Human Creativity Emerging and Enduring Through Embodied–Collective Intelligence

Tatsuya Daikoku

Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan

*** Corresponding author**

Tatsuya Daikoku

Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan, 7-3-1 Hongo, Bunkyo-ku, 113-8656, Tokyo, Japan

Email: daikoku.tatsuya@mail.u-tokyo.ac.jp

Abstract

Creativity in human societies is sustained not by isolated minds but by embodied, collective intelligence. Using music as a model system, we show that creativity rests on the dynamic tuning of uncertainty: moderate, time-varying surprise—often felt as internal bodily sensations—sparks exploration, while group synchrony stabilizes what is new. Because creativity itself is value-neutral, its fate depends on the co-occurrence of ethics, morality, and empathy. Ethical norms bound admissible deviation, and empathic perspective-taking—amplified by bodily and physiological coupling—acts as a social "precision weight" that admits, reshapes, or rejects novelty. From these ingredients emerges field-level intelligence: a process in which groups preserve diversity, selectively integrate deviations, and maintain "just-right" uncertainty under partially connected social structures. This framing explains the resilience of creativity across generations and sets an agenda for empirical work: quantify uncertainty dynamics in embodied interaction, test how ethical and empathic cues regulate admissibility, and identify network conditions that best sustain socially shared creativity. This framing clarifies why creativity endures and suggests a concrete empirical agenda.

Keywords: Collective Intelligence, Uncertainty Dynamics, Embodied Cognition, Ethics, Empathy

1. Introduction

Creativity is a fundamental human capacity that drives societal advancement. It is the ability to generate novel ideas and frameworks and to embed them in forms that can be shared with others. This capacity extends across domains as varied as science and technology, the arts, education, industry, and everyday collaboration (Hennessey & Amabile, 2010). Crucially, creativity is not a transient spark of inspiration; rather, it is a dynamic process that is continuously sustained and renewed across generations and communities. Creativity is not a phenomenon confined to the solitary mind; it emerges, is tested, and ultimately selected through the dynamics of social interaction. Within networks of collaborators and teams, ideas are challenged, refined, and either disseminated or discarded—underscoring that creativity is as relational and collective as it is cognitive (Uzzi & Spiro, 2005; Wuchty et al., 2007). This cumulative renewal of knowledge—often termed the "ratchet effect"—has long been a central theme in cultural evolution research (Dean et al., 2012). Elucidating the mechanisms that sustain creativity is therefore indispensable for understanding the very foundations of sustainable intelligence and the dynamics of cultural development.

A useful contrast for these dynamics is the recent rise of large language models (LLMs). These systems, accessed through natural-language interfaces, have broadened access to knowledge (Brown et al., 2020). Yet their control still relies heavily on post hoc alignment, safety filters by humans, and output-level moderation (Ouyang et al., 2022; Gehman et al., 2020; Weidinger et al., 2022). As a result, vulnerabilities remain, including susceptibility to unexpected behaviors and the reproduction of biases inherited from training data (Bender et al., 2021). Moreover, empirical and theoretical work has shown that when models iteratively incorporate their own generated data into retraining, the distributional diversity erodes and both quality and variability decline—a phenomenon termed "model collapse" (Shumailov et al., 2024). In other words, this trend raises the concern that the very source of creativity—the richness of uncertainty and diversity—may be lost. This contrast forces us to confront a fundamental question: why is creativity in human societies far less vulnerable to collapse? Why, instead, is diversity maintained and even amplified across generations? And what mechanisms sustain such a resilient system?

In this review, we address these questions from the perspective that creativity is sustained not at the level of the individual, but at the level of the collective. We focus in particular on music as a model system. From the standpoint of predictive processing—namely, the interplay between prediction

errors and their precision—music provides an ideal domain in which to examine how novelty is linked to surprise and how new information is assigned "value" or experienced as pleasure (Friston, 2010; Koelsch et al., 2019; Pearce & Wiggins, 2012; Salimpoor et al., 2011). Moreover, decision-making and emotional processes grounded in the norms and violations of musical grammar have been visualized in neural activity (Maess et al., 2001; Koelsch, 2014). The social and collective diversity and universality that arise through music further illuminate a "maintenance system" that enables creativity to be continuously renewed (Mehr et al., 2019). Building on these insights, our aim is to delineate the fundamental mechanisms that sustain the ongoing renewal of creativity in human societies and to propose a theoretical framework that integrates them. Ultimately, we extend this perspective beyond music to domains such as scientific research and human–AI collaboration, thereby advancing an interdisciplinary understanding of the fundamental question: how is creativity not only preserved, but actively amplified across generations?

