How Non-Living Intelligence Brings Life to Music: A Philosophical, Cognitive, and Structural Analysis of AI-Created Sound

Ziad Salah
Independent Researcher
GitHub: github.com/Zierax
X: @Zierax_X

February 2026

Abstract

The question of whether artificially generated music can be considered "alive" in any meaningful aesthetic sense has been framed, almost universally, as a question about the machine. This paper argues that this framing is the problem. When examined through the intersecting lenses of perceptual cognition, structural aesthetics, and systems theory, what becomes clear is that musical vitality has never resided in the generative source , it has always resided in the interpretive encounter between structured sound and a perceiving mind. Artificial intelligence, understood not as a creative agent but as a high-dimensional structural amplifier, does not challenge this relationship; it clarifies it. This paper introduces the concept of *Delegated Vitality* , a framework describing the conditions under which perceived aliveness in sound emerges independently of the originating system's phenomenological status. Across six conceptual axes , music as perceptual construct, non-living creativity, the illusion of liveness, intentionality transfer, the collapse of process-based legitimacy, and AI as instrument , this paper builds an argument for repositioning the philosophical debate around AI music away from questions of authorship and consciousness, and toward questions of structure, constraint design, and interpretive activation. The contribution is not a defence of AI art, nor a critique of it. It is a more precise description of what music actually is, and why that description has always accommodated the non-living.

Keywords: generative music, delegated vitality, structural creativity, intentionality transfer, perceptual aesthetics, cognitive musicology, philosophy of AI

1. Introduction

There is a version of this debate that is not worth having. Whether artificial intelligence "truly" creates, whether its outputs deserve the label of art, whether something without experience can produce something meaningful , these questions, taken at face value, assume answers to much harder problems that philosophy of mind has not resolved. They assume we know what creativity is. They assume we know what experience contributes to meaning. They assume the value of art is traceable to its source rather than its reception.

The more productive framing begins elsewhere: not with what the machine is, but with what music does. If music is examined not as an object produced but as a cognitive event triggered, the question of whether the originating system possesses consciousness becomes, if not entirely irrelevant, then at least secondary to the question of whether the resulting structure is capable of activating the perceptual and interpretive mechanisms through which musical meaning is constructed.

This is not a minor reframing. It displaces the entire ground on which most resistance to AI music stands. Most objections to AI-generated music rest on what this paper will call the source-dependency thesis: the claim that artistic value inheres in or flows from the intentions, experiences, or consciousness of the producing agent. This thesis, examined carefully, does not survive contact with the history of music theory, cognitive science, or aesthetic philosophy. Music has always operated at a level of structural organization that outruns the intentions of any individual producer.

The paper proceeds as follows. Section 2 establishes the conceptual vocabulary and positions the paper relative to existing literature. Section 3 examines music as a perceptual construct, arguing that musical meaning is an emergent property of listener cognition rather than an intrinsic property of sonic objects , and presents a preliminary observational case from compositional practice that bears directly on this claim. Section 4 introduces a tripartite model of creativity that decouples structural creativity from consciousness. Section 5 addresses the illusion of liveness and introduces the *Delegated Vitality* framework. Sections 6 and 7 examine intentionality transfer and the historical pattern by which process-based artistic legitimacy gives way to outcome-based legitimacy. Section 8 formalizes the AI-as-instrument thesis. Section 9 offers the paper's original theoretical contribution as a synthesizing model. Section 10 addresses implications, and Section 11 concludes.

2. Conceptual Foundations

2.1 The Source-Dependency Thesis and Its Vulnerabilities

The dominant philosophical tradition in Western aesthetics has, at least since Romanticism, located artistic value in the expressive act of an individual consciousness. Expression theories of art , from Croce's intuition-expression monism to Collingwood's account of artistic clarification , ground the value of the artwork in the interior activity of the artist (Collingwood, 1938; Croce, 1902/1909). On this view, a work without an experiencing subject behind it is not merely incomplete; it is categorically disqualified.

This tradition is not indefensible, but it carries a significant burden when extended to music specifically. Music, unlike representational painting or narrative literature, does not obviously point to a referent. Its semantic content , to the extent it has any , is contested at every level. The formalist tradition, most rigorously developed by Hanslick, argues that music's content is precisely its form: the dynamic movement of tonal structures in time (Hanslick, 1854/1986). What music expresses, on this account, is not a state of mind but a particular configuration of relationships , tension and resolution, expectation and disruption, pattern and deviation.

If Hanslick is even partially correct, then the question of whether an originating consciousness intended to express a particular inner state becomes far less important than whether the resulting structure embodies the relational properties that activate musical perception. The implications for AI music follow immediately: what matters is whether the generated structure instantiates those relational properties. The consciousness of the generator is, on this account, orthogonal to the question.

2.2 What "Life" in Music Has Ever Meant

Before this paper can argue that non-living intelligence brings life to music, it needs to be precise about what "life" in music actually denotes. In ordinary usage, musicians and listeners describe music as alive when it exhibits properties that correlate with biological vitality: responsiveness, unpredictability within coherence, energy that feels motivated rather than mechanical, and a quality sometimes called *presence*.

