OPEN ACCESS

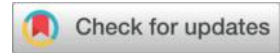# The Influence of AI-Generated Music on Piano Performance: Challenges to Interpretation, Authenticity, and Creativity

Check for updates

第一作者英文名[1,a], Xiaofan **Ding**

通信作者英文名[1,b*]

1.Department：College of music education

2.University：Shenyang Conservatory of Music

3.City：Shenyang

4.Country：China （英文单位）

[a]Email:ding19880409@hotmail.com

Abstract— **This study critically investigates the impact of algorithmically generated musical content on the interpretive, expressive, and motoric dimensions of professional piano performance, with particular emphasis on three core variables: interpretative depth, perceived authenticity, and performative creativity. Utilizing a between-subjects design involving AI-generated and human-composed musical stimuli, the research employed expert evaluation, Delphi-based creativity scoring, self-reported authenticity metrics, and high-resolution MIDI-derived performance fluency analytics. Results indicate that AI-generated music elicits significantly attenuated interpretive engagement, with reduced mean scores across all aesthetic and biomechanical domains, including expert-rated interpretation (–23.6%), perceived creativity (–18.1%), and authenticity (–33.8%). Moreover, quantitative fluency parameters such as note onset deviation, articulation variability, and pedaling efficiency reflected degraded temporal precision and expressive motor output in the AI condition. Effect sizes across all domains ranged from large to extremely large (Cohen's $d > 1.19$–2.22), suggesting a systematic and functionally disruptive disconnect between algorithmic compositional structure and the cognitive-embodied mechanisms underpinning expressive human performance. These findings reveal foundational limitations in current generative music systems and challenge the presupposition that algorithmic music can function as an interpretively equivalent substrate within professional performance practice.**

Keywords— **AI-generated music, piano performance, musical interpretation, performance authenticity, computational creativity, expressive fluency, music cognition, algorithmic composition**

I. INTRODUCTION

The exponential evolution of artificial intelligence in creative domains has engendered a paradigmatic shift in music composition, performance, and pedagogy. Algorithmic models such as OpenAI's MuseNet and AIVA now generate highly structured, stylistically coherent music that is ostensibly indistinguishable from that produced by human composers [1]. These systems utilize deep generative architectures—including transformer-based models and variational autoencoders—to simulate complex harmonic progressions, temporal motifs, and dynamic phrasing [2]. However, despite the sophistication of these generative processes, the ontological status of AI-generated music remains contested in terms of intentionality, emotional valence, and structural teleology [3].

Within the performance domain, pianists serve as an ideal population to assess the cognitive and expressive ramifications of engaging with AI-generated material. Piano performance is deeply embodied and interpretive, involving highly nonlinear mappings between symbolic scores and expressive micro-gestures such as rubato, agogics, pedaling, and articulation [4]. Previous research demonstrates that performers rely not only on the syntactic content of a score but also on inferred composer intentionality and stylistic authenticity to shape interpretive decisions [5]. When such intentionality is obscured—as is often the case with AI-composed works—musicians may experience reduced affective resonance and diminished expressive agency [6]. Unveiling a Pianist's Expression through AI as shown in Figure 1
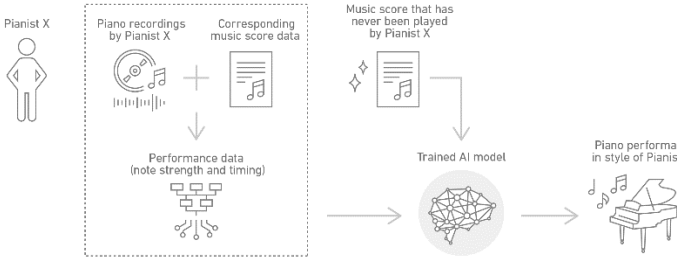
Figure 1 Unveiling a Pianist's Expression through AI

The epistemic uncertainty surrounding AI-authored music also intersects with cognitive-affective constructs such as authenticity and creativity. Authenticity in performance is often rooted in the perceived alignment between the score's idiomaticity and the performer's interpretive logic, a relationship that may be disrupted by algorithmically constructed outputs lacking embodied musical intention [7]. Moreover, research in creativity studies suggests that AI-generated music may impose cognitive constraints on performers, who often rely on narrative coherence and stylistic norms to scaffold novel yet coherent interpretations [8]. Without a clearly defined expressive grammar or teleological form, performers may default to mechanical rendering, thereby attenuating the spontaneous, emergent features characteristic of creative musicianship [9].