2. Sustenance of Creativity in Human Societies

2.1. Dynamics of Uncertainty Emerging from Embodied-Collective Intelligence

The sustained renewal of creativity across generations and communities cannot be explained solely by individual's predictive models. Instead, the evidence points toward a fundamentally collective substrate of human predictive processing, continuously shaped through embodied and social interaction. Here, we use "uncertainty" for how unpredictable the next event is (often quantified as entropy) and "surprise" for how much an event violates expectations (prediction error). Using music as a predictive model system, we have developed an integrative approach that couples predictive processing with behavioral and neuroscientific measurements in humans (Daikoku & Tanaka, 2024). This work has analyzed a large corpus of jazz improvisations spanning the early 20th to the 21st century (456 pieces, 78 musicians) with a hierarchical Bayesian statistical learning model (Daikoku et al., 2023; Daikoku, 2023). This analysis revealed that while the acoustic properties of musical signals themselves remain strikingly stable across eras, the temporal dynamics of prediction error (surprise) and entropy (uncertainty) exhibit gradual change, period-specific signatures (Daikoku, 2024). In particular, pitch-related sequences displayed gradual, community-level shifts in their temporal patterns of uncertainty, whereas rhythmic structures—such as fundamental ratios of 1:1 and 1:2 (e.g., a quarter note followed by another quarter note = 1:1, a quarter note followed by a half note = 1:2)—remained relatively invariant across time. The key point is this: historical change is reflected not in the raw

acoustic signal, but in the probabilistic structure of sequences—transition probabilities, their uncertainty, and how these fluctuate over time. These findings suggest that the dynamics of deviation and alignment within shared predictive models are what drive creativity as a collective, emergent form of intelligence rather than as an individual cognitive process.

Complementary studies reinforce this view. In a large-scale experiment with 353 participants, emotional evaluations and bodily maps were collected in response to chord progressions. By constructing a population-level predictive model of chord progression (Daikoku, Tanaka, & Yamawaki, 2024) and manipulating the temporal dynamics of uncertainty and prediction error, we demonstrated that chord progressions with higher uncertainty elicited stronger subjective feelings of creativity and heightened arousal (Daikoku and Tanaka, 2024). Importantly, this effect was tightly linked to interoceptive sensations in the cardiac region, whereas the perception of beauty and pleasure was more strongly associated with predictable and less uncertain chord progressions. These results indicate that fluctuations in high uncertainty, grounded in collective predictive models, recruit embodied pathways to promote the feeling of creativity, whereas aesthetic value is stabilized by predictability (for theoretical background, see Critchley et al., 2004; Seth, 2013). Crucially, these findings are consistent with broader cross-cultural evidence on the embodied grounding of affect and music. Volynets et al. (2020) demonstrated that bodily maps of basic emotions are culturally universal, suggesting that interoceptive patterns provide a shared biological template for linking internal bodily states with affective experiences. Extending this framework, Putkinen et al. (2024) revealed that bodily maps of musical sensations likewise display systematic and cross-culturally consistent patterns. Together, these results highlight that the coupling between predictive dynamics and embodied sensations is not culturally idiosyncratic but reflects a universal mechanism through which human societies anchor and transmit affective and creative experiences. Taken together, these converging findings underscore a key principle: creativity is not simply sustained by novelty per se, but by the temporal dynamics of uncertainty embedded in collective predictive frameworks, which engage the body and its interoceptive signals as mediators of creative experience. This perspective aligns with recent affective neuroscience and music cognition studies showing that both surprise and uncertainty fluctuations shape emotional intensity and social bonding in music (e.g., Vuust et al., 2022). Moreover, the cross-cultural universality of bodily maps of both emotions and musical sensations may suggest that embodied pathways form a robust and species-wide foundation for the resilience and renewal of creativity in human societies.

These observations reinforce the view that value judgments—including those underlying creativity—do not reside within the isolated individual but emerge in social and collective contexts of resonance. Neural evidence has shown that coupling of neural activity between speakers and listeners, or performers and audiences, is a critical determinant of mutual understanding and joint action in language, music, and gesture (Stephens et al., 2010; Hasson et al., 2012). Likewise, physiological synchronization during collective rituals may be linked to the strengthening of social bonds (Konvalinka et al., 2011; Hove & Risen, 2009; Mogan et al., 2017), while synchronized movement more broadly promotes cooperative behavior (Wiltermuth & Heath, 2009). Together, such findings highlight how "collective intelligence" can emerge from the interactions of diverse individuals (Woolley et al., 2010). Diversity itself plays a decisive role: heterogeneous groups consistently outperform homogeneous ones (Hong & Page, 2004), and success in cultural markets often follows unpredictable and unequal trajectories driven by social influence (Salganik et al., 2006). These strands of evidence converge on a central principle: the value of creativity is distributed across a collective/social "field," shaped through the alignment and deviation of mutual predictions.

