Definition and Value Reconstruction of Human Creativity in the AI Era

Authors: Jiang Ke, Ruizhi Huang, Shen Zhoujie, YAN Wenjing, Jiang Ke, Yan Wenjing

Date: 2025-11-07T00:00:00+00:00

Abstract

In the AI era, traditional definitions of creativity can no longer precisely describe the value distinction of intelligent activities between humans and AI. AI' s "high-score" performance in traditional creativity tasks has fostered an illusion that "AI possesses high creativity." In fact, AI' s excellent performance in board games and precision decision-making based on big data does not equate to "high creativity." Creative activities can be divided into three types according to their output outcomes: (1) Interpolation-type creation, which generates knowledge with application value through structured and refined processing within the cognitive system; (2) Extrapolation-type creation, which produces new knowledge beyond existing knowledge boundaries through inductive and generalized reasoning, or represents the process of applying knowledge from existing cognitive domains to new domains; (3) Transition-type creation, which forms higher-level abstract and generalized statement models based on existing knowledge, achieving transformation of fundamental knowledge models. Traditional creativity testing tasks place greater emphasis on interpolation-type and extrapolation-type creation, while relatively neglecting consideration of transition-type creation; moreover, testing tasks do not precisely distinguish among the three creative modes and provide separate evaluations. Among the three types of creative activities, creativity is defined as the ratio between the value of creative output obtained and the cost (energy consumption) paid for it in each creative activity; that is, for a given value of creative output, the lower the cost paid, the higher the creativity of the system. In the three-dimensional model of creativity assessment, creativity is defined as measurement statistical norms on three dimensions: X (interpolation-type creation), Y (extrapolation-type creation), and Z (transition-type creation). According to this model, the creativity levels and values of different types of creative activities, such as engineering manufacturing, research reviews, and theoretical innovation, can be evaluated uniformly; and it provides measurement standards in the sense of measurement statistics

for distinguishing and uniformly evaluating human and AI creativity.

Full Text

Preamble

Redefining Human Creativity and Its Value in the AI Era

Ke Jiang[1,2], Rui-Zhi Huang[1], Zhou-Jie Shen[3], Wen-Jing Yan[3]*

[1]Department of Psychology, Zhejiang Normal University
[2]School of Teacher Education, Lishui University
[3]School of Psychology, Zhejiang Normal University

Abstract

In the AI era, traditional definitions of creativity no longer precisely capture the value distinctions between human and AI intelligence. AI' s "high scores" on conventional creativity tasks have fostered an illusion that "AI possesses high creativity." In reality, AI' s impressive performance in games like Go and big-data decision-making does not equate to "high creativity." Creative activities can be categorized into three types based on their outputs: (1) **interpolation creativity**, which generates valuable knowledge through internal structural refinement within a cognitive system; (2) **extrapolation creativity**, which extends beyond existing knowledge boundaries through inductive and generalization reasoning, or applies knowledge from one domain to another; and (3) **leap creativity**, which forms higher-level abstract and generalized models on the basis of existing knowledge, achieving a transformation of fundamental knowledge frameworks. Traditional creativity tests emphasize interpolation and extrapolation while relatively neglecting leap creativity, and fail to precisely differentiate and separately evaluate these three modes. In this three-dimensional creativity assessment model, creativity is defined as the ratio of creative output value to the cost (energy expenditure) incurred in each type of creative activity. That is, for a given creative output value, the lower the cost, the higher the system' s creativity. The model defines creativity as standardized statistical measures along three dimensions: X (interpolation creativity), Y (extrapolation creativity), and Z (leap creativity). Based on this model, different creative activities—such as engineering manufacturing, literature reviews, and theoretical innovation—can be evaluated within a unified framework. Moreover, it provides psychometric measurement standards for distinguishing and uniformly evaluating human and AI creativity.

Keywords: artificial intelligence; creativity; paradigm revolution; genetic epistemology; Piaget; free energy

Introduction

The two major themes of 21st-century human development are health (World Health Organization, 2000) and creativity (Delors, 1998). The intervention of artificial intelligence has brought unprecedented opportunities for human development, but may also pose risks.

Since the beginning of the 21st century, AI has been entering human life at an exponential rate. Since 2022 in particular, the marketization of large language models has threatened traditional human survival values. Geoffrey Hinton, the "father of artificial intelligence," explicitly stated in his July 2025 speech at the World Artificial Intelligence Conference in Shanghai that AI development may pose the danger of "nurturing a tiger that will eventually harm us"[1]. This potential threat is primarily manifested as a challenge to human creativity. When AI can "create" artworks and efficiently complete work tasks, the "irreplaceability" of being human seems to be gradually eroding.

