuroaesthetics. The aesthetic triad organized evidence around sensory-motor, emotion-valuation, and meaning-knowledge systems^[6]; later revisions of the cognitive-stage model emphasized feedback, expertise, context, and multiple possible outcomes of the aesthetic episode^[7]. Broad syntheses warned that aesthetic experience is distributed across ordinary cognitive and affective systems rather than localized in a single “beauty center”^[8]. The Vienna Integrated Model of Art Perception (VIMAP) formalized reciprocal top-down and bottom-up processing, discrepancy, self-related change, and transformative outcomes^[9], while the Distancing-Embracing model explained how art contexts allow negative emotion to contribute to enjoyment^[10]. More recently, neuroimaging reviews^[11] and the sensory valuation account^[12] have reframed aesthetic appreciation as valuation of sensory information under contextual and organismic constraints.

At the same time, computational image aesthetics has expanded from hand-crafted feature engineering to deep, probabilistic, personalized, and vision-language architectures. This literature can predict average ratings and score distributions at unprecedented scale, but it often operationalizes aesthetics as an output label rather than explaining why a judgment occurs. A review of computational neuroaesthetics identifies the conceptual gap between brain-based accounts and computer-science benchmarks^[13]. The challenge is not to reject computation, but to specify what should be computed, for which observer and context, and at which explanatory level.

This article provides a structured integrative review of the progress of aesthetic models from formal and psychophysical accounts to process, neural, contextual, and computational frameworks. It advances three arguments. First, no single tradition has replaced its predecessors; each captures a different level of the aesthetic episode. Second, the most consequential advance is the shift from object-centered prediction to relational explanation involving time, viewer, and context. Third, the next generation of aesthetic models should be situated, multilevel, uncertainty-aware, and explanation-oriented. To make this proposal concrete, the review concludes with a Situated Multilevel Aesthetic Model (SMAM) and a validation agenda linking theory to behavioral, physiological, neural, and computational evidence.

2. Review Scope and Method

This is a structured integrative review rather than a formal meta-analysis. The objective is conceptual synthesis across literatures that use different outcomes, stimuli, and methods, not estimation of a single pooled effect. Searches were conducted for publications available through August 2026 using combinations of the terms aesthetic model, aesthetic appreciation, empirical aesthetics, neuroaesthetics, sensory valuation, predictive processing, image aesthetics assessment, computational aesthetics, personalized aesthetics, context, expertise, aesthetic emotion, transformative experience, and aesthetic judgment. Candidate records were identified through journal and publisher pages, Crossref metadata, backward reference tracing from major reviews, and forward-oriented searches for recent high-level studies. Priority was given to peer-reviewed articles in internationally indexed psychology, cognitive neuroscience, and art-perception journals, supplemented by technically foundational computer-vision papers when necessary to explain the computational transition.

Studies were retained when they made one of four contributions: they proposed or substantially revised an explanatory model; provided empirical evidence capable of discriminating among models; introduced a benchmark or architecture that changed the operationalization of aesthetic prediction; or tested how contextual information, expertise, or individual differences change aesthetic response. Work was excluded when “aesthetic” referred only to surface styling without a human judgment

outcome, when no clear theoretical or empirical contribution could be identified, or when bibliographic identity could not be verified. Every DOI in the final evidence base was programmatically queried against Crossref, and the returned title, author list, venue, year, volume, issue, and pagination or article number were compared with the cited record.

The review uses an evidence-layer logic. Conceptual models are assessed for construct coverage and causal clarity; behavioral studies for task design, stimulus control, and individual variation; neural studies for spatial and temporal inference boundaries; and computational studies for target definition, dataset structure, generalization, and interpretability. This prevents a common category error: high predictive performance at one level does not validate a psychological mechanism at another. A model that predicts average ratings may be useful engineering even if it does not explain individual aesthetic experience, just as a reliable neural correlate does not by itself identify a unique aesthetic computation.

The period covered is deliberately broad. Foundational work is required because contemporary models still inherit assumptions about fluency, typicality, novelty, and staged processing. However, the critical emphasis falls on developments from 2014 onward, particularly integrative neuroaesthetic theories, contextual and transformative accounts, large-scale computational datasets, and personalized assessment. The synthesis is organized by changes in the unit of analysis: from object, to processing episode, to distributed valuation, and finally to person-context interaction.

Table 1. Comparative architecture of major aesthetic-model traditions

Tradition	Primary unit	Main strength	Central limitation
Formal and psycho-physical	Object-feature relations	Controlled, testable variables	Context and meaning under-specified
Cognitive-process	Temporal aesthetic episode	Separates perception, mastery, emotion, judgment	Stage boundaries can be task-dependent
Neuroaesthetic	Distributed neural configuration	Constrains sensory, valuation, and self-related processes	Correlates do not uniquely identify mechanisms
Contextual and expertise	Person-object-situation relation	Explains criterion and viewer variability	Requires multilevel sampling
Computational	Predictive representation	Scales to large datasets and rating distributions	Prediction can outrun explanation
SMAM synthesis	Five recurrent interacting levels	Links constructs, time, uncertainty, and context	Requires prospective validation

3. From Object Properties to Relational Regularities

The earliest scientific ambition in aesthetics was to convert beauty into a measurable relation among stimulus properties. Although such approaches are often portrayed as naive formalism, they established two lasting ideas. First, aesthetic judgments exhibit statistical regularities rather than being wholly arbitrary. Second, these regularities are relational: a feature becomes effective through the perceptual and cultural system that receives it. Prototypicality research demonstrated that preference may increase when a stimulus fits an acquired structural schema ^[1]. Processing fluency later generalized this insight by proposing that ease of identification, integration, or interpretation carries positive affect that can be attributed to the object ^[2]. These accounts remain influential because they generate manipulable predictions: symmetry, contrast, repetition, familiarity, or figure-ground clarity can alter processing ease, which can in turn alter liking.