These descriptions, examined structurally, are not mystical. They refer to specific perceptual experiences: the sense that a musical phrase is going somewhere, that it responds to what preceded it, that its next move is constrained but not yet determined. These are properties of temporal structure , not of the psychological state of the composer at the moment of creation. A Mozart symphony performed by a computer-controlled orchestra with no human performers exhibits these properties as fully as the same symphony performed by living musicians. The "life" perceived is in the structure, activated by the perceiver.

This observation is foundational to everything that follows.

3. Cognitive and Perceptual Analysis: Music as Constructed Event

3.1 The Perceiver as Co-Author

Musical meaning is not transmitted from composer to listener. It is constructed by the listener using the musical structure as a scaffold. This is not a controversial claim in cognitive musicology; it is the field's central finding. Meyer's foundational work on emotion and meaning in music (1956) established that musical affect is generated by the dynamic between expectation and its fulfillment or violation. The listener brings to any musical encounter a dense network of learned patterns , stylistic conventions, harmonic schemas, rhythmic expectations , and the music activates, confirms, delays, or defeats these patterns in ways that produce structured emotional responses.

At the cognitive level, this means musical meaning is always a joint production. The structure provides the constraint; the perceiver provides the interpretive framework; the meaning emerges from their interaction. Crucially, it follows that the composer's original intention is only one input into this process , and not necessarily the most significant one.

This point is not merely theoretical. It was demonstrated with some specificity in an observational case arising from the author's concurrent practice in philosophical fiction and generative music composition.

Observational Case: Structural Activation Without Narrative Context

A long-form philosophical narrative , structured around the cosmological rebellion of a sovereign demonic figure , required an original score to accompany specific dramatic junctures. The compositional demands of the narrative's climactic sequence, designated internally as *The Last Piece*, presented a methodologically useful opportunity: this segment depicts the final phase of the central conflict, specifically a pursuit sequence that closes the war arc of the narrative.

Two compositional approaches were employed in parallel for this sequence: a manually composed version, produced through conventional compositional judgment, and several AI-generated versions, produced through iterative constraint specification within a generative system. The versions were developed independently, with each output evaluated against the structural and dramatic requirements of the sequence as the author understood them.

Following the completion of both versions, an informal perceptual study was conducted among ten individuals drawn from the author's immediate social network. Critically, none of the participants possessed any prior knowledge of the narrative, its thematic content, its genre, or the identity of the scene the music was intended to accompany. Participants were also non-specialist listeners , individuals with regular music consumption habits but without formal training in composition, music theory, or audio production. The absence of industry expertise was deliberate:

the aim was to assess naive perceptual response to structural properties, unconditioned by technical frameworks or genre familiarity.

Participants were presented with the music and asked a single open-ended question: *What is the first image or situation that comes to mind when you listen to this?*

The results, while drawn from too small a sample to carry statistical weight, are structurally suggestive in a way that is directly relevant to the paper's argument. Seven of the ten participants independently identified a pursuit or chase scenario. Two described a scene of warfare or large-scale conflict. The remaining participant described close-range combat. No participant was directed toward any of these responses; the question was left open to any associative domain whatsoever.

What makes this observation analytically significant is not the majority response in isolation, but the distribution of all three responses considered together. Pursuit, warfare, and combat are not identical concepts , they differ in scale, in spatial organization, in the number of agents implied, and in their temporal structure. But they share a formal signature: escalating urgency, asymmetric agency between parties, forward momentum under threat, and temporal compression toward a point of resolution. The participants did not converge on a single image; they converged on a *family* of structural experiences, all of which map precisely onto the dramatic logic of the sequence the music was designed to accompany.

This suggests that the AI-generated compositional output embedded structural properties , tempo trajectory, rhythmic density, harmonic tension patterns, dynamic contour , that were sufficient to activate congruent interpretive frameworks across listeners who had no access to any narrative context. The meaning was not transmitted through the story. It was constructed by each listener independently, using the musical structure as the sole available scaffold.

The finding does not resolve the central argument of this paper; it would be methodologically irresponsible to claim otherwise from ten responses. What it does is sharpen the question considerably. If a generative system, operating under constraints specified by a human orchestrator holding a particular dramatic intention, produces a structure that reliably activates semantically congruent perceptual responses in naive listeners , responses that align with the narrative logic the composer was working from , then the relevant question is no longer simply whether AI can produce aesthetically valid music. The more productive question is whether AI can function as a medium for the transmission of *structural intentionality*: whether the constraint space, designed by the human operator, encodes sufficient aesthetic intelligence to participate in genuine communicative acts between structure and perceiver, without requiring the listener to know anything about the source.

This is the question the *Delegated Vitality Framework*, introduced in Section 5, is designed to address.

Huron's elaboration of Meyer's framework into the ITPRA model (Imagination, Tension, Prediction, Reaction, Appraisal) sharpens the theoretical picture considerably (Huron, 2006). Musical

experience is structured by predictive processing: the brain continuously generates probabilistic models of what will happen next, and the quality of musical experience is largely a function of how those predictions are managed. Clark's work on predictive processing as a general feature of cognition provides the broader theoretical backdrop here , the brain, on this account, is fundamentally a prediction machine, and musical experience is one of the most structurally dense exercises of that capacity (Clark, 2013; 2016). Surprise, satisfaction, and tension are not properties of the music; they are properties of the relationship between the music's structure and the listener's predictive apparatus.