How should humanity view this growing "little tiger"? How should we evaluate the "creativity" exhibited by AI? As AI advances step by step, should we, as humans, re-examine the meaning and value of human creativity?

2.1 Does AI Actually Have "Creativity"?

In classical creativity theory, researchers typically measure creativity using indicators such as fluency, flexibility, and originality (see Table 1). Today, AI' s performance is challenging these standards. In well-structured tasks (where task structure is highly clear and feedback is immediately verifiable), AI' s outputs and structural recombination strategies are often praised by professionals as "unprecedented" or "astonishing." For example, on March 10, 2016, in Seoul, South Korea, AlphaGo' s "Move 37" in the second game of its match against Lee Sedol broke through conventional expectations of human Go strategy (commentators estimated its prior probability at about "one in ten thousand")[2], significantly altering human players' concepts of layout and influence. The following year, the new generation AlphaZero, trained through "self-play from scratch with only the rules provided," reached and surpassed AlphaGo' s level within dozens of hours (Jiang, 2018), demonstrating that new-generation AI can produce "highly creative" solutions and problem-solving strategies in rule-closed, adjudicable game spaces through reinforcement learning and tree search (Silver et al., 2016; Silver et al., 2018; Silver et al., 2017; Wikipedia, 2025a).

On the other hand, creators and critics from literature, art, and other fields hold a more cautious attitude toward the "creativity" displayed by AI. While they generally acknowledge AI's advantages in diversity, technical features, efficiency, and cost control, they also point out that AI' s outputs remain combinations within existing style spaces, lacking genuine originality. When it comes to value judgment, emotional resonance, and deep aesthetic evaluation, AI' s works appear very immature or even make absurd errors. For instance, although GPT-4 scores high on the Torrance Tests of Creative Thinking (TTCT), its "creative" texts still suffer from semantic repetition and templated generation (Cropley, 2023; Guzik et al., 2023). Similar criticisms are echoed in scientometric research on "disruptiveness" and paradigm shifts (Park et al., 2023; Wu et al., 2019).

It appears that whether AI possesses true "creativity"remains a contentious issue. This paper aims to address the question of "whether AI has genuine creativity" by analyzing AI' s performance in creative tasks, comparing AI and human outputs in such tasks, and exploring the essential characteristics and value of human creativity. Simultaneously, this paper attempts to critique traditional creativity concepts, reflect on traditional definitions and measurement indicators of creativity, and reconstruct a definitional framework and evaluation strategy for creativity in the AI era.

2.2 What Can AI "Create" and What Can It Not?

Contemporary mainstream generative AI is typically constituted by three coupled strategies: (1) **high-probability fitting**, which learns the conditional distribution of outputs through probability sampling in large-scale corpora (e.g., large language models); (2) **reinforcement learning and search training**, which forms optimal response strategies through self-play and Monte Carlo tree search in rule-closed, goal-clear problem spaces (e.g., board games) (AlphaGo/AlphaZero); and (3) **retrieval-generation hybrid (RAG) strategies**, which improve the determinism and coverage of generated results by calling external knowledge bases or vector retrieval techniques (Lewis et al., 2020; Silver et al., 2016; Silver et al., 2018; Silver et al., 2017). These strategies often demonstrate significant advantages in paradigm-determined or well-structured tasks; they can efficiently reorganize existing elements and optimize candidate outputs at low marginal cost while maintaining formal correctness and semantic coherence. However, in complex tasks emphasized by classical creativity theory—such as remote association, problem finding/redefinition, introduction of new concepts, and meaning integration—AI tends to expose limitations like path dependency in generation, weak generalization ability, and low sensitivity to semantic mutations (Lake & Baroni, 2018; Mednick, 1962; Mumford et al., 2012; Ontanon et al., 2022; Wold et al., 2025).