Yet fluency is not a universal law of simplicity. Excessively predictable stimuli can be boring, while difficulty can be rewarding when it promises insight. A dual-process account distinguishes the immediate positive response to fluent processing from more reflective pleasure in successful elaboration ^[14], and appraisal research shows that interest depends on novelty-complexity together with perceived comprehensibility ^[15]. The industrial-design principle of being “most advanced, yet acceptable” captures this balance: novelty is attractive when it remains intelligible relative to a category ^[4]. Prediction-error accounts deepen the claim by treating aesthetic reward as progress in sense-making rather than low error per se ^[5,16]. The satisfying moment may occur when an initially surprising object becomes compressible, when a hidden regularity

is discovered, or when a viewer's model is productively revised. The aesthetic-Aha effect provides direct evidence that perceptual insight can produce a temporally localized increase in liking ^[17]. This distinction between predictability and prediction improvement helps reconcile why both familiar and challenging art can be valued.

Object-centered theories also face the problem of multidimensionality. Formal features do not act independently: complexity can be pleasurable under expertise but aversive under cognitive load; novelty can invite exploration when the setting is safe but withdrawal when it signals threat; symmetry may support rapid organization but conflict with expressive or genre-specific goals. The same feature therefore has no stable aesthetic sign outside task, category, expectation, and viewer. This does not make formal analysis irrelevant. It means that formal properties should be modeled as inputs whose effects are moderated by learned priors and current goals.

The strongest legacy of psychophysical aesthetics is thus methodological rather than reductionist. Formal theories encouraged controlled manipulation and quantitative measurement. Contemporary models should preserve that discipline while rejecting the assumption that low-level features determine value directly. A useful formal variable must be embedded in a process account specifying when it is extracted, how it interacts with expectations, and which response it influences. For example, edge density may increase perceptual load, semantic richness, or both; color contrast may facilitate segmentation while simultaneously changing arousal. Without an intermediate process model, the same feature can support incompatible post hoc explanations.

The transition from object measures to relational regularities is the first major advance in aesthetic modeling. Beauty is not a scalar stored in the artwork. It is an evaluation generated when stimulus structure encounters perceptual capacities, learned categories, prediction dynamics, and evaluative goals. A dynamic account of appreciation similarly argues that repeated elaboration can change evaluation over time ^[18]. This relational premise becomes the foundation for process, neural, and computational models, even when later work does not state it explicitly.

4. Process Models: From Stages to Recurrent Aesthetic Episodes

The staged model of aesthetic appreciation proposed by Leder and colleagues gave empirical aesthetics a process architecture ^[3]. Its importance lies less in a rigid sequence than in the separation of operations that are often collapsed into a single rating. Perceptual analysis concerns formal organization; implicit memory integration concerns familiarity and prototypicality; explicit classification concerns style and content; cognitive mastering concerns interpretation; and evaluation relates the outcome to goals and satisfaction. The model also distinguishes aesthetic emotion from aesthetic judgment, allowing an object to be evaluated as artistically important without being pleasant, or enjoyed affectively without being judged technically accomplished.

The ten-year reassessment of the model made several necessary revisions ^[7]. Expertise was treated not as a demographic nuisance but as a source of qualitatively different processing. Context was recognized as constitutive rather than peripheral. The aesthetic episode became more recurrent, with feedback among classification, mastering, affect, and evaluation. These changes weakened a strictly serial interpretation and opened the model to interactions between bottom-up features and top-down knowledge. They also clarified that a rating obtained after a few seconds is a compressed report on a temporally extended episode.

VIMAP expanded this dynamic view by explicitly modeling mismatches, control, self-relevance, and transformative outcomes ^[9]. Instead of assuming that successful processing ends in pleasure, it allows disruption, ambiguity, rejection, insight, and schema change. The model is especially useful for modern and contemporary art, where difficulty or violation is often intentional. Its central contribution is to represent aesthetic experience as a trajectory through states rather than a single endpoint. This trajectory can include initial perceptual reaction, cognitive discrepancy, attempts at reclassification, affective change, and eventual transformation or disengagement.

The Distancing-Embracing model addresses a complementary problem: why people seek sadness, fear, disgust, or moral conflict in art ^[10]. Psychological distance marks the experience as occurring in an art frame, reducing direct action demands. Embracing mechanisms - including compositional pleasure, meaning construction, mixed emotion, and genre expectations - then allow negative affect to contribute to enjoyment. This framework shows why valence alone is an inadequate target for

aesthetic models. A negative emotional response may be evidence of failure in one context and deep engagement in another. Recent work further links aesthetic experience to learning and cognition. The “stopping for knowledge” hypothesis treats beauty as a signal that temporarily interrupts action in favor of perceptual learning^[19]. The wider cognitive-science agenda argues that aesthetics offers a privileged domain for studying interactions among attention, emotion, prediction, memory, and social meaning^[20]. Research on crying and transformation demonstrates that intense aesthetic response can combine negative affect, insight, and self-change^[21]. A systematic review of transformative aesthetic experiences similarly finds that transformation is not reducible to immediate liking and requires attention to self-related and longer-term outcomes^[22]. Aesthetic cognitivism adds that art may promote understanding by reorganizing concepts, perspective, or social knowledge^[23]. These developments redefine model adequacy. A model should not be judged solely by whether it predicts beauty ratings. Component research distinguishes aesthetic fascination, appraisal, and emotion^[24], while theoretical and measurement work demonstrates that aesthetic emotions include epistemic, prototypical, pleasing, negative, and self-forgetful responses^[25-26]. A model should therefore specify which outcome is being modeled: perceptual pleasure, interest, liking, meaningfulness, being moved, insight, memorability, willingness to revisit, or transformative impact. Different outcomes may occur at different times and may even diverge. The progress of process models lies in replacing a one-dimensional aesthetic response with an evolving episode that can support multiple endpoints.

5. Neuroaesthetics: Distributed Valuation without a Beauty Center

Neuroaesthetics initially gained visibility through studies identifying neural correlates of beauty. Early functional imaging linked beauty judgments to reward- and valuation-related activity^[27], while prefrontal activation was reported during visual aesthetic perception^[28]. Research on judgments of abstract patterns implicated networks involved in evaluation and cognitive decision-making^[29]. Studies of classical sculpture showed that proportion and explicit judgment recruit partly overlapping but distinguishable systems^[30], and viewing artworks under different instructions demonstrated contributions of both perceptual facilitation and cognitive control^[31]. These studies were foundational, but the field’s maturation has depended on moving beyond localization.