The implication for AI-generated music is not trivial. If musical meaning is a function of structural properties that interact with listener cognition, and if AI systems can produce structures that instantiate those properties, then the phenomenological status of the AI is not a relevant variable in the musical equation. The question is whether the structure works , not whether the system that generated it is alive.

3.2 Memory, Pattern Completion, and the Construction of Continuity

Musical experience is inseparable from memory. A melody is not experienced as a series of isolated pitches; it is experienced as a trajectory, each note acquiring meaning from what preceded it while anticipating what follows. This temporal integration is not a peripheral feature of musical perception; it is constitutive of what music is. Without the memory structures that allow a listener to hold a phrase in active perception while orienting toward its continuation, there would be no music , only a sequence of sounds.

This has a consequence that bears directly on the AI question: the sense that a piece of music is unified, directed, and coherent , that it feels like a single developing thing rather than a series of disconnected events , is an achievement of the listener's cognitive apparatus, not a property straightforwardly resident in the acoustic signal. A listener unfamiliar with tonal harmony will not experience a dominant seventh chord as inherently tension-generating; the tension is produced by the interaction of the chord with a schema the listener has internalized (Sloboda, 1985; Zbikowski, 2002).

AI systems trained on large corpora of music have, in a meaningful functional sense, internalized those same schemas. Not as conscious knowledge, but as statistical regularities encoded in model parameters. When such a system generates a phrase that conforms to and then departs from a particular harmonic schema, it produces exactly the kind of structural event that triggers the listener's expectation-management systems. Schaefer, Overy, and Nelson (2013) demonstrate that these predictive mechanisms in musical behaviour are not uniform , they operate differentially across structural dimensions, which has direct implications for how AI-generated structures might be tuned to exploit specific perceptual channels more precisely than conventional compositional intuition would suggest.

The perceiver does not ask, and cannot determine from the structure alone, whether the schema was deployed with conscious intent.

4. Structural Creativity in Non-Living Systems

4.1 Decoupling Creativity from Consciousness: A Tripartite Model

The conflation of creativity with consciousness is a philosophical assumption that has been insufficiently interrogated, largely because, prior to AI, all known instances of creativity were produced by conscious beings. The availability of AI as a case study makes it possible , indeed, necessary , to examine whether creativity requires consciousness, or whether consciousness has simply been the contingent substrate in which creativity has historically occurred.

This paper proposes a formal tripartite distinction:

Expressive creativity refers to the capacity to externalize an inner state, transforming phenomenological experience into a communicable form. This is the domain in which consciousness is genuinely necessary: there must be an inner state to externalize. Autobiographical songwriting, confessional poetry, and representational painting in the Romantic tradition are paradigm cases.

Structural creativity refers to the capacity to generate novel configurations of elements within a formal system that exhibit internal coherence and measurable departure from prior patterns. Structural creativity is defined entirely at the level of relationships between elements; it requires no inner state, only a generative process capable of producing configurations not reducible to simple recombination of existing templates. Mathematical proofs, chess combinations, and , critically , most of what we call "musical composition" at the technical level fall within this category (Boden, 2004).

Interpretive creativity refers to the capacity to produce readings of existing structures that disclose possibilities not previously recognized. This is the creativity of the analyst, the performer, and the listener who hears something in a piece that its composer did not consciously place there.

The claim of this paper is that AI operates primarily, and with increasing sophistication, in the structural domain. This is neither a limitation nor an achievement to be celebrated; it is a precise description. AI systems do not externalize inner states (expressive creativity is unavailable to them) and they do not engage in genuine interpretation in the phenomenological sense. But structural creativity , the generation of novel, coherent, constraint-satisfying configurations , does not require consciousness, and AI systems have by now demonstrated this conclusively.

The significance of this distinction is that structural creativity is not a lesser form of creativity. The majority of what music theorists analyze when they examine a Bach fugue or a Beethoven sonata is

structural: the relationships between voices, the logic of the development section, the coherence of the tonal plan. These are not properties that depend on Bach's or Beethoven's consciousness for their existence or their value.

4.2 The Generative Mechanism Is Not the Measure

A persistent objection to AI creativity holds that because AI systems operate by statistical pattern matching , predicting what comes next based on what they have seen , they are not genuinely creative but merely recombinatory. This objection, examined structurally, proves either too much or too little.

If recombination disqualifies an output from creative status, then the vast majority of human musical composition is also disqualified. Music theory is, in substantial part, the study of rules governing which combinations are permissible, conventional, or surprising within a given stylistic tradition. A composer working within the Viennese Classical style is not freely inventing from nothing; they are operating within an extraordinarily dense constraint system, and their creativity consists in what they do within that system, not in the system itself.

The difference between human and AI composition is not that one recombines and the other does not. The difference lies in how constraint is applied and what kinds of structural novelty are achievable. A human composer can deliberately violate a convention with full metacognitive awareness of its historical significance. An AI system generates violations statistically, without that reflective orientation. Whether this difference matters aesthetically , whether it actually affects the structural properties of the output in ways that a listener can detect , is an empirical question, not a philosophical axiom. Dahlstedt (2021) frames this directly as a question about the relationship between knowledge-driven and material-driven creativity, arguing that computational systems occupy a distinct but legitimate position in this space rather than a deficient one.