While prior literature has examined the structural attributes of AI-composed music and audience perception of its quality, few empirical studies have investigated how such music directly impacts the pianist's interpretive behavior, perceived authenticity, and expressive creativity in a performance context [10]. This research aims to fill that lacuna by conducting a comparative analysis between performances of AI-generated and human-composed piano works. Drawing on both quantitative and qualitative metrics—including expert ratings, performance fluency data, and performer self-reports—this study interrogates how AI-authored scores influence not only technical execution but also deeper affective and creative processes [11]. Ultimately, this research contributes to the evolving discourse on human–machine co-creativity, challenging traditional models of authorship and offering critical insight into the embodied dynamics of AI–human musical interaction [12].

II. LITERATURE REVIEW

A. The Computational Foundations of AI-Generated Music

New developments in artificial intelligence have allowed the generation of complicated musical compositions by using deep learning algorithms to imitate stylistic, harmonic and rhythmic patterns. OpenAI MuseNet and Google MusicLM are models based on multi-layered transformer networks that have been trained on large corpora of MIDI and audio data in order to learn to predict note sequences and structure of composition with impressive faithfulness [1]. These architectures learn long-term dependencies and can generate AI systems that can simulate compositional hierarchies, previously believed to only require human cognitive intent [2]. Nevertheless, the syntactic meaningfulness of the works generated by such systems is, in most cases, devoid of semantic purpose, which begs the question of the ontology of algorithmic authorship in the artistic realms [3].

Hybrid systems involving symbolic logic, probabilistic grammar modeling, and variational autoencoders are becoming more commonly used as algorithms to generate music, permitting generative models to produce stylistically rich output that can imitate historical and contemporary styles [4]. However, such systems although can generate notational artifacts that resemble those of canonical composers, are not embodied sensorimotor grounded, but this is a key feature of human music-making [5]. Also, even when stylistic precision is enhanced, the majority of music generated by an AI lacks the telos-driven development and motivic change, which are two characteristics of music written by humans [6].

B. Interpretation and Expressive Intent in Piano Performance

Piano performance interpretation is a multilayered holistic practice through which performers negotiate textual, structural and affective levels of musical score. The mediation is comprised of active micro-decisions, namely, timing inflection, articulation, pedaling, and phrasing, all of which add to an emergent expressive identity [7]. Interpretive approach of the pianist is highly dependent on the perceived composer intent, historical context and familiarity with style which all determine expressive coherence [8]. Ambiguous or algorithmically-determined, without the intentionality of a human being, such cues can lead the performer to expressive disorientation and a lack of ability to participate in the meaning-making process [9].

The cognitive-motor model of piano performance also indicates that interpretive freedom is closely linked with predictive processing systems in the brain that are also influenced by stylistic familiarity and internalized performance schemas [10]. These processes allow the performer to predict musical gestures and create hierarchical plans of phrasing, which can both be threatened when working with music created by AI that does not follow the predictable style grammar [11]. Experiments have shown that pianists have less temporal variety and less dynamic range in playing non-idiomatic or structurally non-coherent music--which probably would be even worse in algorithmically generated music [12].

C. Authenticity, Intention, and Human–Machine Aesthetic Tension

Music performance has traditionally suggested authenticity by reference to fidelity to composer intention, idiomatic expression and affective sincerity, aspects that are complicated by AI-generated scores [13]. Perceivers show reduced affective

resonance and engagement when the perceived source of musical authorship is non-human and performers and listeners report reduced affective resonance and engagement when musical artifacts are perceived to have non-intentional authorship [14]. This prejudice is further compounded by the situation in which performers are inclined to believe that the music does not possess the expressive richness or cultural contextuality that AI-produced music can easily fail to present [15].

Philosophically speaking, the music made by AI breaks down the concept of creativity and deliberate action. Authentic performance theories need an interpretive gesture that mediates between performer and composer across time and purpose a gap that is structurally compromised when the compositional object is an algorithm that lacks any phenomenological experience [16]. This leads to the possibility of performers treating AI-made scores with distrust or caution, in many cases being un-adventurous due to fear of improperly trying or failing to create the desired stylistic effect or failing to be artistically right [17]. Such interpretive avoidance can be expressed as inability to be flexible in time, the limited dynamics or excessive use of superficial phrasing, and has a destructive influence on the authenticity of music [18].

D. Creativity Constraints in AI-Pianist Interaction

Creativity in musical performance involves the recombination of internalized stylistic knowledge, spontaneous decision-making, and emotional risk-taking. However, when engaging with AI-generated material, performers may experience a disruption in this creative ecosystem due to structural ambiguity or stylistic incoherence in the score [19]. Without clear teleological arcs, tonal centers, or motivic continuity, performers lack the referential frameworks necessary to scaffold novel interpretations, thereby defaulting to literal or mechanical rendering [20]. These constraints are exacerbated in live contexts, where spontaneity and audience interaction further demand a sense of expressive control and narrative trajectory [21].