The aesthetic triad is a major step in this direction^[6]. It treats aesthetic experience as interaction among sensory-motor systems, emotion-valuation systems, and meaning-knowledge systems. The account is compatible with multiple art forms and with the observation that expertise or context can change how the same input is processed. Broader reviews describe neuroaesthetics as the cognitive neuroscience of aesthetic experience rather than a search for a dedicated beauty module^[8]. Quantitative synthesis finds partially consistent engagement of valuation and perceptual networks when people view paintings^[32], while cross-modal meta-analysis suggests that aesthetic appraisal shares components across vision, audition, and other sensory domains^[33]. Research on architectural contour further demonstrates that perceptual organization can influence both aesthetic judgment and approach-avoidance decisions^[34], linking formal structure to embodied response.

Research on the default-mode network adds an important self-related dimension. Intense responses to art can recruit default-mode regions that are normally suppressed by externally oriented tasks^[35]. Follow-up work argues that art can reach within by linking perceptual input to personal relevance and internally directed cognition^[36]. A later study showed that default-mode representations of aesthetic appeal generalize across visual domains^[37]. These results support a model in which strong aesthetic experience emerges when external sensory organization becomes integrated with self-related meaning. They do not imply that the default-mode network is uniquely aesthetic; rather, aesthetic episodes exploit general systems for self-referential and integrative cognition.

The field has also confronted conceptual overreach. The argument for a “farewell to art” in psychology and neuroscience calls for studying aesthetic valuation across ordinary objects and sensory events rather than treating canonical art as the defining category^[38]. This critique improves model generality but creates a counter-risk: if every sensory preference becomes aesthetic, the concept loses discriminative power. A productive resolution is to distinguish basic sensory valuation from culturally framed aesthetic episodes. The sensory valuation account makes this distinction tractable by grounding aesthetic value in general valuation operations while allowing context, learning, and goals to shape the evaluated sensory information^[12]. Neuroimaging reviews likewise stress that apparent aesthetic regions participate in domain-general valuation, attention,

emotion, and meaning^[11].

This distributed view changes what neural evidence can establish. Activation in orbitofrontal or default-mode regions cannot by itself prove beauty, depth, or artistic quality. EEG alpha or gamma changes cannot be assigned a fixed aesthetic meaning without reference to task, timing, scalp distribution, and behavioral convergence. Neural measures are most informative when they discriminate process predictions: whether context acts early or late, whether insight follows uncertainty reduction, and whether expertise changes representational integration. Neural data should constrain aesthetic models, not decorate them.

The strongest current position is therefore neither neuro-reductionist nor anti-neural. Aesthetic experience is realized through distributed systems that support sensation, action, valuation, emotion, memory, self-processing, and knowledge. The distinctive phenomenon lies in their configuration over time under an aesthetic stance. Work connecting aesthetic emotion to psychological well-being illustrates why the downstream consequences of this configuration also matter^[39]. The configuration is compatible with the sensory valuation account^[12], the triadic framework^[6], recurrent process models^[7,9], and learning-centered theories^[19]. Their convergence suggests that the next advance will come from specifying interactions and temporal transitions rather than naming additional regions.

6. Context, Expertise, Meaning, and the Situated Viewer

Context research provides some of the clearest evidence against object-only models. Information about an artist, title, gallery, authenticity, or interpretive background can alter evaluation even when the visual stimulus remains unchanged. Experiments with works attributed to Max Ernst and Pablo Picasso found that contextual information changes aesthetic appreciation^[40]. Neuroimaging evidence indicates that object-context relationships modulate aesthetic judgment, demonstrating that contextual effects are not merely post hoc verbal rationalizations^[41]. More recent work shows that context and painting attributes interact with expertise^[42], while appreciation context can influence both aesthetic evaluation and perceived duration^[43]. Theory-of-mind engagement can increase appreciation by encouraging viewers to infer intentions behind visual works^[44].

These findings motivate a distinction among at least four contexts. Perceptual context includes surrounding stimuli, display scale, lighting, and temporal sequence. Semantic context includes titles, explanations, genre categories, and curatorial narratives. Social context includes co-viewers, group identity, education, and culturally shared interpretive conventions. Institutional context includes museums, markets, platforms, and educational settings that authorize particular modes of attention. Each can modify the priors used to interpret an object, the criteria used to judge it, or the value attached to the outcome.

Expertise changes more than the magnitude of liking. Experts possess richer categories, stronger style schemas, and different attentional routines. They can tolerate ambiguity because they have more hypotheses available for resolving it. They may also separate technical quality, historical importance, originality, and personal pleasure more readily than novices. Evidence from everyday designed objects connects aesthetic life experience and expertise to both judgment and emotion^[45]. The interaction between context and expertise^[42] therefore supports a hierarchical model: low-level organization remains available to all viewers, but higher-level interpretation depends on learned representational resources. A universal model that omits expertise may predict population averages while misunderstanding the process that generates them.

Dynamic and embodied approaches extend context into time. Ecological research shows that laboratory control and real-world art experience need not be treated as mutually exclusive^[46], while direct comparison of museum and laboratory settings demonstrates that place changes the quality of aesthetic engagement and memory^[47]. Embodied work on abstract art further highlights bodily and action-related dimensions that conventional ratings overlook^[48]. These results disaggregate outcome dimensions and contexts: a feature that supports immediate perceptual organization may have a different relation to later interest, understanding, or insight. “Preference” therefore cannot stand in for the whole aesthetic episode.

The situated viewer also carries stable and temporary conditions. Need for cognition is associated with aesthetic competence and may mediate how experience develops into knowledge^[49]. Current fatigue, stress, mood, cultural identity, and task goals can change the resources available for interpretation. A unifying account of visual experience shows that recurrent perceptual constraints and cultural context can be represented within the same model rather than treated as competing explanations^[50]. These variables are often treated as noise because they complicate prediction, but they are theoretically central. A model of

human aesthetic experience should represent between-person variation and within-person state variation rather than averaging them away.