5. The Illusion of Liveness: Delegated Vitality

5.1 Why Humans Attribute Life to Structure

The attribution of vitality to non-living systems is not a cognitive error; it is a feature of the perceptual systems through which humans navigate a world full of agents. Animacy perception , the tendency to attribute goal-directedness and agency to moving objects under certain conditions , is among the most primitive and robust features of human visual cognition (Heider & Simmel, 1944). Its auditory analogue is equally pervasive. Listeners attribute intention, emotion, and directedness to musical phrases with no more evidence than their structural properties.

This is not naïveté. It is the application of the most powerful interpretive strategy available to a social organism: agent modeling. The brain, confronted with a structured temporal sequence that exhibits the statistical signatures of goal-directed behavior , that moves toward something, that responds to preceding events, that shows coherent variation , activates the cognitive systems used for modeling intentional agents. The result is the experience of "life" in the music.

This attribution is triggered by structure, not by the actual presence of an agent. Natural phenomena , the sound of wind through trees, the rhythm of ocean waves, the crackling of fire , contain no intentional agents, yet can produce, under certain listening conditions, the same quality of perceived vitality that music produces. This is not because listeners confuse wind for a composer; it is because the structural properties of these phenomena share features with goal-directed sequences at the level of perceptual processing.

5.2 Delegated Vitality: A Framework

This paper introduces the concept of *Delegated Vitality* to describe the process by which perceived aliveness in a musical work is generated through the interaction of three factors: (1) the structural properties of the sound, (2) the perceptual and cognitive systems of the listener, and (3) the intentional framework within which the listener positions the work.

The term "delegated" is chosen carefully. Vitality in music is not intrinsic to the acoustic signal; it is delegated to the signal by the listener's perceptual systems, which are themselves shaped by the listener's prior experience of how intentional agents produce structured sound. When a listener hears an AI-generated composition, the vitality they perceive is real , it is a genuine feature of their experience , but it is produced by the activation of cognitive systems that do not require the actual presence of a living agent. The observational case presented in Section 3.1 is directly relevant here: the ten listeners who assigned kinetic, conflict-based meanings to the AI-generated sequence were not making errors of attribution. They were doing precisely what Heider and Simmel's subjects did in 1944 , reading structural properties as signatures of agency and directed movement, because that is what human perceptual systems are built to do.

What distinguishes the Delegated Vitality account from previous accounts of musical perception is the explicit introduction of the third factor: the intentional framework. The same piece of music heard under different framings , as a composition by a human, as a piece generated by an AI, as a natural sonic phenomenon , does not necessarily produce the same perceptual experience, not because the structure changes, but because the listener's interpretive posture changes. The intentional framework shapes which features of the structure are attended to and how heavily they are weighted.

This has a counterintuitive implication: knowing that a piece was generated by AI does not necessarily diminish its perceived vitality, and in some cases may enhance it. When a listener frames an AI composition as a demonstration of structural capacity rather than expressive communication, different perceptual strategies come into play , strategies more closely aligned with how a musician listens to a technical performance or how a theorist approaches a contrapuntal exercise. The structure is apprehended in a mode that is not primarily empathic but formally analytic, and the vitality perceived is of a different but not lesser kind.

5.3 Natural Analogue and Structural Sufficiency

The comparison to natural phenomena is not merely illustrative; it is structurally significant. The Japanese aesthetic concept of *ma* , the meaningful interval or pause in sound , operates identically whether instantiated by a human performer, a natural soundscape, or a generative system. The structural property (the interval, the silence that creates forward tension, the breath between

phrases) produces its perceptual effect regardless of its source. The structural property is the thing; the source is not.

This suggests that structural sufficiency , the property of a sound sequence being adequately structured to activate musical perception , is achievable by non-living systems, and has always been achievable by them. What has changed with AI is not the principle but the degree of structural sophistication achievable, and the explicitness with which the human role is one of design rather than direct production.

6. Intentionality and Interpretation

6.1 Intentionality Transfer: Encoding Human Aesthetic Intent in Generative Systems

The claim that AI music lacks human intention is, at best, a claim about proximate causation. At the level of distal causation , the chain of decisions that produced the system, selected the training data, designed the loss functions, specified the constraints, and framed the generative task , human intentionality is pervasive throughout. The question is not whether human intention is present in AI music, but where it is located in the causal chain.

This paper proposes the concept of *intentionality transfer* to describe the process by which human aesthetic intent is encoded into a generative system and subsequently activated in its outputs. Intentionality transfer operates at multiple levels: at the level of corpus selection (which examples of human musical production are used to shape the system's statistical dispositions), at the level of architectural design (which structural regularities the system is capable of capturing and reproducing), at the level of constraint specification (which parameters are set by the human operator for a given generative task), and at the level of curation (which outputs are selected for presentation and which are discarded).

At each of these levels, human aesthetic judgment is operative. The AI system does not generate in a vacuum; it generates within a space defined, shaped, and delimited by human choices. The resulting output is not the product of an autonomous creative agent; it is the product of a system whose generative space has been carefully engineered by human agents with specific aesthetic ends. This is precisely what was operative in the compositional case described in Section 3.1: the human operator's knowledge of the narrative sequence , its dramatic logic, its kinetic character, its emotional register , shaped the constraint specification that produced the AI outputs. The listeners' responses suggest that this intent was not merely encoded but transmitted, structurally, through the generated material to naive perceivers.