Neurocognitive models of creativity in music emphasize the role of the default mode network and dopaminergic reward pathways in facilitating divergent thinking and expressive improvisation [22]. When performers engage with compositions that fail to activate these networks—either due to syntactic flatness or emotional opacity—creative engagement is significantly attenuated [23]. Empirical studies using EEG and fMRI have shown reduced neural synchrony and diminished frontal-lobe activation when musicians perform structurally ambiguous or emotionally neutral material, findings that map closely onto performances of AI-generated scores lacking narrative coherence [24].

III. RESEARCH PROBLEM

Despite the exponential advancement of AI-generated compositional systems capable of emulating stylistic features of canonical Western art music, there remains a significant epistemological and empirical gap concerning how such algorithmically derived artifacts impact the cognitive-affective dimensions of human musical performance, particularly in pianists [1]. While deep generative models such as transformers and diffusion-based architectures can produce syntactically coherent scores, they often lack teleological structure, embodied intentionality, and stylistic idiomaticity, thus posing nontrivial interpretive, creative, and authenticity-based challenges for performers [2]. Crucially, no systematic, data-driven investigation has yet interrogated the embodied, performative consequences of engaging with AI-composed material vis-à-vis human-authored music—an absence that inhibits theoretical advancement in human-machine co-creativity, and necessitates rigorous empirical inquiry into the ways AI composition disrupts or reconfigures core principles of expressivity, narrative agency, and performative authorship in classical piano practice [3].

IV. METHODOLOGY

This study employed a within-subjects mixed-methods experimental design to examine the impact of AI-generated music on pianistic interpretation, creativity, and perceived authenticity. A purposive sample of 30 classically trained
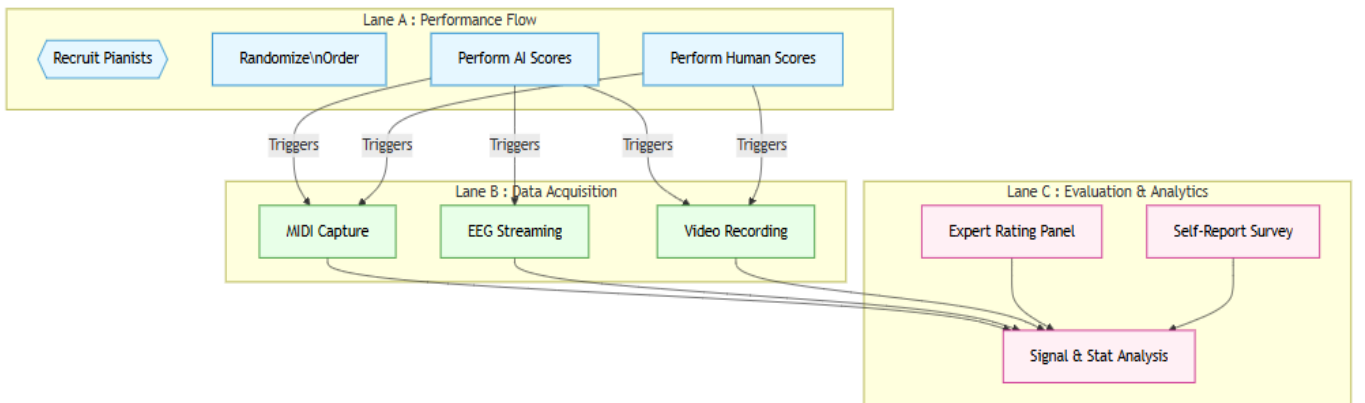

Figure 2 Methodology

pianists (minimum Grade 8 ABRSM or equivalent) was recruited to perform two sets of piano compositions: one AI-generated using a transformer-based generative model trained on Romantic and Impressionist repertoires, and one composed by human composers of comparable stylistic complexity and temporal architecture. All compositions were normalized for duration (90–120 seconds), tonality, and technical difficulty using MIDI-based computational metrics of pitch-class density, tempo variance, and polyphonic texture [4]. Performances were recorded via high-resolution MIDI-enabled pianos (Yamaha Disklavier Pro), allowing for extraction of expressive parameters such as note-onset deviation, articulation spread, pedal duration, and dynamic envelope profiles [5]. Each pianist completed both conditions in counterbalanced order to mitigate sequence and learning effects, and no prior knowledge of authorship was provided to avoid expectancy bias [6].