Meaning is the most difficult component to operationalize, yet it cannot be omitted. Meaning includes recognized content, cultural symbolism, personal association, narrative implication, and conceptual understanding. It can increase value even when perceptual fluency is low, and it can reverse initial responses. Process models describe meaning construction as cognitive mastering ^[3,7], VIMAP describes it as possible schema change ^[9], default-mode findings connect it to self-relevance ^[35-37], and aesthetic cognitivism connects it to understanding ^[23]. Theoretical work on the self as an aesthetic effect adds that aesthetic encounters can participate in the organization of subjectivity itself ^[51]. The convergence suggests that meaning is not an optional interpretive layer added after perception; it is a recurrent source of prediction and valuation.

7. Computational Aesthetics: From Feature Scores to Personalized Representations

Computational aesthetics made aesthetic judgment scalable. The AVA database established a large benchmark linking images to distributions of human ratings and photographic style or content labels ^[52]. AADB added attributes and content adaptation, supporting models that connect global ratings with interpretable photographic properties ^[53]. NIMA predicted rating distributions rather than only mean scores, preserving some information about disagreement ^[54]. A unified probabilistic formulation further treated aesthetic assessment as an uncertain inference problem ^[55]. These developments are substantive because they change the target from a deterministic beauty label to a distribution generated by multiple viewers.

The next transition was from generic to theme-sensitive and personalized assessment. TAD66K was designed around theme-specific aesthetic criteria and benchmarked models across diverse content categories ^[56]. PARA incorporated rich image attributes and viewer information, enabling personalized image aesthetics assessment ^[57]. Such datasets make it possible to ask whether a model generalizes across themes, individuals, and rating conditions rather than merely fitting a pooled score. They also reveal a conceptual truth that psychological models had long anticipated: disagreement is not always annotation noise. It can reflect legitimate differences in experience, knowledge, preference, or criterion use.

Deep models improve predictive capacity, but their psychological interpretation remains limited. A convolutional or transformer representation may capture composition, color, semantics, genre, and dataset artifacts simultaneously. High correlation with ratings does not identify which of these factors caused a judgment. A concise synthesis of aesthetics as a scientific topic highlights the need to connect pleasure, perception, and judgment rather than reducing them to one score ^[58]. Vision-language systems increase semantic reach; a CLIP-based approach can connect aesthetic assessment to textual concepts and prompts ^[59]. Yet interpretability requires more than a saliency map. Explanations should correspond to constructs that can be independently measured in viewers, such as fluency, novelty, affect, uncertainty, self-relevance, or semantic understanding.

Computational work also broadens the measurement side of the problem. Learned visual representations can capture patterns that are difficult to specify with hand-crafted features, and controlled image transformation can support stronger tests of composition, color, complexity, and style. For aesthetic science, the crucial issue is experimental: computational tools can manipulate visual properties with increasing precision, but technical control and aesthetic explanation remain separate achievements. A system may predict that observers will like an image without containing a psychologically adequate account of why they like it.

The relationship between computational neuroaesthetics and image assessment remains underdeveloped ^[13]. Computer-vision benchmarks usually use static ratings obtained after brief presentation, whereas neural and process theories describe time-varying perception, valuation, and meaning. Bridging the fields requires matched constructs and tasks. For example, uncertainty reduction could be operationalized through time-resolved ratings, viewing duration, eye movements, or neural dynamics; self-relevance could be modeled from individualized semantic associations; and context effects could be tested by holding the visual object constant while changing interpretive information or viewing goals. The goal is not to force every psychological construct into a neural network, but to design predictions that are commensurable across levels.

Several methodological issues limit current benchmarks. First, mean opinion scores conceal multimodality: a mean of five can indicate consensus or polarization. Second, random image splits may allow near-duplicate content or creator style to leak

across training and test sets. Third, ratings collected from a narrow platform population may not generalize across cultures or expertise. Fourth, aesthetic labels often correlate with technical image quality, making it difficult to separate beauty from exposure, sharpness, or professional production. Fifth, model explanations are rarely validated against human reasons. A credible next-generation benchmark should report distributional calibration, cross-domain and cross-cultural generalization, individual-level reliability, and explanatory alignment.

Computational aesthetics has therefore progressed through four stages: hand-crafted formal descriptors, deep feature learning, distributional prediction, and personalized or context-aware modeling. Its next stage should be construct-aware modeling. Rather than asking only which architecture best predicts scores, researchers should ask which theoretically defined information improves prediction, for whom, under which context, and with what uncertainty. This shift would make computational models partners in theory testing rather than isolated engineering systems.

8. Toward a Situated Multilevel Aesthetic Model

The reviewed traditions can be integrated without pretending that they are interchangeable. The proposed Situated Multilevel Aesthetic Model (SMAM) treats an aesthetic judgment as a probabilistic report generated by interactions among five levels: object representation, predictive processing, affective valuation, self-meaning integration, and social-institutional context. The levels are analytically distinct but recurrently coupled.

The object-representation level contains low- and mid-level features, composition, material cues, movement, genre, and semantic content. It corresponds to the strongest domain of formal analysis and computational image models. The predictive-processing level evaluates fluency, novelty, typicality, uncertainty, and progress in resolving structure^[1-2,4-5]. The affective-valuation level integrates reward, arousal, action tendency, negative emotion, and sensory valuation^[6,10-12]. The self-meaning level links the object to memory, knowledge, identity, intention, and understanding^[9,19,21,23,35-37]. The social-institutional level includes cultural norms, co-viewing, museums, markets, educational settings, curatorial framing, and shared conventions^[40-45]. SMAM assumes that the reported response is conditional on viewer and situation. Let y denote an outcome such as liking, beauty, interest, meaningfulness, or willingness to pay. A general formulation is:

$$E(y_{i,t} | x_i, v_j, c_k) = g[O(x_i), P(x_i, v_j, t), A(x_i, v_j, c_k, t), M(x_i, v_j, c_k, t), S(c_k, v_j, t)],$$

where O is object representation, P is predictive processing, A is affective valuation, M is self-meaning integration, S is social-institutional context, v_j represents viewer characteristics, c_k represents context, and t represents time. The function g is not assumed to be linear. Feedback is expected: a title can change semantic interpretation, which changes affect, which redirects attention to different visual features. The model also distinguishes expected response from uncertainty. A complete model should estimate both $E(y)$ and $\text{Var}(y)$, because disagreement may be theoretically meaningful.