6.2 Authorship vs. Orchestration

The concept of authorship , the attribution of a work to a single originating consciousness whose intentions the work expresses , is a historically contingent construct. Its dominance in Western art reflects specific economic and legal structures (copyright as the legal expression of authorial

intent) and specific ideological commitments (Romantic individualism, the genius myth) rather than any deep truth about the nature of creative production (Moruzzi, 2018).

In practice, musical production has always been more distributed than the authorship model acknowledges. A composer writes a score; a publisher edits it; a conductor interprets it; an orchestra realizes it; a recording engineer shapes its sonic presentation; a streaming algorithm determines the context in which it is heard. The listener receives the product of this entire chain, and attributes it to "the composer" in a move that is interpretively convenient but analytically inadequate.

AI music does not create this problem; it makes it visible. When the generative system is foregrounded, the distributed nature of musical authorship can no longer be elided. What replaces the author, on this account, is the *orchestrator*: the agent who designs the generative conditions, shapes the constraint space, curates the outputs, and positions the work within an interpretive context. Orchestration in this sense is not a lesser form of authorship; it is a more accurate description of what sophisticated musical production has always been.

The distinction between authorship and orchestration is not merely conceptual; it has consequences for how we evaluate AI-assisted music. The relevant question is not whether the composer touched every note, but whether the constraint space was designed with sufficient aesthetic intelligence to produce outputs worth attending to. This is a question about the quality of the orchestration, and it is answerable by the same evaluative criteria we apply to other forms of musical judgment.

6.3 The Receiver's Role as a Third Axis of Intention

Gadamer's hermeneutic circle, applied to music, suggests that the meaning of a work is not fixed at its production but continually constituted in the act of reception (Gadamer, 1960/1989). The listener is not a passive decoder of the composer's message; they are an active participant in the production of meaning, and their particular position , historical, cultural, experiential , shapes what the work means in ways that no composer can fully anticipate or control.

This has a specific implication for AI music: the receiver's intentional framework contributes to the meaning of the work as much as the generative system's structural outputs. A musicologist listening to an AI-generated fugue brings to it the entire apparatus of their analytical training; they hear voice-leading, motivic development, structural proportion. A casual listener brings a different set of schemas. Neither reading is more authentic than the other, because there is no originating consciousness whose intent could serve as the authoritative arbiter of correctness.

The observational case in Section 3.1 demonstrates this from the other direction: even without any analytical framework, naive listeners constructed coherent interpretive responses that were structurally convergent with the compositional intention. This convergence did not require them to know anything about the source. It required only that the structure be sufficient to activate their existing perceptual apparatus , which it was.

7. The Collapse of Process Importance

7.1 The Historical Pattern of Legitimacy Shift

Every major technological intervention in music production has followed the same pattern: initial resistance based on process-purity arguments, followed by gradual acceptance as the outputs demonstrate sufficient structural quality, followed by retrospective reframing in which the new technology is absorbed into the tradition it was claimed to threaten.

Electronic music provides the clearest case. The objections raised against early electronic music , Stockhausen's tape compositions, musique concrète , were precisely the objections now raised against AI music: that the process lacked the authenticity of live performance, that the machine could not express genuine human experience, that the resulting sounds were cold and mechanical. These objections did not survive contact with the actual outputs, which, when examined structurally, demonstrated the same properties of coherence, tension, and resolution that defined the tradition they were allegedly violating.

Sampling culture provides a second case. The argument that sampling is not real composition , that assembling fragments of existing recordings is categorically different from creating new ones , rested on process-purity intuitions that the subsequent history of recorded music has comprehensively undermined. The question the market and the canon eventually settled on was not "did you produce every sound?" but "does the result have structural integrity and aesthetic impact?"

Auto-tune provides a third case, distinguished from the others by the explicitness of the debate around process purity. The objection to auto-tune was specifically that it corrected a feature , pitch accuracy , previously understood as a marker of authentic performance. What was "authentic" imperfection became, in this framing, a value in itself; its removal was read as inauthenticity. This argument failed, not because pitch accuracy stopped mattering, but because listeners and musicians discovered that constrained imperfection and processed precision could both serve legitimate structural musical purposes. The process was not the point; the structural result was the point.

7.2 Why the AI Objection Follows the Same Pattern

The objection to AI music is structurally identical to these prior cases: it is a process-purity argument dressed in the language of consciousness and intent. The underlying move is the same , the claim that a particular feature of the generative process (human experience, live performance, manual composition) is necessary for musical legitimacy , and the historical pattern suggests it will meet the same fate.

This is not to say the objections are without philosophical interest; the question of what AI music is and what it means is genuinely complex. It is to say that the intuition driving the objections , that something is missing in AI-generated music that is present in human-composed music , will not,

by itself, settle the aesthetic question. The structural properties of the outputs are more relevant than the phenomenological status of the process by which they were generated.