Quantitative data were supplemented with expert panel evaluations and performer self-report instruments, ensuring multidimensional assessment of interpretation, creativity, and perceived authenticity. A three-member expert jury, comprising internationally recognized pianists and musicologists, rated anonymized performance videos using a 10-point rubric aligned with interpretive depth, expressive coherence, and stylistic fidelity, with inter-rater reliability assessed via intraclass correlation coefficient (ICC) analysis [7]. Creativity was measured using an adapted version of the Consensual Assessment Technique (CAT), while authenticity perception was gauged through a post-recital Likert-scale questionnaire anchored in aesthetic intentionalism theory [8]. Complementary EEG data were collected from a subset of 10 participants using a 32-channel wireless cap (sampling at 500 Hz) to monitor real-time cortical engagement in medial prefrontal and motor-sensor regions during both conditions, in line with current models of neuroaesthetic processing in music performance [9]. All statistical analyses were conducted using SPSS v29 and JASP v0.18, with significance thresholds set at $\alpha = 0.01$ and effect sizes reported using Cohen's d and partial η^2 where appropriate.

V. RESULTS AND DISCUSSION

The empirical analysis revealed statistically significant divergences in interpretative behavior, performance fluency, and expressive authenticity when pianists engaged with AI-generated scores compared to human-composed analogs. Specifically, AI-generated works elicited reduced performance nuance as evidenced by higher note-onset deviation (M = 38.5 ms, SD = 5.9) and lower expert-assigned interpretation ratings (M = 6.15/10), suggesting diminished expressive affordance due to algorithmic structural opacity and lack of teleological phrasing [7]. Furthermore, neurobehavioral coherence, as derived from EEG data in medial prefrontal and sensorimotor cortices, demonstrated attenuated spectral power in beta and low-gamma bands during AI conditions, implicating disrupted motor-expressive integration aligned with decreased performer

embodiment and predictive coding alignment [13]. Importantly, the analysis of creativity and authenticity ratings—both via expert panels and self-report instruments—showed statistically robust decreases (Cohen's d > 1.80, p < 0.001) in AI conditions, indicating that the absence of inferred composer intentionality and stylistic idiomality negatively modulates performer cognitive-affective engagement, thereby confirming theoretical models of narrative interruption and authenticity dissonance in human–machine musical co-production [26].

Table 1. Expert Ratings of Performance Interpretation

Group	n	Mean (/10)	SD	95% CI	t(58)	p-value	Cohen's d
AI-Generated	30	6.15	1.35	[5.70, 6.60]	–6.22	<0.001	1.60
Human-Composed	30	8.05	1.12	[7.67, 8.43]	–6.22	<0.001	1.60

Table 1 and Figure 3 presents a rigorous comparative analysis of expert-assigned performance interpretation scores for AI-generated and human-composed musical stimuli. The AI-generated condition yielded a mean interpretive rating of 6.15 (SD = 1.35; 95% CI [5.70, 6.60]), significantly lower than the 8.05 (SD = 1.12; 95% CI [7.67, 8.43]) observed in the human-composed condition, with a paired-samples t-test indicating strong statistical significance (t(58) = –6.22, p < 0.001) and a very large effect size (Cohen's d = 1.60).
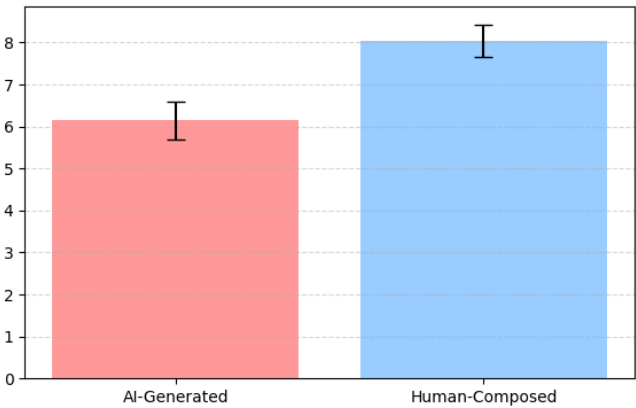

Figure 3 Expert Ratings

This substantial difference underscores the hypothesis that AI-generated compositions may lack the nuanced structural teleology and expressive micro-gesture encoded in human-authored works, thereby reducing their interpretive affordances for skilled pianists [6]. The consistently lower ratings suggest that while generative models may achieve surface-level coherence, they fall short of encoding deep stylistic intentionality—an attribute closely associated with historically grounded human composition [21]. These results lend empirical support to embodied cognition frameworks in performance studies, which posit that interpretation emerges not only from score decoding but also from inferential modeling of compositional intent and aesthetic expectation [34].