The model generates falsifiable predictions. First, image-only models should lose accuracy when interpretive contexts or viewing goals are experimentally manipulated; adding context variables should restore part of that loss. Second, experts and novices should differ not only in average ratings but in which paths connect object features to meaning and value. Third, transformative outcomes should be better predicted by temporal change, self-relevance, and meaning than by initial fluency alone. Fourth, rating distributions should be multimodal for stimuli that activate conflicting cultural schemas. Fifth, cross-cultural transfer should fail selectively at the social and meaning levels while retaining some lower-level perceptual regularities. Sixth, explanations aligned with psychological constructs should generalize better across datasets than unconstrained post hoc feature attributions.

SMAM also clarifies causal inference. Aesthetic datasets are observational and can identify predictive relations but not mechanisms. Controlled experiments are needed to manipulate context, complexity, viewing instruction, or explanatory information. Mediation analysis is appropriate only when temporal order and measurement quality justify it. Neural measures can test when and where a manipulation changes processing, but causal neural claims require intervention or convergent designs. Computational ablation can show that a feature set improves prediction, but not that human viewers use the same representation. Evidence should therefore be triangulated across behavioral manipulation, process measure, neural constraint, and out-of-sample prediction.

The value of SMAM is not that it introduces five entirely new constructs. Its contribution is architectural: it states which

levels must be separated, how they interact, and what kind of evidence can support each connection. It converts the recurring conclusion that “context matters” into a model in which context has identifiable pathways. It also gives computational aesthetics a psychologically interpretable target and gives humanistic criticism a structure for testing claims about agency, meaning, and institutions.

9. Methodological Priorities for the Next Generation

Future research should first improve construct definition. Studies routinely use beauty, liking, aesthetic appeal, quality, creativity, profundity, and value as if they were interchangeable. They are not. Beauty can be high while interest is low; a disturbing work can be disliked yet meaningful; technical quality can be recognized without personal preference; willingness to pay can depend on scarcity or effort rather than experience. Researchers should preregister primary and secondary outcomes and explain how each maps onto the proposed process.

Second, designs should separate stimulus properties from interpretive context. Studies often compare naturally different image sets and attribute all differences to a single formal dimension. A stronger factorial design manipulates a visual property while holding content stable, then independently varies viewing goal or contextual information. This design distinguishes perceptual effects from interpretive effects and their interaction. Contextual experiments ^[40-44] illustrate the value of this separation.

Third, aesthetic episodes should be measured over time. Continuous ratings, response trajectories, eye movements, dwell time, pupil response, EEG, fNIRS, or repeated judgments can reveal whether context changes early attention, later interpretation, or memory. Dynamic feature research ^[45] shows that the same property may have different consequences across viewing phases. Longitudinal designs are also needed to test whether repeated exposure produces learning, fatigue, or durable transformation rather than transient preference.

Fourth, sampling must move beyond convenience populations. Cultural background, art expertise, age, education, and domain familiarity should be modeled, not merely listed. Multilevel models can separate image, viewer, and context variance. Measurement invariance should be tested before comparing cultures or expertise groups. The sensory valuation account ^[12] predicts shared valuation principles, but context studies ^[40-44] predict substantial variation in priors and criteria. Cross-cultural research can determine which components are stable and which are institutionally learned.

Fifth, computational benchmarks should preserve disagreement and collection context. Datasets should report rater-level records, uncertainty, demographic and expertise descriptors where ethically permissible, editing history, presentation conditions, viewing task, and collection interface. Evaluation should include calibration, subgroup performance, cross-theme transfer, cross-dataset transfer, and sensitivity to manipulated context. Distributional methods ^[54-55] and personalized datasets ^[57] are important foundations, but they need stronger psychological annotation.

Sixth, explanation must be validated. Human participants can rate whether model explanations correspond to their reasons, but agreement alone is not sufficient because verbal reports are incomplete. Stronger validation combines reason annotation, experimental manipulation, counterfactual image transformation, and process measures. If a model claims that symmetry drives a judgment, altering symmetry while preserving semantics should change the outcome in the predicted direction. If interpretive context is decisive, changing only that context should produce a measurable effect. Explanations should survive interventions, not merely sound plausible.

Seventh, researchers should protect against analytical flexibility. Aesthetic studies often include many ratings, features, time windows, electrodes, regions, and subgroups. Preregistration, correction for multiple testing, explicit separation of confirmatory and exploratory analyses, effect sizes, uncertainty intervals, and multiverse or robustness analyses are essential. Small samples may be useful for method development, but they cannot support stable individual-difference or high-order interaction claims. High-dimensional neural or computational results should be evaluated on held-out data.

Finally, ethical and humanistic variables must enter the model as objects of research rather than afterthoughts. Museum authority, market valuation, cultural appropriation, unequal access to education, and platform curation can alter which objects are encountered and which criteria are normalized. A situated model can study these effects empirically without reducing ethical judgment to preference. It can also preserve the distinction between explaining why people value an object and

deciding what valuation practices are ethically justified.

10. Critical Implications for Humanistic Social Theory

The progress of aesthetic models carries a broader theoretical implication: value is neither located wholly in objects nor freely projected by isolated subjects. It is produced through relations among material form, embodied capacities, learned expectations, social narratives, and institutions. This position avoids two extremes. Pure formalism ignores the social conditions under which forms become legible; pure constructivism cannot explain why some perceptual organizations and learning dynamics recur across contexts. A multilevel model treats universality and cultural specificity as empirical questions located at different levels.