The relevant historical variable is not whether the objections are philosophically serious, but whether the outputs achieve structural sufficiency. In the cases of electronic music and sampling, they did, and the objections receded accordingly. The evidence is accumulating that AI-generated music is achieving structural sufficiency across an increasing range of styles and forms , and the observational case discussed in this paper, modest as it is, contributes a data point to that accumulation.

8. AI as a Creative Instrument, Not an Agent

8.1 Constraint Design as the New Virtuosity

The framing of AI as a creative agent , as a system that composes autonomously , is both philosophically imprecise and practically misleading. More accurate, and more illuminating, is the framing of AI as a high-dimensional instrument: a system whose generative behavior is determined by the parameters set by its operator, in the same way that the sonic behavior of a violin is determined by the bowings, fingerings, and posture of its player.

The analogy is not perfect, but its imperfections are informative. A violin's range of possible outputs is constrained by its physical construction; an AI system's range of possible outputs is constrained by its training data, architecture, and operating parameters. A violinist achieves expressive results through fine-grained physical control; an AI operator achieves expressive results through constraint design , the specification of parameters that shape the generative space. In both cases, the human is not merely activating a system but shaping its outputs through a form of applied expertise. Pressing (1988), in his analysis of improvisation, describes skilled musical agency as fundamentally a process of navigating within constraint structures , a formulation that maps with some precision onto what the AI operator does, albeit at a different level of abstraction.

What changes with AI is the nature of the expertise required. Violin virtuosity is a matter of embodied skill, developed through years of physical training. AI virtuosity , if we can call it that , is a matter of *constraint intelligence*: the capacity to specify the generative conditions that will produce outputs with desired structural properties. This requires a different kind of musical knowledge: not performance technique, but an understanding of how generative parameters correspond to structural outcomes, and how structural outcomes correspond to perceptual effects.

8.2 From Notation to Constraint: The Continuity of Mediation

The history of Western music is, in substantial part, a history of mediating technologies , systems that stand between the composer's intention and the sonic realization of that intention. Musical

notation is the paradigm case: a system of symbolic encoding that allows compositional intentions to be transmitted across time and space, realized by performers who were not present at the composition. The introduction of notation did not make music less musical; it expanded the scope of what musical composition could achieve (Roads, 1996).

Digital audio workstations represent a second mediating layer: systems that allow compositional decisions to be encoded in formats that produce sonic outputs without the intervention of live performers. The objections raised against early DAW-based music , that it was cold, mechanical, lacking the expressive variation of live performance , are structurally identical to the objections raised against AI music today.

AI systems represent a third layer of mediation: systems in which the compositional decisions are encoded at the level of generative constraints rather than specific note choices. The human composer specifies not "this note, then this note" but "a structure with these properties, operating in this space, under these constraints." The step back from direct specification is real, but it is continuous with the step back represented by notation and DAW composition. In each case, the human role shifts from direct sonic production to the design of systems and instructions that produce sonic results. The distance from the sound increases; the sophistication of the design task does not decrease.

8.3 The Constraint Space as Aesthetic Object

There is a further implication of the instrument framing that deserves attention. When AI is understood as an instrument whose expressive range is defined by its constraint space, the constraint space itself becomes an object of aesthetic consideration. A poorly designed constraint space produces outputs that are technically competent but aesthetically undifferentiated , the equivalent of a skilled performance of a trivial exercise. A richly designed constraint space produces outputs that are structurally interesting across multiple listenings and responsive to the specific dramatic or expressive demands of the compositional context.

This means the evaluation of AI music cannot be fully separated from the evaluation of the generative conditions that produced it. A piece of AI music that is the product of a thoughtfully designed constraint space carries a different kind of aesthetic significance than one produced by default parameters. The structural intelligence of the operator is, in some sense, present in the output , not as expressed inner state, but as the shape of the generative possibility space. The compositional case described in Section 3.1 illustrates this: the convergence of listener responses around a shared structural family was not an accident. It was a consequence of a constraint space designed to produce a specific kind of kinetic, tension-escalating musical logic. The design worked. That is an evaluable fact about the quality of the orchestration.

9. Theoretical Model: The Delegated Vitality Framework

9.1 Statement of the Model

The foregoing analysis converges on a theoretical model that this paper proposes as an original contribution: the *Delegated Vitality Framework* (DVF). The DVF provides a systematic account of how musical vitality , the perceived quality of aliveness, presence, and directed energy in sound , is generated in AI-produced music, and why it varies across listeners, listening contexts, and compositional approaches.

The DVF operates on three analytical axes:

Axis 1: Structural Activation Potential (SAP) , the degree to which a sound structure contains the formal properties (tension-resolution dynamics, expectation management, motivic coherence, temporal directionality) that activate the listener's musical perception systems. SAP is a property of the structure, not of the generating system, and is assessable independently of the system's phenomenological status. In the observational case presented in Section 3.1, the convergent listener responses suggest a high SAP: the structure contained sufficient formal properties to activate congruent interpretive frameworks across ten independently responding perceivers.

Axis 2: Intentional Framework Contribution (IFC) , the degree to which the listener's positioning of the work within an interpretive framework shapes what is perceived and how it is evaluated. IFC captures the hermeneutic dimension: the same structure heard as "a composition" versus "a demonstration" versus "a natural phenomenon" activates meaningfully different perceptual strategies. AI music does not necessarily reduce IFC; under appropriate framing, it can redirect it toward structurally productive modes of attention that a traditional authorship framing would not afford.