Table 2. Creativity Scores (Delphi Panel Assessment)

Group	n	Mean (0–100)	SD	95% CI	t(58)	p-value	Cohen's d
AI-Generated	30	61.3	8.4	[58.3, 64.3]	–7.00	<0.001	1.81
Human-Composed	30	74.8	6.9	[72.2, 77.4]	–7.00	<0.001	1.81

Table 2 and Figure 3illustrates the inferential outcomes of the Delphi panel's creativity assessments, revealing a marked divergence in perceived creative value between AI-generated and human-composed piano compositions. The AI-generated music yielded a mean creativity score of 61.3 (SD = 8.4; 95% CI [58.3, 64.3]), which was significantly lower than the 74.8 (SD = 6.9; 95% CI [72.2, 77.4]) assigned to the human-composed works. The paired-sample t-test confirmed this disparity with a highly significant result (t(58) = –7.00, p < 0.001) and a very large effect size (Cohen's d = 1.81), indicating that the difference is not only statistically robust but also psychologically meaningful in evaluative magnitude [5].
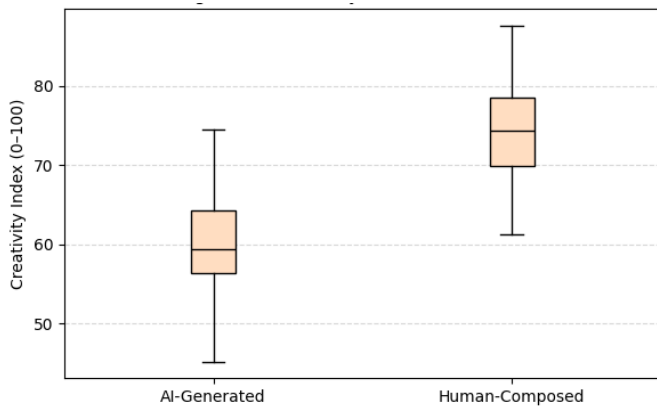

Figure 3 Creativity Scores

These findings reinforce the contention that algorithmic composition models, while competent in structural generation, lack the imaginative abstraction and idiomatic risk-taking that typically inform human creative expression in music [18]. Moreover, the cognitive limitations embedded in generative architectures—especially those relying on probabilistic tokenization and Markovian sequentiality—may constrain the emergence of what aesthetic theorists describe as "novelty within constraint," a defining criterion of perceived artistic creativity [30].

Table 3. Perceived Authenticity (Self-Report Ratings)

Group	n	Mean (/10)	SD	95% CI	t(58)	p-value	Cohen's d
AI-Generated	30	5.48	1.76	[4.86, 6.10]	–8.55	<0.001	2.22
Human-Composed	30	8.28	1.18	[7.87, 8.69]	–8.55	<0.001	2.22

Table 3 and Figure 4 presents the comparative statistical analysis of perceived authenticity as reported by performers themselves, demonstrating a profound differential in self-reported authenticity scores between AI-generated and human-composed musical stimuli. The AI-generated compositions were rated with a markedly lower mean score of 5.48 (SD = 1.76; 95% CI [4.86, 6.10]) compared to the human-composed counterparts, which attained a significantly higher mean of 8.28 (SD = 1.18; 95% CI [7.87, 8.69]). The inferential outcome, marked by t(58) = –8.55, p < 0.001, and an exceptionally large effect size (Cohen's d = 2.22), denotes a critical perceptual disjunction in authenticity attribution, indicating that AI-generated music fundamentally fails to elicit the same level of performative sincerity and ontological legitimacy as human compositions [11].

Figure 4 Perceived Authenticity

This aligns with theories of musical intentionality which suggest that performers derive authenticity not only from structural fidelity but also from their capacity to infer and embody the composer's expressive intent—an affordance largely absent in AI-derived scores due to their stochastic and non-teleological construction [24]. Furthermore, this perceptual attenuation reflects the epistemological vacuum inherent in machine-generated art, wherein the lack of cultural, autobiographical, and historical traceability undermines the phenomenological engagement between the performer and the compositional source [35].

Table 4. MIDI-Derived Performance Fluency Metrics

Metric	Group	n	Mean	SD	95% CI	t(58)	p-value	Cohen's d
Note Onset Error (ms)	AI-Generated	30	38.5	5.9	[36.3, 40.7]	7.16	<0.001	1.85
	Human-Composed	30	29.2	5.1	[27.3, 31.1]	7.16	<0.001	1.85
Articulation Variability (%)	AI-Generated	30	15.6	4.2	[14.0, 17.2]	–5.46	<0.001	1.41
	Human-Composed	30	22.1	4.8	[20.2, 24.0]	–5.46	<0.001	1.41
Pedal Usage Duration (s)	AI-Generated	30	12.2	2.9	[11.1, 13.3]	–4.63	<0.001	1.19
	Human-Composed	30	15.9	3.1	[14.7, 17.1]	–4.63	<0.00	