The computational basis of such dynamics can be situated within predictive processing frameworks, which emphasize the representation and weighting of uncertainty in Bayesian inference (Knill & Pouget, 2004; Körding & Wolpert, 2004). Humans appear to prefer an intermediate rate of information—neither too little nor too much—a phenomenon known as the "Goldilocks effect," observed in infants' allocation of visual and auditory attention (Kidd et al., 2012). This preference for "just-right" uncertainty resonates with musical groove, where moderate deviations in prediction sustain engagement and drive affective responses (Vuust et al., 2022). It also suggests a precision-tuning mechanism that regulates the optimal balance of pleasure, arousal, and immersion that defines creative experience. At the societal level, the dynamics of maintaining uncertainty without excess or depletion resemble probabilistic search algorithms. For instance, the Metropolis–Hastings algorithm accepts even rare deviations with a finite probability, thereby preventing premature convergence and preserving diversity (Metropolis et al., 1953; Hastings, 1970). Experimental studies of cultural evolution further support this analogy: partially connected networks, which avoid overly dense synchronization, foster the accumulation of diverse and complex solutions (Derex & Boyd, 2016).

Taken together, these findings suggest that human creativity is sustained not by optimizing individual internal models, but by collective processes of acceptance and selection mediated through bodily synchronization, emotional resonance, and reciprocal prediction. Just as the nature of human

existence cannot be fully understood by examining cells in isolation, the essence of intelligence and creativity resides in emergent field-level structures that transcend the individual. These structures flexibly regulate the temporal profile of surprise and uncertainty, allowing norms and deviations to coexist in dynamic tension. It may be precisely this capacity for continual adjustment that safeguards prediction uncertainty as the wellspring of creativity.

2.2. Co-occurrence of Ethics, Morality, and Empathy

The social and collective dynamics that sustain creativity do not terminate within individual brain. We propose that three processes intertwine within collective dynamics, forming what can be viewed as an "intelligence of fields": (i) innovative ideas that underlie uncertainty; (ii) ethical/moral constraints that stabilize predictability by delimiting what is acceptable; and (iii) empathy and perspective taking, which bridge the two by rendering innovation intelligible and admissible to others.

Empirical evidence supports the key components of this framework. Individuals with stronger creative problem-solving skills show superior ethical decision-making (N = 258), and creative reasoning can unlock moral insights that reconcile value conflicts—particularly when actors generate multiple "What could I do?" options (Mumford et al., 2010; Zhang, Gino & Margolis, 2018). On the other hand, creativity has a dark side. Creativity can facilitate the self-justification of unethical conduct (e.g., creatively finding loopholes) and are associated with increases in dishonest behavior (Gino & Ariely, 2012). Furthermore, malevolent creativity—novel and effective ideas serving harmful purposes—has been theorized and functionally modeled in domains such as terrorism and crime (Cropley, Kaufman & Cropley, 2008). These findings imply that creativity per se is value-neutral; the field conditions that align innovations with social norms—and thereby shape the probability of acceptance—determine whether novelty crystallizes as ethical innovation or deviates into harmful ingenuity.

Within this regulation, empathy may function as central moderators. Evidence indicates that empathic concern robustly supports prosocial behavior (Eisenberg et al., 2010). Neural evidence also showed that interoceptive processes that can be reflected in the activity of anterior insula (Critchley et al., 2004) and the heartbeat-evoked potentials (HEP) (Fukushima, Terasawa & Umeda, 2011) contribute to the neural basis of empathy. Interoception is the ability to perceive and self-recognize bodily sensations. Notably, the interoceptive inference provides a theoretical account of how prediction

errors couple to affective value (Seth, 2013). In aggregate, empathy (a model of the other) operates as social precision weights that tune the likelihood that novel deviations will be admitted rather than pathologized, thereby curbing malign trajectories. A further lever is cognitive perspective taking, the ability to adopt another person's standpoint and to understand their feelings and ways of thinking—which serves as a foundation for generating empathy. In negotiation, perspective taking facilitates the discovery of hidden agreements and the creation of joint value (Galinsky et al., 2008) and improves outcomes in mixed-motive settings (Gilin et al., 2013). It also increases willingness for intergroup contact and promotes cohesion across social divides (Wang et al., 2014). Put differently, perspective taking may transform the mismatch introduced by exploratory creativity into forms that make sense to others, especially when combined with interoceptive bodily sensation and body synchrony—thereby helping to balance creativity with ethical acceptability.

These mechanisms are particularly vivid when performing music, where interactions among creative behavior, morality/ethics, and empathy prevent both collapse from excessive novelty and stagnation from rigid adherence to norms. In a jazz session, the groove established by the rhythm section scaffolds the soloist's bold and novel insertions; the ensemble and audience momentarily hold their breath as a bodily response to surprise, but then stabilize interpretation and predictability by catching the phrase harmonically and metrically. In these processes, performers, listeners, and co-listeners alike become physically coupled to the music, their bodies moving spontaneously (Janata et al., 2012). Such bodily entrainment naturally gives rise to interpersonal synchrony, which in turn enhances affiliation (Hove & Risen, 2009). Building on this foundation, evidence shows that joint music-making promotes prosocial behavior (Kirschner & Tomasello, 2010), while group singing and synchronous dance strengthen social bonding (Tarr, Launay & Dunbar, 2014; Weinstein et al., 2016). From lullabies and choral singing to the dance floor, this loop re-emerges across contexts, sustaining a dynamic balance between novelty and stability—likely grounded in musical moral–ethical sensibilities and collectively negotiated through embodied empathy and perspective taking.