For critical humanistic social theory, the central task is to examine how aesthetic systems distribute visibility, legitimacy, and value. Museums, schools, markets, media platforms, and disciplinary canons do not merely transmit neutral objects; they organize attention and make some interpretive criteria appear natural. Computational predictors trained on platform ratings can reproduce dominant taste and render it technically authoritative. Personalized systems may broaden access but also enclose individuals within established preference histories. Aesthetic modeling should therefore make institutional conditions explicit, especially when its outputs guide recommendation, curation, education, or cultural funding.

The model also reframes the relation between universality and difference. Evidence for recurrent perceptual preferences need not imply that aesthetic value is culturally invariant. Shared sensory constraints can coexist with historically learned categories, unequal cultural capital, and divergent interpretive practices. Conversely, evidence of contextual variation does not entail that formal organization has no causal role. The analytic task is to locate stability and variability at the appropriate level rather than forcing the whole aesthetic episode into either biological universalism or cultural relativism.

This approach gives aesthetic disagreement positive theoretical status. Disagreement may reflect measurement error, but it may also reveal competing norms, identities, memories, or meanings. Institutions often transform one distribution of judgments into an official value through exhibition, criticism, price, or curricular inclusion. A multilevel aesthetic model can describe how individual responses aggregate, how authority weights some responses more heavily than others, and how repeated institutional exposure changes later perception. These are empirical and normative questions that require cooperation among psychology, sociology, art theory, and computational research.

The humanistic contribution is therefore not a rejection of quantitative modeling. It is a demand for adequate objects of explanation. Ratings, gaze, neural signals, and embeddings are partial traces of an aesthetic relation. Their meaning depends on who produced the object, who judged it, under what description, in which institution, for what purpose, and with what consequences. A scientifically strong model becomes more, not less, humanistic when it represents these conditions and states the boundary of its claims.

11. Limitations of the Present Review

This review has several limitations. It is structured and transparent but not a database-complete systematic review, and it does not estimate pooled effect sizes. The included literatures use heterogeneous stimuli, outcomes, and populations that cannot be summarized by a single quantitative result. The review prioritizes visual art and image aesthetics, although several theoretical models were designed to generalize across music, design, literature, nature, and multisensory experience. Computer-vision conference papers were included only where they introduced field-defining datasets or architectures; their role is technical and is kept distinct from psychological evidence.

The DOI audit verifies bibliographic identity and resolution, not the correctness of every theoretical inference or the permanence of journal indexing status. The literature search is weighted toward English-language publications and may underrepresent models developed in other intellectual traditions. Finally, SMAM is a synthetic framework derived from converging evidence. Its paths and moderators require prospective testing; it should not be presented as an empirically validated final theory.

12. Conclusion

Aesthetic models have progressed through a sequence of enlargements in their unit of explanation. Formal and

psychophysical research identified lawful relations between stimulus structure and preference. Cognitive models reconstructed aesthetic response as a staged and recurrent episode. Neuroaesthetics replaced the search for a beauty center with distributed interactions among perception, valuation, emotion, self, and knowledge. Contextual research demonstrated that learned categories, expertise, and institutional framing participate in the response. Computational aesthetics scaled prediction, preserved rating distributions, and began to model personalization.

The field's central lesson is that predictability is not equivalent to explanation. Image features can predict part of aesthetic judgment, but the judgment is generated by a viewer in time, under a description, within a culture and institution. The proposed Situated Multilevel Aesthetic Model organizes this complexity into object representation, predictive processing, affective valuation, self-meaning integration, and social-institutional context. It calls for distributional outcomes, multilevel sampling, causal manipulation of context, time-resolved measurement, and explanation validated through intervention.

The most promising future is neither a return to unmeasurable subjectivism nor an expansion of score prediction alone. It is an evidence-integrated aesthetics in which psychological theory defines meaningful constructs, neuroscience constrains process, computation tests generalization, and humanistic inquiry identifies the social conditions of value. Such a framework can explain not only why an image is liked, but how aesthetic experience becomes meaningful, contested, and culturally consequential.

Funding

This research was supported by the Watershed Vernacular Built Environment Shanxi Provincial Key Laboratory Open Research Fund (WaBEL2024-05) and the 2025 Shanxi Province Art Science Planning Project (25G126).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

- [1] Smith, J. D. & Melara, R. J. (1990). Aesthetic preference and syntactic prototypicality in music: 'Tis the gift to be simple. *Cognition*, 34(3), 279-298. [https://doi.org/10.1016/0010-0277\(90\)90007-7](https://doi.org/10.1016/0010-0277(90)90007-7)
- [2] Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Review*, 8(4), 364-382. https://doi.org/10.1207/s15327957pspr0804_3
- [3] Leder, H., Belke, B., Oeberst, A., et al. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489-508. <https://doi.org/10.1348/0007126042369811>
- [4] Hekkert, P., Snelders, D., & Van Wieringen, P. C. W. (2003). 'Most advanced, yet acceptable': Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, 94(1), 111-124. <https://doi.org/10.1348/000712603762842147>
- [5] Van de Cruys, S. & Wagemans, J. (2011). Putting Reward in Art: A Tentative Prediction Error Account of Visual Art. *i-Perception*, 2(9), 1035-1062. <https://doi.org/10.1068/i0466aap>
- [6] Chatterjee, A. & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences*, 18(7), 370-375. <https://doi.org/10.1016/j.tics.2014.03.003>
- [7] Leder, H. & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments : The aesthetic episode – Developments and challenges in empirical aesthetics. *British Journal of Psychology*, 105(4), 443-464. <https://doi.org/10.1111/bjop.12084>
- [8] Pearce, M. T., Zaidel, D. W., Vartanian, O., et al. (2016). Neuroaesthetics. *Perspectives on Psychological Science*, 11(2), 265-279. <https://doi.org/10.1177/1745691615621274>
- [9] Pelowski, M., Markey, P. S., Forster, M., et al. (2017). Move me, astonish me... delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, 21, 80-125. <https://doi.org/10.1016/j.plrev.2017.02.003>