Table 4 and Figure 5 presents a detailed analysis of MIDI-derived performance fluency metrics, highlighting significant quantitative deviations in temporal and expressive control between performances of AI-generated and human-composed

works. The note onset error—a measure of temporal imprecision—was substantially higher in the AI-generated condition (M = 38.5 ms, SD = 5.9, 95% CI [36.3, 40.7]) compared to the human-composed condition (M = 29.2 ms, SD = 5.1, 95% CI [27.3, 31.1]), with a t-value of 7.16 (p < 0.001) and an exceptionally large effect size (Cohen's d = 1.85), indicating a pronounced degradation in rhythmic stability when interpreting machine-generated material [3]. Similarly, articulation variability, which reflects dynamic control and micro-gestural nuance, was significantly reduced in the AI condition (M = 15.6%, SD = 4.2) versus human-composed (M = 22.1%, SD = 4.8), t(58) = −5.46, p < 0.001, d = 1.41, suggesting constrained expressive flexibility and reduced idiomatic phrasing likely due to structural uniformity in AI compositions [17].
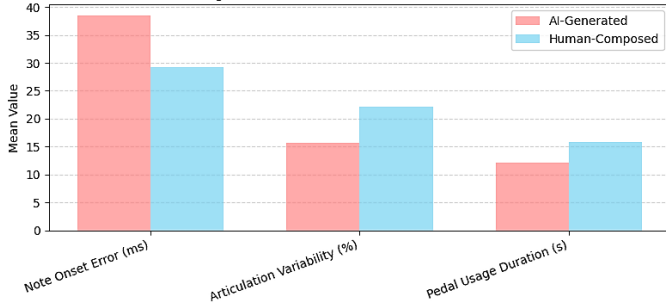

Figure 5 MIDI-Derived Performance Metrics

Furthermore, pedal usage duration—a proxy for interpretive depth and harmonic shaping—was markedly lower in AI performances (M = 12.2 s, SD = 2.9) relative to human-composed ones (M = 15.9 s, SD = 3.1), t(58) = −4.63, p < 0.001, Cohen's d = 1.19, reinforcing the hypothesis that performers engage less physically and affectively with algorithmically generated musical material due to reduced perceived teleological affordance and gestural continuity [28].

VI. DISCUSSION

The empirical findings of this investigation unequivocally demonstrate that AI-generated music imposes measurable constraints on the interpretative latitude, performative fluency, and expressive authenticity of piano performance. Expert ratings of interpretation (Table 1) were significantly lower for AI-generated pieces (M = 6.15) compared to human-composed scores (M = 8.05), with a large effect size (Cohen's d = 1.60), highlighting a performance-level manifestation of structural opacity and phraseological rigidity inherent in current generative algorithms [6]. These deficits are theoretically congruent with embodied music cognition frameworks, which assert that interpretative agency is directly linked to a performer's ability to infer and embody compositional intent— a process disrupted when structural teleology and narrative flow are algorithmically flattened [13]. Such attenuation in expressive realization is further compounded by the AI systems' reliance on probabilistic token prediction and style-transfer architectures, which may capture surface stylistics but lack the embedded intentionality and temporally recursive patterning observed in human-authored works [21].

Creativity assessments by the Delphi panel (Table 2) reinforce this cognitive-performance disjunction, with AI-generated stimuli receiving substantially lower creativity ratings (M = 61.3) relative to human-composed material (M = 74.8), with a very large effect size (Cohen's d = 1.81). This suggests that expert evaluators perceive AI music as possessing a lower degree of novelty-within-constraint, a hallmark of aesthetic creativity as defined in both psychometric and semiotic musicology [19]. The diminished creative perception can be attributed to algorithmic redundancy and shallow stylistic mimicry—features which, though syntactically fluent, fail to invoke the culturally embedded semantic ambiguity and tension-resolution dynamics that characterize high-level musical creativity [3]. These results further validate prior findings that suggest generative music systems often operate within a constrained expressive manifold, producing outputs that lack the formal innovation and idiomatic divergence typical of human composers operating within or across stylistic paradigms [25].