Meta-analyses indicate that interpersonal synchrony in musical activities (choral singing, ensemble performance, dance) increases cooperation, altruism, and cohesion (Mogan, Fischer & Bulbulia, 2017; Rennung & Göritz, 2016). Synchrony heightens self–other overlap and—sometimes via endogenous opioids—boosts trust and cooperation (Lang et al., 2017). In this light, the ethics of musical practice can be viewed as the coordination rules that sustain "*creative exploration within a shared temporal action space*" by satisfying: (1) care (protecting others' agency and welfare), (2)

fairness (equitable distribution of participation and resources—time, loudness, solos, credit), and (3) authority/norm-respect (adherence to agreed forms, rituals, and copyrights) as foundations in moral foundations theory (Graham et al., 2018). These may be scaffolded by bodily synchrony and perspective taking. The result is the maintenance of socially optimized uncertainty, steering between the twin failure modes of over-stability (rigidity) and over-uncertainty (chaos).

3. Collective Intelligence Emerging from Uncertainty Dynamics Mediated by Creativity, Ethics, and Empathy

The foregoing arguments indicate that the sustainment of creativity cannot be explained from an individual's internal model. Rather, it depends on collective interaction grounded in embodied sensibility. In music, for example, the timing of prediction and deviation is socially shared; the resulting mismatch and surprise are experienced as a communicable sense of creativity through bodily synchrony and resonance among performers and listeners. Crucially, what supports creativity is not novelty per se, but the temporal dynamics of predictive uncertainty. Novelty functions as a probabilistic fluctuation that is continuously modulated by the group and distributed through bodily experience, rather than as a single, isolated jolt.

Creativity is likewise inseparable from ethics/moral/insight and empathy. While it can generate moral insight, it also carries the risk of norm violations and self-justification. This ambivalence is governed by social regulatory mechanisms—bodily synchrony, empathy, and perspective taking. Whether a creative deviation is admitted or rejected by a community is decided within a social field through embodied empathic understanding and its alignment with ethical norms. By "collective," we therefore do not mean a mere aggregation of parallel individuals. The term designates an emergent, field-level intelligence that arises from the dynamics of bodily coordination and misalignment. This intelligence of the field continuously fine-tunes between stabilizing norms and exploratory deviations, avoiding both rigidity and chaos. Through this ongoing micro-regulation, creativity should be continually renewed and transmitted across generations and cultures.

Jazz improvisation exemplifies this process. When a soloist launches a bold phrase, listeners' bodies register a brief rise in uncertainty (e.g., faster heartbeat). The ensemble then absorbs the deviation by adjusting harmony and meter, turning it into shared, intelligible invention. Rather than

rejecting deviation and novelty, the group incorporates momentary misfit into the field, allowing musical creativity to be updated while retaining ethical and empathic meaning. A similar logic operates in choral singing and joint rhythmic activity. Synchrony of voices and movement creates strong collective cohesion. Individual predictions are calibrated with respect to others, and small deviations are folded back into the overall harmony. Here, normative constraints (pitch, rhythm and musical grammar) remain salient, yet micro-deviations are tolerated as the signature of living music, often enriching creative vitality. Bodily synchrony thus sustains ethical order while keeping creativity alive.

The same dynamics appear in human–AI collaboration, particularly with generative systems. Model outputs often contain unexpected deviations or noise. Through human interpretation and empathic uptake, these fragments can be converted into creative ideas; without such alignment to empathic and ethical constraints, they are dismissed as harmful or meaningless. Hence, even in human–AI settings, embodiment- and empathy-based social regulation is pivotal, opening the possibility of a genuinely co-creative field intelligence spanning humans and machines.

Taken together, uncertainty dynamics mediated by creativity, ethics, and empathy are observable in musical practice. What persists is not a simple sum of individual knowledge but a collective intelligence of the field that rises through the interplay of misalignment and coordination. This structure plausibly underwrites the robustness and sustainability of human creativity.

4. Discussion

Intelligence in human societies is rarely the mere sum of individual abilities. Rather, it frequently arises from the space of interaction—the dynamic field of exchanges among people. This emergent capacity is often termed collective intelligence, and group performance on complex tasks is frequently explained less by the distribution of individual IQs than by the quality of coordination within the group (Woolley et al., 2010). Beyond a simple aggregation of minds, intelligence is distributed across bodies, temporal synchrony, and environments, thereby functioning as a field-level phenomenon that extends cognition into the world (Clark & Chalmers, 1998). This paper refers to this distributed, interaction-born capacity as field-level (collective) intelligence.