Axis 3: Design Intelligence Index (DII) , the degree to which the generative conditions of the work reflect sophisticated aesthetic judgment: the quality of constraint design, the selection and curation of outputs, the positioning of the work within a listening context. DII is the primary axis along which the human contribution to AI music is assessable, and it corresponds to what this paper has called orchestration. It is the axis that distinguishes AI music that results from genuine compositional intelligence from AI music that results from default parameters and undirected generation.

Musical vitality in AI-generated work is, on this model, a function of all three axes jointly. A work with high SAP but low IFC , a structurally competent piece framed in ways that obstruct genuine musical engagement , will be experienced as technically impressive but aesthetically inert. A work with high SAP and high IFC but low DII , a structurally rich piece well framed but produced by unsophisticated generative conditions , will be experienced as accidental: a structural windfall rather than an aesthetic achievement. The full vitality experience requires all three axes to be functioning, which is precisely why the same AI-generated output can produce radically different responses in different listeners or different listening contexts.

9.2 What the DVF Explains That Existing Accounts Do Not

Existing philosophical accounts of AI music divide, roughly, into two camps: those that locate the problem in the absence of consciousness or intentionality in the generating system, and those that argue for the irrelevance of generative source to aesthetic value. Both camps have difficulty explaining the phenomenon of *variable* aesthetic response to AI music , why the same structurally

similar AI-generated pieces can produce radically different responses in different listeners or different listening contexts.

The DVF explains this by making the listener's intentional framework an explicit variable in the analysis (the IFC axis), while simultaneously providing a basis for non-trivially evaluating the quality of the human contribution to AI music through the DII axis. It avoids the source-dependency error of the first camp without making the overcorrection of the second (the claim that source is entirely irrelevant). The human contribution is real and evaluable; it is simply located differently in the generative chain than traditional authorship models suggest.

The DVF also provides a basis for distinguishing between AI music that is aesthetically significant and AI music that is not , a distinction that both dismissive and celebratory accounts of AI art have difficulty drawing rigorously. The relevant criteria are structural (SAP), hermeneutic (IFC), and orchestrative (DII). Each is assessable through established methodologies within musicology, cognitive science, and compositional theory respectively. And the observational case this paper presents , modest, preliminary, and offered as illustration rather than proof , demonstrates that the framework maps onto actual compositional and perceptual experience in a way that is at least consistent and potentially productive for further investigation.

9.3 Boundaries and Limitations of the Model

The DVF does not resolve all questions in the philosophy of AI music. It does not address whether structural creativity and expressive creativity differ in value , it brackets that question as dependent on contested assumptions about the relationship between inner states and aesthetic significance, assumptions that are not settled in the existing literature. It does not address the legal questions around AI music, which are determined by considerations largely orthogonal to the aesthetic ones this paper examines. And it does not predict, with specificity, which AI systems will produce the most aesthetically significant outputs , that is partly an empirical question and partly a function of the constraint intelligence of individual operators.

The DVF is proposed as a framework for analysis, not a theory of value. It provides the conceptual vocabulary for a more precise discussion of what is happening when AI generates music that listeners experience as vital, present, and aesthetically compelling. Whether such music is "good" in some deeper evaluative sense remains a question for the encounter between specific works and specific listeners operating within specific interpretive traditions. The framework clarifies the conditions of that encounter; it does not predetermine its outcome.

10. Implications for Music Theory and Philosophy of Art

10.1 For Music Theory

If the arguments of this paper are accepted, music theory faces a specific challenge: the dominant analytical vocabularies were developed in the context of human composition, and they carry implicit assumptions about intentionality and meaning that may require revision when applied to

AI-generated work. Schenkerian analysis, for example, operates on the premise that a tonal composition has a structural deep structure that represents the realization of a compositional intention; the concept of structural unity, central to tonal analysis, has historically been understood as the coherence of an intention realized through compositional craft.

Applied to AI music, these concepts can still do analytical work , but their theoretical grounding needs to be revised. Structural unity in AI music is not the unity of an intention; it is the unity produced by the interaction of generative constraints operating on trained statistical regularities. The analytical results may look similar on the page, but the theoretical interpretation of what those results signify requires regrounding in the framework this paper has outlined.

More broadly, music theory stands to benefit from developing analytical tools specifically designed for constraint-designed music , tools that evaluate not just the structural properties of outputs but the relationship between those properties and the generative conditions that produced them. This would be a form of analysis that operates simultaneously at the level of the sound structure and the level of the constraint space that generated it. Lerdahl and Jackendoff's generative theory of tonal music (1983) already works in this double register, though it was not designed with AI generation in mind; its underlying architecture is more compatible with the DVF than might initially appear.

10.2 For Philosophy of Art

The philosophy of art has been dominated, for reasons that are partly sociological and partly philosophical, by questions about the nature and value of human creative expression. AI art , and AI music specifically , provides a test case that forces those questions to become more precise. The intuition that human creative expression has special value is not obviously wrong; but it needs to be defended at a level of philosophical precision that vague appeals to consciousness, intentionality, and experience do not provide.