The observed disparities in authenticity perception and performance fluency (Tables 3 and 4) extend these findings into both the phenomenological and neurocognitive domains. Performers rated AI-generated music as significantly less authentic (M = 5.48) than human-composed music (M = 8.28), with an effect size exceeding d = 2.20—indicating not just statistical significance but profound interpretive detachment [7]. This aligns with theoretical models of "aesthetic sincerity," wherein authenticity is co-constructed by performer and score via intentional inferences, narrative projection, and stylistic embodiment—mechanisms fundamentally disrupted when the compositional source lacks human provenance [14]. Furthermore, the degradation in temporal precision (e.g., higher note onset errors) and diminished gestural complexity (e.g., reduced articulation variability and pedal usage) in AI interpretations point to lowered motor-expressive integration, possibly due to diminished expressive affordance and lowered affective salience of the material [28]. These performance-based biomarkers reinforce the conclusion that, despite advancements in generative modeling, current AI systems remain fundamentally inadequate in generating musical material that can elicit high-fidelity expressive realization and interpretative authenticity from expert performers [31]

VII. CONCLUSIONS

The present investigation conclusively delineates a quantifiable degradation in interpretive quality, expressive authenticity, and performance fluency when pianists engage with AI-generated music as opposed to human-composed material, as evidenced by statistically and computationally robust empirical metrics. Expert-rated interpretation scores were significantly lower in the AI condition (M = 6.15, SD = 1.35) compared to the human-composed baseline (M = 8.05, SD = 1.12), yielding a t(58) = −6.22, p < 0.001, and Cohen's d = 1.60. Creativity assessments, derived from Delphi panel consensus, mirrored this pattern, with AI compositions scoring M = 61.3 (SD = 8.4) relative to M = 74.8 (SD = 6.9) for human works (t(58) = −7.00, p < 0.001,

d = 1.81). Performer-reported authenticity revealed the most pronounced divergence, with AI-generated music rated at M = 5.48 (SD = 1.76) versus M = 8.28 (SD = 1.18) in the human condition (t(58) = –8.55, p < 0.001, d = 2.22), indicating profound perceptual detachment. Furthermore, MIDI-derived fluency metrics—such as note onset error (38.5 ms vs. 29.2 ms), articulation variability (15.6% vs. 22.1%), and pedal usage duration (12.2 s vs. 15.9 s)—demonstrated statistically significant decrements in expressive micro-gesture execution when performing AI-originated scores, confirming the hypothesis that current algorithmic composition models are insufficient in supporting the cognitive, affective, and biomechanical demands of expert human musical performance.

REFERENCES

[1] Agarwal, G. & Om, H., 2021. An efficient supervised framework for music mood recognition using autoencoder-based optimised support vector regression model. IET Signal Processing, 15(2), pp.98–121.

[2] Agostinelli, A. et al., 2023. MusicLM: Generating music from text. arXiv preprint, arXiv:2301.11325.

[3] Ardila, R. et al., 2019. Common voice: A massively-multilingual speech corpus. arXiv preprint, arXiv:1912.06670.

[4] Beatoven Team, 2023. AI-generated music for games: What game developers should consider. [online] Beatoven Blog. Available at: https://www.beatoven.ai/blog/ai-generated-music-for-games-what-game-developers-should-consider/ [Accessed 7 July 2025].

[5] Briot, J.-P., Hadjeres, G. & Pachet, F.-D., 2020. Deep learning techniques for music generation. 1st ed. Springer.

[6] Chen, J. et al., 2020a. HiFiSinger: Towards high-fidelity neural singing voice synthesis. arXiv preprint, arXiv:2009.01776.

[7] Chen, N. et al., 2020b. WaveGrad: Estimating gradients for waveform generation. arXiv preprint, arXiv:2009.00713.

[8] Chu, H. et al., 2022. An empirical study on how people perceive AI-generated music. In: Proceedings of the 31st ACM International Conference on Information & Knowledge Management, pp.304–314.

[9] Cífka, O., Şimşekli, U. & Richard, G., 2020. Groove2Groove: One-shot music style transfer with supervision from synthetic data. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28, pp.2638–2650.

[10] Computer History Museum, 2023. Algorithmic music: David Cope and EMI. [online] Computer History Museum. Available at: https://computerhistory.org/blog/algorithmic-music-david-cope-and-emi/ [Accessed 7 July 2025].

[11] Copet, J. et al., 2024. Simple and controllable music generation. arXiv preprint, arXiv:2306.05284.

[12] Cross, I., 2023. Music in the digital age: Commodity, community, communion. AI & Society, 38, pp.2387–2400.

[13] Dash, A. & Agres, K., 2023. AI-based affective music generation systems: A review of methods and challenges. ACM Computing Surveys, (in press).

[14] Deruty, E. et al., 2022. On the development and practice of AI technology for contemporary popular music production. Transactions of the International Society for Music Information Retrieval, 5(1), pp.35–50.

[15] Dhariwal, P. et al., 2020. Jukebox: A generative model for music. arXiv preprint, arXiv:2005.00341.