Contemporary neuroscience frames the brain as a predictive organ that minimizes prediction error (Friston, 2010; Knill & Pouget, 2004). Crucially, this predictive machinery is not confined within

individuals: during communication, neural activity couples between speakers and listeners, supporting shared understanding and joint action (Stephens et al., 2010; Hasson et al., 2012). In this sense, prediction and error correction are not only intra-cranial processes but also shared across the social field that links brains together. Music is a particularly tractable model system for examining such collective, field-level intelligence. Musical experience hinges on probabilistic "grammars": listeners form expectations about upcoming notes and chords, and affective responses arise from the balance of confirmation and violation (Koelsch, 2014; Pearce & Wiggins, 2012). A moderate degree of fluctuation of predictability—neither too little nor too much—maximizes pleasure and movement tendencies (Vuust et al., 2022; Stupacher et al., 2022). Harmony violations recruit regions that also support syntactic processing in language, including Broca's area (Maess et al., 2001). Put simply, when listening to Western tonal music, many listeners strongly expect a tonic after a dominant. Resolution produces feelings of beauty and relief; deliberate detours produce surprise and tension which, when subsequently integrated back into the prevailing grammar, are often shared and appraised as creative (Daikoku, Tanaka & Yamawaki, 2024). Thus, musical affect emerges from time-varying play between prediction and deviation that is scaffolded by culturally shared grammatical regularities. Music thereby functions as a collective prediction device—a window onto the mechanisms of field-level intelligence.

The substrate of this field is bodily. Interoceptive pathways—particularly those involving the anterior insula—support predictions about cardiac and visceral states and the processing of interoceptive prediction errors, shaping emotion and the sense of self (Critchley et al., 2004; Seth, 2013). People also share broadly similar bodily maps of emotion—for example, heat or pressure in the chest for anger or sadness—suggesting a common bodily "lexicon" for affect (Nummenmaa et al., 2014; Daikoku, Minatoya & Tanaka, 2025). These embodied templates provide a shared substrate through which affective experiences can be communicated and socially stabilized, thereby supporting collective creativity. Recent work extends these insights directly to music. Cross-cultural mapping shows that bodily sensations evoked by music exhibit systematic topographies that align across Western and East Asian participants—for example, danceable, joyful tracks map more strongly to the limbs, while gentle, sad tracks map to the chest (Putkinen et al., 2024). This pattern generalizes Nummenmaa's bodily maps of basic emotions to the musical domain, suggesting that the interaction among prediction of external phenomenon, bodily sensation, and emotion is, in broad outline, culturally shareable (Putkinen et al., 2024). Such a shared bodily communication makes it easier for performers and audiences to achieve understanding and synchrony, providing a scaffold for collective intelligence. Complementing this, body-map studies focused on the felt sense of creativity show that

the experience of creativity tracks the temporal dynamics of uncertainty and surprise in chord progressions. Higher entropy and surprise increase ratings of "felt creativity" and are accompanied by stronger sensations in the cardiac region; by contrast, beauty and pleasantness peak under more predictable, low-uncertainty progressions. Individuals high in interoceptive sensibility report stronger creative feelings and clearer bodily sensations (Daikoku & Tanaka, 2024). Hence, creative experience is likely mediated not only by informational novelty but also by interoceptive prediction-error processing and its conscious access. In musical settings, momentary mismatches can thus be absorbed via the body into empathic synchrony, and thereby transformed into socially shared judgments of "creative and interesting".

Taken together, (1) cross-culturally shared emotion–body mappings (Putkinen et al., 2024) and (2) the coupling between uncertainty dynamics and interoception in the experience of creativity (Daikoku & Tanaka, 2024) jointly explain why music readily functions as a collective prediction device. On a widely legible bodily substrate, moderate fluctuations in predictability ignite a sense of creativity that is then distributed through bodily synchrony and empathy into the shared field. This circulation—the rise of creative feeling on a shared bodily code, its uptake through synchrony, and its stabilization in collective appraisal—captures a core mechanism of field-level intelligence.

A distinctive strength of collective intelligence is its ability to maintain diversity while keeping uncertainty at a just-right level. Heterogeneous groups tend to outperform homogeneous high-ability groups (Hong & Page, 2004). In cultural evolution experiments, partially connected networks (in which not everyone is tightly linked to everyone else, slowing premature consensus and preserving diversity)—avoiding over-dense coupling—accumulate more diverse and complex solutions than fully connected ones (Derex & Boyd, 2016). Algorithmically, such dynamics echo search schemes that accept rare deviations with finite probability (e.g., Metropolis–Hastings), thereby preventing premature convergence (Metropolis et al., 1953; Hastings, 1970). The Collective Predictive Coding (CPC) framework (Taniguchi et al., 2016; Taniguchi, 2025; Taniguchi et al., 2025) seeks to formalize these ideas computationally. CPC emphasizes that predictive models are not only individual constructs but are also generated and calibrated within the social field, which regulates whether deviations are admissible as shareable innovations. This view aligns with theories of symbol emergence, in which meanings and norms arise not top-down but through interaction, progressively stabilizing within the field (Taniguchi et al., 2016).