The DVF, together with the conceptual distinctions developed throughout this paper, provides resources for that project. By separating structural creativity from expressive creativity, and by providing an account of how perceived vitality in music is generated that does not depend on the consciousness of the generator, this paper makes room for both the genuine value of human expressive creativity and the genuine aesthetic significance of AI structural creativity , without conflating them and without requiring us to rank them on a single scale.

The observational case this paper presents also points toward an empirical research agenda that philosophy of art has not yet adequately developed: the systematic study of listener responses to AI-generated music under varied framing conditions. This is, at its core, a question about IFC , about how intentional framework shapes perceptual experience , and it is answerable through the kinds of methods that cognitive musicology and experimental aesthetics have been developing for decades. The philosophical questions set the terms of the inquiry; the empirical questions fill them in.

11. Conclusion

The question with which this paper began , how non-living intelligence brings life to music , turns out, on examination, to be a question about the nature of musical life itself. Life in music has never been a property of the generator; it has always been a property of the encounter between structure and perception. The generator matters insofar as it produces structures with the right formal properties. The perceiver matters insofar as they bring the cognitive and interpretive apparatus through which those structures are activated. The human orchestrator of an AI music system matters insofar as the constraint space they design shapes the generative possibility space from which those structures emerge.

What is new about AI is not the principle, but the explicitness. The distributed, mediated, co-constituted nature of musical meaning has always been the case; AI music simply makes it impossible to ignore. The composer who sits at a piano is, in a sense, also an orchestrator , of conventions, schemas, bodily habits, and cultural expectations that they did not invent and cannot fully control. The AI operator works further back in the causal chain, but the chain is continuous. The observational case presented in this paper illustrates one instance of that continuity: a narrative intention, encoded into a constraint space, transmitted through a generative system, and reconstructed , with structural accuracy , by listeners who had no access to the narrative at all.

The concept of Delegated Vitality captures this situation with, this paper argues, more precision than either the source-dependency account or the complete-irrelevance account. Musical vitality is delegated to sound structures by perceivers whose cognitive systems are calibrated to detect the signatures of goal-directed, temporally organized behavior. Those structures can be produced by living composers, by natural phenomena, and by AI systems. In each case, the vitality the listener experiences is real. What varies is the nature and location of the human contribution to the conditions that made it possible.

That contribution, in the AI case, is orchestrative intelligence: the capacity to design generative conditions that reliably produce structures of sufficient aesthetic complexity to activate musical perception in a prepared , or even an unprepared , listener. This is a form of musical expertise. It is not the only form, and it does not replace the others. But it is the form most relevant to the analysis of AI music, and it is the form that will become increasingly central as generative systems develop and as the compositional community develops the constraint intelligence to use them well.

The philosophical work ahead is not to determine whether AI music is legitimate. That question will be settled, as it has always been settled, by the accumulated experience of listeners and musicians over time. The philosophical work is to develop the conceptual tools adequate to describing what is actually happening when it is , and to ensure that those tools are precise enough to distinguish what is genuinely new from what only appears to be.

References

Berkowitz, A. E. (2024). Artificial intelligence and musicking: A philosophical inquiry. *Music Perception, 41*(5), 393–412. https://doi.org/10.1525/mp.2024.41.5.393

Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). Routledge.

Clancy, M. (Ed.). (2022). *Artificial intelligence and music ecosystem*. Routledge.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Collingwood, R. G. (1938). *The principles of art*. Oxford University Press.

Croce, B. (1902/1909). *Aesthetic as science of expression and general linguistic* (D. Ainslee, Trans.). Macmillan.

Dahlstedt, P. (2021). Musicking with algorithms: Thoughts on artificial intelligence, creativity, and agency. In E. R. Miranda (Ed.), *Handbook of artificial intelligence for music: Foundations, advanced approaches, and developments for creativity* (pp. 873–914). Springer. https://doi.org/10.1007/978-3-030-72116-9_31

Davies, S. (1994). *Musical meaning and expression*. Cornell University Press.

Gadamer, H.-G. (1960/1989). *Truth and method* (2nd rev. ed., J. Weinsheimer & D. G. Marshall, Trans.). Crossroad.

Hanslick, E. (1854/1986). *On the musically beautiful* (G. Payzant, Trans.). Hackett.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*(2), 243–259. https://doi.org/10.2307/1416950

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.

Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.

Moruzzi, C. (2018). Creative AI: Music composition programs as an extension of the composer's mind. In V. C. Müller (Ed.), *Philosophy and theory of artificial intelligence 2017* (pp. 69–72). Springer. https://doi.org/10.1007/978-3-319-96448-5_8

Pressing, J. (1988). Improvisation: Methods and models. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 129–178). Oxford University Press.

Roads, C. (1996). *The computer music tutorial*. MIT Press.

Schaefer, R. S., Overy, K., & Nelson, P. (2013). Affect and non-uniform characteristics of predictive processing in musical behaviour. *Behavioral and Brain Sciences, 36*(3), 226–227. https://doi.org/10.1017/S0140525X12002373

Scruton, R. (1997). *The aesthetics of music*. Oxford University Press.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. Oxford University Press.

Zbikowski, L. (2002). *Conceptualizing music: Cognitive structure, theory, and analysis*. Oxford University Press.

Zylinska, J. (2020). *AI art: Machine visions and warped dreams*. Open Humanities Press.