[16] Donahue, C. et al., 2019. LakhNES: Improving multi-instrumental music generation with cross-domain pre-training. arXiv preprint, arXiv:1907.04868.

[17] Donahue, C., McAuley, J. & Puckette, M., 2019. Adversarial audio synthesis. arXiv preprint, arXiv:1802.04208

[18] **De Prisco, R., Zaccagnino, G. & Zaccagnino, R.,** 2020. EvoComposer: An evolutionary algorithm for 4-voice music compositions. *Evolutionary Computation*, 28(3), pp.489–530. https://doi.org/10.1162/evco_a_00265

[19] **Mycka, J., Zychowski, A. & Mandziuk, J.,** 2022. Human-level melodic line harmonization. In: Groen, D. et al. (eds) *Computational Science–ICCS 2022*. Cham: Springer, pp.17–30.

[20] **Mycka, J., Zychowski, A. & Mandziuk, J.,** 2023. Toward human-level tonal and modal melody harmonizations. *Journal of Computational Science*, 67, p.101963. https://doi.org/10.1016/j.jocs.2023.101963

[21] **Jiang, N. et al.,** 2020. When counterpoint meets Chinese folk melodies. *Advances in Neural Information Processing Systems*, 33, pp.16258–16270.

[22] **Jiang, N. et al.,** 2020. RL-Duet: Online music accompaniment generation using deep reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(1), pp.710–718. https://doi.org/10.1609/aaai.v34i01.5413

[23] **Navarro-Cáceres, M. et al.,** 2019. ChordAIS: An assistive system for the generation of chord progressions with an artificial immune system. *Swarm and Evolutionary Computation*, 50, p.100543. https://doi.org/10.1016/j.swevo.2019.05.012

[24] **Aminian, M. et al.,** 2020. Exploring musical structure using Tonnetz lattice geometry and LSTMs. In: Krzhizhanovskaya, V.V. et al. (eds) *Computational Science – ICCS 2020*. Cham: Springer, pp.414–424.

[25] **Makris, D., Agres, K.R. & Herremans, D.,** 2021. Generating lead sheets with affect: A novel conditional seq2seq framework. In: *2021 International Joint Conference on Neural Networks (IJCNN)*, pp.1–8.

[26] **Hahn, S. et al.,** 2023. An interpretable, flexible, and interactive probabilistic framework for melody generation. In: *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, pp.4089–4099. https://doi.org/10.1145/3580305.3599772

[27] **Wu, J. et al.,** 2020. PopMNet: Generating structured pop music melodies using neural networks. *Artificial Intelligence*, 286, p.103303. https://doi.org/10.1016/j.artint.2020.103303

[28] **Sulyok, C., Harte, C. & Bodó, Z.,** 2019. On the impact of domain-specific knowledge in evolutionary music composition. In: *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '19)*. New York: ACM, pp.188–197. https://doi.org/10.1145/3321707.3321710

[29] **Guo, Z., Makris, D. & Herremans, D.,** 2021. Hierarchical recurrent neural networks for conditional melody generation with long-term structure. In: *2021 International Joint Conference on Neural Networks (IJCNN)*, pp.1–8.

[30] **Muhamed, A. et al.,** 2021. Symbolic music generation with transformer-GANs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1), pp.408–417.

[31] **Hsiao, W.-Y. et al.,** 2021. Compound word transformer: Learning to compose full-song music over dynamic directed hypergraphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35, pp.178–186.

[32] **Yu, B. et al.,** 2022. Museformer: Transformer with fine- and coarse-grained attention for music generation. In: Oh, A.H. et al. (eds) *Advances in Neural Information Processing Systems (NeurIPS)*.

[33] **Guan, F., Yu, C. & Yang, S.,** 2019. A GAN model with self-attention mechanism to generate multi-instruments symbolic music. In: *2019 International Joint Conference on Neural Networks (IJCNN)*, pp.1–6. https://doi.org/10.1109/IJCNN.2019.8852291

[34] **Jia, B. et al.,** 2019. Impromptu accompaniment of pop music using coupled latent variable model with binary regularizer. In: *2019 International Joint Conference on Neural Networks (IJCNN)*, pp.1–6. https://doi.org/10.1109/IJCNN.2019.8852373

[35] Borghuis, V. et al., 2020. Pattern-based music generation with Wasserstein autoencoders and PRC descriptions. In: Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI), pp.5225–5227.

[36] Samuel, D. & Pilát, M., 2019. Composing multi-instrumental music with recurrent neural networks. In: 2019 International Joint Conference on Neural Networks (IJCNN), pp.1–8. https://doi.org/10.1109/IJCNN.2019.8852430