These “musical collective intelligence” principles may sketch design cues for future AI. Current AI systems—especially LLMs—lack humanlike bodies and therefore struggle to ground symbols in perception and action. This long-standing symbol grounding problem highlights a structural gap between current AI and embodied human cognition (Harnad, 1990). Decades of work in grounded and extended cognition show that human understanding is deeply embedded in body and environment, implying qualitative differences between human learning and prevailing AI pipelines (Barsalou, 2008). The gap is stark for affect: inferring emotion from surface labels or facial/vocal cues alone is insufficient, casting doubt on whether present AI understands emotion in a human-relevant sense (Barrett et al., 2019). Although affect detection using physiological signals has progressed (Calvo & D’Mello, 2010; Koelstra et al., 2011), AI systems themselves lack interoceptive generative models that predict and correct internal bodily states—the very machinery that supports affect and self in humans (Critchley et al., 2004; Seth, 2013). Consequently, they may poorly be coupled to the bodily synchrony and empathy that scaffold field-level intelligence.

Music suggests how to close part of this gap. In musical practice, deviations are not discarded as noise but admitted as creativity when they are absorbed through bodily sensation and empathic uptake. Translating this into AI design requires going beyond rule-based governance of morality and ethics to a sensory–affective alignment layer in which systems optionally respond to bio-backed signals—heart rate, respiration, or HRV—under strict privacy-preserving protocols (Critchley et al., 2004; Thayer et al., 2012). To avoid model collapse caused by training on self-generated data (Shumailov et al., 2024), novelty should be filtered through multi-agent, heterogeneous reception rather than a single average preference, leveraging diversity’s first-order benefits and partially connected evaluation networks that preserve exploration (Derex & Boyd, 2016). In parallel, interactions among humans and AIs can incorporate synchrony-like signals—overlap in attentional allocation, coordinated response latencies, and measures of interpretive agreement—coupled to ethical metrics so that the admissibility of deviations is tuned by the state of the shared field. A further lesson is to regulate “just-right” uncertainty. Inspired by the Goldilocks effect, systems should maintain moderate predictive variability—neither boring nor chaotic—by adaptively controlling sampling temperature, search breadth, and presentation order (Kidd et al., 2012). Conceptually, these ingredients align with Collective Predictive Coding: meanings and norms emerge through interaction (Taniguchi et al., 2016; Taniguchi et al., 2025), and the social field dynamically assigns precision to deviations—accepting, reshaping, or rejecting them in context. Operationalizing these mechanisms would begin to instantiate field-level intelligence in AI, opening a path toward systems that cohere with human social

dynamics rather than merely optimizing in isolation.

Several limitations qualify the scope of this review. First, although we have used music as a model system, the extent to which the proposed mechanisms generalize to other coordination-intensive domains—scientific collaboration, education, or collective decision-making—remains uncertain. Even where cross-cultural similarities in bodily maps for music are observed, cultural ecologies may still shape norms for "admissible deviation" and sanctioning in ways we have not fully captured. Second, constructs such as field-level intelligence and admissibility of deviation risk remaining theoretical unless tied to operational definitions. They should be validated against longitudinal outcomes—innovation uptake, retention of knowledge—and aligned with existing findings that partially connected networks help preserve diversity (Derex & Boyd, 2016), thereby providing a coherent evaluative framework. Third, while the "Goldilocks" hypothesis for creative engagement is promising, the optimal uncertainty is likely task-, expertise-, and risk-dependent, plausibly nonlinear and hysteretic. Infant learning and musical groove studies establish the principle (Kidd et al., 2012; Vuust et al., 2022), but parameterizing it for complex adult decisions will require closed-loop, adaptively tuned experimental designs. Fourth, our AI proposals—synchrony-like interaction signals and bio-informed sensory–affective alignment—are, at best, proxies for interoception rather than true interoceptive generative models. Their limits must be quantified against strong non-embodied baselines to demonstrate the incremental value of embodied proxies. Finally, there is no standard benchmark for field-level intelligence. Community-agreed tasks and metrics are needed to jointly test collapse resistance under self-generated data, diversity preservation, and socially accepted novelty under partial connectivity; without such shared benchmarks, cumulative progress will be difficult.

Acknowledging these constraints, this paper nevertheless advances a framework that centers field-level intelligence—beyond individual internal models—as a key to the persistence of creativity in human societies. Using music as a tractable model, we have highlighted how dynamic fluctuations between prediction and deviation are socially shared via bodily sensation and empathy, and how this circulation sustains creative capacity. We further argued that these insights can be extended to scientific collaboration, education, and human–AI co-creativity, offering a theoretical basis for explaining why creativity does not merely endure across generations but is actively renewed.

6. Conclusion

Creativity persists in human societies not as a property of isolated minds but as a phenomenon distributed across bodies, artifacts, and social relations. Viewing creativity through the lens of predictive processing and musical practice reveals a simple but powerful logic: moderate, time-varying uncertainty ignites exploration; interoceptive sensation and empathic uptake translate deviation into shared value; and partially connected social structures preserve diversity while preventing premature convergence. In this dynamic, music functions as a collective prediction device, making visible the operational signature of field-level intelligence.