- [10] Menninghaus, W., Wagner, V., Hanich, J., et al. (2017). The Distancing-Embracing model of the enjoyment of negative emotions in art reception. *Behavioral and Brain Sciences*, 40, e347. <https://doi.org/10.1017/S0140525X17000309>
- [11] Skov, M. (2019). Aesthetic Appreciation: The View From Neuroimaging. *Empirical Studies of the Arts*, 37(2), 220-248. <https://doi.org/10.1177/0276237419839257>
- [12] Nadal, M. & Skov, M. (2024). The sensory valuation account of aesthetic experience. *Nature Reviews Psychology*, 4(1), 49-63. <https://doi.org/10.1038/s44159-024-00385-y>
- [13] Li, R. & Zhang, J. (2020). Review of computational neuroaesthetics: bridging the gap between neuroaesthetics and computer science. *Brain Informatics*, 7(1), 16. <https://doi.org/10.1186/s40708-020-00118-w>
- [14] Graf, L. K. M. & Landwehr, J. R. (2015). A Dual-Process Perspective on Fluency-Based Aesthetics. *Personality and Social Psychology Review*, 19(4), 395-410. <https://doi.org/10.1177/1088868315574978>
- [15] Silvia, P. J. (2005). What Is Interesting? Exploring the Appraisal Structure of Interest. *Emotion*, 5(1), 89-102. <https://doi.org/10.1037/1528-3542.5.1.89>
- [16] Frascaroli, J., Leder, H., Brattico, E., et al. (2024). Aesthetics and predictive processing: grounds and prospects of a fruitful encounter. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 379(1895), 20220410. <https://doi.org/10.1098/rstb.2022.0410>
- [17] Muth, C. & Carbon, C. C. (2013). The Aesthetic Aha: On the pleasure of having insights into Gestalt. *Acta Psychologica*, 144(1), 25-30. <https://doi.org/10.1016/j.actpsy.2013.05.001>
- [18] Carbon, C. C. (2011). Cognitive Mechanisms for Explaining Dynamics of Aesthetic Appreciation. *i-Perception*, 2(7), 708-719. <https://doi.org/10.1068/i0463aap>
- [19] Sarasso, P., Neppi-Modona, M., Sacco, K., et al. (2020). "Stopping for knowledge": The sense of beauty in the perception-action cycle. *Neuroscience & Biobehavioral Reviews*, 118, 723-738. <https://doi.org/10.1016/j.neubiorev.2020.09.004>
- [20] Wassiliwizky, E. & Menninghaus, W. (2021). Why and How Should Cognitive Science Care about Aesthetics? *Trends in Cognitive Sciences*, 25(6), 437-449. <https://doi.org/10.1016/j.tics.2021.03.008>
- [21] Pelowski, M. (2015). Tears and transformation: feeling like crying as an indicator of insightful or "aesthetic" experience with art. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01006>
- [22] Pizzolante, M., Pelowski, M., Demmer, T. R., et al. (2024). Aesthetic experiences and their transformative power: a systematic review. *Frontiers in Psychology*, 15, 1328449. <https://doi.org/10.3389/fpsyg.2024.1328449>
- [23] Christensen, A. P., Cardillo, E. R., & Chatterjee, A. (2025). Can art promote understanding? A review of the psychology and neuroscience of aesthetic cognitivism. *Psychology of Aesthetics, Creativity, and the Arts*, 19(1), 1-13. <https://doi.org/10.1037/aca0000541>
- [24] Marković, S. (2012). Components of Aesthetic Experience: Aesthetic Fascination, Aesthetic Appraisal, and Aesthetic Emotion. *i-Perception*, 3(1), 1-17. <https://doi.org/10.1068/i0450aap>
- [25] Menninghaus, W., Wagner, V., Wassiliwizky, E., et al. (2019). What are aesthetic emotions? *Psychological Review*, 126(2), 171-195. <https://doi.org/10.1037/rev0000135>
- [26] Schindler, I., Hosoya, G., Menninghaus, W., et al. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLOS ONE*, 12(6), e0178899. <https://doi.org/10.1371/journal.pone.0178899>
- [27] Kawabata, H. & Zeki, S. (2004). Neural Correlates of Beauty. *Journal of Neurophysiology*, 91(4), 1699-1705. <https://doi.org/10.1152/jn.00696.2003>
- [28] Cela-Conde, C. J., Marty, G., Maestú, F., et al. (2004). Activation of the prefrontal cortex in the human visual aesthetic perception. *Proceedings of the National Academy of Sciences*, 101(16), 6321-6325. <https://doi.org/10.1073/pnas.0401427101>
- [29] Jacobsen, T., Schubotz, R. I., Höfel, L., et al. (2006). Brain correlates of aesthetic judgment of beauty. *NeuroImage*, 29(1), 276-285. <https://doi.org/10.1016/j.neuroimage.2005.07.010>
- [30] Di Dio, C., Macaluso, E., & Rizzolatti, G. (2007). The Golden Beauty: Brain Response to Classical and Renaissance