This review synthesized evidence that prediction–deviation dynamics are stabilized and renewed by bodily synchrony, shared grammars, and socially negotiated admissibility. It also argued that these mechanisms generalize as design principles: calibrate uncertainty toward a “just-right” regime, filter novelty through heterogeneous, multi-agent reception, and couple interaction to synchrony-like signals that reflect the state of the shared field.

A concrete research agenda follows. First, formalize field-level constructs—social precision weighting, admissibility of deviation, collapse resistance—into operational metrics and community benchmarks. Second, develop closed-loop, ecologically valid experiments that manipulate uncertainty and synchrony while tracking longitudinal outcomes such as innovation uptake and knowledge retention. Third, prototype sensory–affective alignment layers that are strictly opt-in, privacy-preserving, and auditable, and evaluate their value over strong non-embodied baselines. Finally, embed these components within collective predictive coding frameworks so that meanings and norms can emerge and stabilize through interaction rather than prescription.

In sum, the persistence of creativity is supported not solely within individual brains but across the field—through interoceptive, empathic coupling and culturally shared musical grammars and bodily maps. Music makes this dynamics especially visible: moderate uncertainty elicits a sense of creativity that, via bodily synchrony and empathy, is socially shared and stabilized. This is the operational signature of field-level intelligence.

Competing Interests

The authors declare no competing financial interests.

Author Contributions

T.D. wrote the draft of the manuscript and figure. T.D. edited and finalized the manuscript.

Acknowledgements

This research was supported by the JSPS KAKENHI (25K03206, 25H00448, 24K22302), Japan. The funding sources had no role in the decision to publish or prepare the manuscript.

Reference

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?🦜. In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (pp. 610-623).

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. IEEE Transactions on affective computing, 1(1), 18-37.

Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7–19.

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. Nature neuroscience, 7(2), 189-195.

Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2008). Malevolent creativity: A functional model of creativity in terrorism and crime. Creativity Research Journal, 20(2), 105-115.

Daikoku, T., Minatoya, M., & Tanaka, M. (2025). Mapping Emotional Feeling in the Body: A Tripartite Framework for Understanding the Embodied Mind. arXiv preprint arXiv:2504.14865.

Daikoku, T., Kamermans, K., & Minatoya, M. (2023). Exploring cognitive individuality and the underlying creativity in statistical learning and phase entrainment. EXCLI journal, 22, 828.

Daikoku, T. (2023). Temporal dynamics of statistical learning in children's song contributes to phase entrainment and production of novel information in multiple cultures. Scientific Reports, 13(1), 18041.

Daikoku, T. (2024). Temporal dynamics of uncertainty and prediction error in musical improvisation across different periods. Scientific Reports, 14, 22297.

Daikoku, T., & Tanaka, M. (2024). *Body Maps for Feeling of Creativity in Musical Chord Progression*. arXiv:2410.11885.

Daikoku, T., & Tanaka, M. (2024). Protocol to visualize a bodily map of musical uncertainty and prediction. STAR protocols, 5(4), 103473.

Daikoku, T., Tanaka, M., & Yamawaki, S. (2024). Bodily maps of uncertainty and surprise in musical chord progression and the underlying emotional response. Iscience, 27(4), 109498.

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. Science, 335(6072), 1114-1118. doi: 10.1126/science.1213969

Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. Proceedings of the National Academy of Sciences, 113(11), 2982-2987.

Eisenberg, N., Eggum, N. D., & Di Giunta, L. (2010). Empathy-related responding: Associations with prosocial behavior, aggression, and intergroup relations. Social issues and policy review, 4(1), 143-180.

Fukushima, H., Terasawa, Y., & Umeda, S. (2011). Association between interoception and empathy: evidence from heartbeat-evoked brain potential. International Journal of Psychophysiology, 79(2), 259-265.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*, 127–138.

Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. Psychological science, 19(4), 378-384.

Gehman, S., Gururangan, S., Sap, M., Choi, Y., & Smith, N. A. (2020). RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models. Findings of EMNLP 2020, 3356–3369.

Gilin, D., Maddux, W. W., Carpenter, J., & Galinsky, A. D. (2013). When to use your head and when to use your heart: The differential value of perspective-taking versus empathy in competitive interactions. Personality and social psychology bulletin, 39(1), 3-16.

Gino, F., & Ariely, D. (2012). The dark side of creativity: original thinkers can be more dishonest. Journal of personality and social psychology, 102(3), 445.

Graham, J., Haidt, J., Motyl, M., Meindl, P., Iskiwitch, C., & Mooijman, M. (2018). Moral foundations theory: On the advantages of moral pluralism over moral monism. In K. Gray & J. Graham (Eds.), Atlas of moral psychology (pp. 211–222). The Guilford Press.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. Trends in cognitive sciences, 16(2), 114-121.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109.

Hennessey BA, Amabile TM. Creativity. Annu Rev Psychol. 2010;61:569-98. doi: 10.1146/annurev.psych.093008.100416.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proceedings of the National Academy of Sciences, 101(46), 16385-16389.

Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition, 27*(6), 949–961.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. Journal of experimental psychology: general, 141(1), 54.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. TRENDS in Neurosciences, 27(12), 712-719.

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. Nature, 427(6971), 244-247.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. PloS one, 7(5), e36399.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. Evolution and human behavior, 31(5), 354-364.

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. IEEE transactions on affective computing, 3(1), 18-31.

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience, 15*(3), 170–180.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences, 23*(1), 63–77.

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E. M., Wallot, S., ... & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. Proceedings of the National Academy of Sciences, 108(20), 8514-8519.

Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. Biological psychology, 127, 191-197.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4*(5), 540–545.

Mehr, S. A., et al. (2019). Universality and diversity in human song. Science, 366(6468), eaax0868.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. The journal of chemical physics, 21(6), 1087-

1092.

Mogan, R., Fischer, R., & Bulbulia, J. (2017). To be in synchrony or not? Meta-analysis of synchrony's effects on behavior, perception, cognition and affect. Journal of Experimental Social Psychology, 72, 13–20.

Mumford, M. D., Waples, E. P., Antes, A. L., Brown, R. P., Connelly, S., Murphy, S. T., & Devenport, L. D. (2010). Creativity and ethics: The relationship of creative and ethical problem-solving. Creativity Research Journal, 22(1), 74-89.

Nummenmaa, L., Glerean, E., Hari, R., & Hietanen, J. K. (2014). Bodily maps of emotions. Proceedings of the National Academy of Sciences, 111(2), 646–651.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in neural information processing systems, 35, 27730-27744.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory Expectation: The Information Dynamics of Music Perception and Cognition. *Topics in Cognitive Science, 4*(4), 625–652.

Putkinen, V., Zhou, X., Gan, X., Yang, L., Becker, B., Sams, M., & Nummenmaa, L. (2024). Bodily maps of musical sensations across cultures. Proceedings of the National Academy of Sciences, 121(5), e2308859121.

Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony. Zeitschrift für Psychologie.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. science, 311(5762), 854-856.

Salimpoor, V. N., et al. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience, 14*(2), 257–262.

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. Trends in cognitive sciences, 17(11), 565-573.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. Proceedings of the national academy of sciences, 107(32), 14425-14430.

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. Nature, 631(8022), 755–759.

Stupacher, J., Witek, M. A. G., Vuust, P., Schroeder, S. C., & Kringelbach, M. L. (2022). Music, rhythm, and the brain: Synchronization, anticipation, and the pleasure of groove. Neuroscience & Biobehavioral Reviews, 137, 104619.

Taniguchi, T., Takagi, S., Otsuka, J., Hayashi, Y., & Hamada, H. T. (2025). Collective predictive coding as model of science: Formalizing scientific activities towards generative science. Royal Society Open Science, 12(6), 241678.

Taniguchi, T., Nagai, T., Nakamura, T., Iwahashi, N., Ogata, T., & Asoh, H. (2016). Symbol emergence in robotics: A survey. Advanced Robotics, 30(11-12), 706–728.

Taniguchi, T. (2025). Beyond individuals: Collective predictive coding for memory, attention, and the emergence of language. Cognitive Neuroscience, 1-2. https://doi.org/10.1080/17588928.2025.2518942

Tarr, B., Launay, J., & Dunbar, R. I. (2014). Music and social bonding:"self-other" merging and neurohormonal mechanisms. Frontiers in psychology, 5, 1096.

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers III, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neuroscience & Biobehavioral Reviews, 36(2), 747-756.

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology, 111*(2), 447–504.

Volynets, S., Glerean, E., Hietanen, J. K., Hari, R., & Nummenmaa, L. (2020). Bodily maps of emotions are culturally universal. Emotion, 20(7), 1127.

Vuust, P., Witek, M. A. G., Dietz, M. J., Kringelbach, M. L., & Overy, K. (2022). Now you hear it: A predictive coding model of music perception. Psychological Review, 129(2), 234–259.

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. Nature Reviews Neuroscience, 23(5), 287-305.

Wang, C. S., Kenneth, T., Ku, G., & Galinsky, A. D. (2014). Perspective-taking increases willingness to engage in intergroup contact. PloS one, 9(1), e85681.

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. S., Mellor, J., ... & Gabriel, I. (2022, June). Taxonomy of risks posed by language models. In Proceedings of the 2022 ACM conference on fairness, accountability, and transparency (pp. 214-229).

Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I., & Stewart, L. (2016). Singing and social bonding: changes in connectivity and pain threshold as a function of group size. Evolution and Human Behavior, 37(2), 152-158.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science, 316*(5827), 1036–1039.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. Psychological science, 20(1), 1-5.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. science, 330(6004), 686-688.

Zhang, T., Gino, F., & Margolis, J. D. (2018). Does “could” lead to good? On the road to moral insight. Academy of Management Journal, 61(3), 857-895.