- Sculptures. PLoS ONE, 2(11), e1201. <https://doi.org/10.1371/journal.pone.0001201>
- [31] Cupchik, G. C., Vartanian, O., Crawley, A., et al. (2009). Viewing artworks: Contributions of cognitive control and perceptual facilitation to aesthetic experience. *Brain and Cognition*, 70(1), 84-91. <https://doi.org/10.1016/j.bandc.2009.01.003>
- [32] Vartanian, O. & Skov, M. (2014). Neural correlates of viewing paintings: Evidence from a quantitative meta-analysis of functional magnetic resonance imaging data. *Brain and Cognition*, 87, 52-56. <https://doi.org/10.1016/j.bandc.2014.03.004>
- [33] Brown, S., Gao, X., Tisdelle, L., et al. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage*, 58(1), 250-258. <https://doi.org/10.1016/j.neuroimage.2011.06.012>
- [34] Vartanian, O., Navarrete, G., Chatterjee, A., et al. (2013). Impact of contour on aesthetic judgments and approach-avoidance decisions in architecture. *Proceedings of the National Academy of Sciences*, 110(supplement_2), 10446-10453. <https://doi.org/10.1073/pnas.1301227110>
- [35] Vessel, E. A., Starr, G. G., & Rubin, N. (2012). The brain on art: intense aesthetic experience activates the default mode network. *Frontiers in Human Neuroscience*, 6. <https://doi.org/10.3389/fnhum.2012.00066>
- [36] Vessel, E. A., Starr, G. G., & Rubin, N. (2013). Art reaches within: aesthetic experience, the self and the default mode network. *Frontiers in Neuroscience*, 7. <https://doi.org/10.3389/fnins.2013.00258>
- [37] Vessel, E. A., Isik, A. I., Belfi, A. M., et al. (2019). The default-mode network represents aesthetic appeal that generalizes across visual domains. *Proceedings of the National Academy of Sciences*, 116(38), 19155-19164. <https://doi.org/10.1073/pnas.1902650116>
- [38] Skov, M. & Nadal, M. (2020). A Farewell to Art: Aesthetics as a Topic in Psychology and Neuroscience. *Perspectives on Psychological Science*, 15(3), 630-642. <https://doi.org/10.1177/1745691619897963>
- [39] Mastandrea, S., Fagioli, S., & Biasi, V. (2019). Art and Psychological Well-Being: Linking the Brain to the Aesthetic Emotion. *Frontiers in Psychology*, 10, 739. <https://doi.org/10.3389/fpsyg.2019.00739>
- [40] Swami, V. (2013). Context matters: Investigating the impact of contextual information on aesthetic appreciation of paintings by Max Ernst and Pablo Picasso. *Psychology of Aesthetics, Creativity, and the Arts*, 7(3), 285-295. <https://doi.org/10.1037/a0030965>
- [41] Kirk, U. (2008). The Neural Basis of Object-Context Relationships on Aesthetic Judgment. *PLoS ONE*, 3(11), e3754. <https://doi.org/10.1371/journal.pone.0003754>
- [42] Li, Q. (2025). How Context and Painting Attributes Affect Aesthetic Judgment Across Expertise. *Empirical Studies of the Arts*, 43(1), 402-423. <https://doi.org/10.1177/02762374241262606>
- [43] Arai, S. & Kawabata, H. (2016). Appreciation Contexts Modulate Aesthetic Evaluation and Perceived Duration of Pictures. *Art and Perception*, 4(3), 225-239. <https://doi.org/10.1163/22134913-00002052>
- [44] Iosifyan, M. (2021). Theory of Mind Increases Aesthetic Appreciation in Visual Arts. *Art & Perception*, 9(2), 113-133. <https://doi.org/10.1163/22134913-bja10011>
- [45] Yeh, Y. C. & Peng, Y. Y. (2019). The Influences of Aesthetic Life Experience and Expertise on Aesthetic Judgement and Emotion in Mundane Arts. *International Journal of Art & Design Education*, 38(2), 492-507. <https://doi.org/10.1111/jade.12213>
- [46] Carbon, C. C. (2020). Ecological Art Experience: How We Can Gain Experimental Control While Preserving Ecologically Valid Settings and Contexts. *Frontiers in Psychology*, 11, 800. <https://doi.org/10.3389/fpsyg.2020.00800>
- [47] Specker, E., Tinio, P. P. L., & van Elk, M. (2017). Do you see what I see? An investigation of the aesthetic experience in the laboratory and museum. *Psychology of Aesthetics, Creativity, and the Arts*, 11(3), 265-275. <https://doi.org/10.1037/aca0000107>
- [48] Cox, R. F. A. & van Klaveren, L. M. (2024). The Embodied Experience of Abstract Art: An Exploratory Study. *Ecological Psychology*, 36(2), 111-122. <https://doi.org/10.1080/10407413.2024.2355901>
- [49] Świątek, A. H., Szcześniak, M. G., Stempień, M., et al. (2024). The mediating effect of the need for cognition between aesthetic experiences and aesthetic competence in art. *Scientific Reports*, 14(1), 3408. <https://doi.org/10.1038/s41598->

024-53957-6

- [50] Redies, C. (2015). Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. *Frontiers in Human Neuroscience*, 09. <https://doi.org/10.3389/fnhum.2015.00218>
- [51] Larrain, A. & Haye, A. (2019). Self as an Aesthetic Effect. *Frontiers in Psychology*, 10, 1433. <https://doi.org/10.3389/fpsyg.2019.01433>
- [52] Murray, N., Marchesotti, L., & Perronnin, F. (2012). AVA: A large-scale database for aesthetic visual analysis. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2408-2415). IEEE. <https://doi.org/10.1109/CVPR.2012.6247954>
- [53] Kong, S., Shen, X., Lin, Z., et al. (2016). Photo Aesthetics Ranking Network with Attributes and Content Adaptation. In *Computer Vision - ECCV 2016* (pp. 662-679). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_40
- [54] Talebi, H. & Milanfar, P. (2018). NIMA: Neural Image Assessment. *IEEE Transactions on Image Processing*, 27(8), 3998-4011. <https://doi.org/10.1109/TIP.2018.2831899>
- [55] Zeng, H., Cao, Z., Zhang, L., et al. (2020). A Unified Probabilistic Formulation of Image Aesthetic Assessment. *IEEE Transactions on Image Processing*, 29, 1548-1561. <https://doi.org/10.1109/TIP.2019.2941778>
- [56] He, S., Zhang, Y., Xie, R., et al. (2022). Rethinking Image Aesthetics Assessment: Models, Datasets and Benchmarks. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence* (pp. 942-948). International Joint Conferences on Artificial Intelligence Organization. <https://doi.org/10.24963/ijcai.2022/132>
- [57] Yang, Y., Xu, L., Li, L., et al. (2022). Personalized Image Aesthetics Assessment with Rich Attributes. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 19829-19837). IEEE. <https://doi.org/10.1109/CVPR52688.2022.01924>
- [58] Briellmann, A. A. & Pelli, D. G. (2018). Aesthetics. *Current Biology*, 28(16), R859-R863. <https://doi.org/10.1016/j.cub.2018.06.004>
- [59] Li, Z., Yan, X., Wei, X., et al. (2025). IAACLIP: Image Aesthetics Assessment via CLIP. *Electronics*, 14(7), 1425. <https://doi.org/10.3390/electronics14071425>