An intuitive comparison is AI' s performance difference between board game tasks and poetry/couplet tasks. In closed-rule, clear-goal, adjudicable game systems like Go or chess, AI can stably output "novel" response strategies verifiable through win rates (such as AlphaGo' s "Move 37"). However, in language games like Chinese couplets and classical poetry creation, although AI can well follow formal constraints such as tonal patterns, parallelism, rhyming, and character arrangement, and quickly generate multiple candidate solutions, the generated content universally suffers from the problem of "easy formal compliance, difficult semantic coherence" (Yi, Li, & Sun, 2018; Yan, Li, Hu, & Zhang, 2016; Chen, & Cao, 2024; Jiang, & Zhou, 2008). More similar studies also point out that AI in such tasks often exhibits semantic far-fetchedness, loose intention, misuse of allusions, and context mismatch (He et al., 2012; Yan et al., 2016; Liu et al., 2022).

Similar problems frequently appear in visual image generation tasks. For example, AI can generate images with clear composition and coordinated colors, but frequently makes errors that violate common sense, such as "six-fingered hands" [3]. Therefore, AI models do not lack image processing capability, but rather the ability to integrate meaning. AI can reconstruct shapes and imitate brushstrokes, but cannot understand the "meaning consistency" between finger number and function because it has never possessed or experienced a "hand." This lack of experience determines AI' s disembodied characteristic. AI lacks a body, senses, and action experience in the existential sense, and thus cannot connect symbolic structures with life meaning. In human cognition, meaning generation depends not only on language and logic, but is deeply rooted in bodily experience and situational interaction. Without this embodied experience, even if AI can simulate language and visual rules, it is difficult to truly "understand" the life logic and experiential meaning behind their combinations (Bender, & Koller, 2020; Bisk, Holtzman, Thomason, et al. 2020).

Therefore, whether it' s semantic defocusing in language generation or common sense mismatch in image generation, these AI performances reflect the same dilemma: AI cannot organize local forms into meanings obtained by humans through embodied experience.

In summary, under the classical creativity definition framework, AI models can usually reliably satisfy indicators such as "fluency, flexibility, originality, and elaboration," but show clear deficiencies in more complex creative dimensions, especially in remote association, problem reconstruction, meaning integration, and value judgment (Koestler, 1964; Mednick, 1962; Mumford et al., 2012).

2.3 Why Can AI Score High on Classical Creativity Measurement Tasks?

The classical definition of creativity is the ability of a subject to produce ideas, works, or solutions that are both novel and useful (Runco & Jaeger, 2012). Based on this framework, classical creativity assessment tools, particularly the Torrance Tests of Creative Thinking (TTCT), evaluate performance across four dimensions: fluency, flexibility, originality, and elaboration (Torrance, 1966/1974; Warne et al., 2022). Over the past decade, the five most commonly used creativity measurement standards in global psychology journals, with their core indicators, can be considered the most consensual "benchmark scales" today (Carson, Peterson, & Higgins, 2005; Diaz, Nelson, Beaujean, Green, & Scullin, 2024; Dollinger, & Shafran, 2005; Gough, 1979; Hadas, & Hershkovitz, 2025; Hongdizi, Cui, & Zhou, 2023; Kapoor, Zheng, Reiter-Palmon, & Kaufman, 2023; Ogurlu, Acar, & Ozbey, 2023; Runco, & Acar, 2012; Zhang, Zhong, Zhang, Ren, & Zhang, 2023).

Table 1 Creativity Measurement Tools and Evaluation Indicators

Category	Representative Tools	Typical Scoring Indicators	Representative Literature
Divergent Thinking	Alternative Uses Task (AUT); TTCT Verbal/Figural	Fluency (output count), Originality (rarity), Flexibility (category count), Elaboration (detail level)	Ogurlu & Runco, 2023; Hadas et al., 2025; Hongdizi et al., 2023
Convergent Thinking	Remote Associates Test (RAT) & variants (Compound RAT, etc.)	Correct answer count, reaction time, efficiency in finding cross-domain associations	Wu et al., 2020; Diaz et al., 2024
Creative Achievement	Creative Achievement Questionnaire (CAQ)	Domain cumulative achievement scores (arts, science, writing, etc.)	Carson et al., 2005; Zhang et al., 2025
Creative Product	Consensual Assessment Technique (CAT); Creative Product Semantic Scale (CPSS)	Expert ratings on Novelty × Usefulness × Aesthetics/Surprise, etc.	Dollinger & Shafran, 2003
Self-Report/Traits	Kaufman Domains of Creativity Scale (K-DOCS); Creative Personality Scale (CPS)	Multi-domain creative behavior frequency; creative personality adjective scores	Zhang et al., 2023; Kapoor et al., 2021

Recent studies have found that generative AI, represented by large language models (LLMs), often scores significantly higher than human averages on standardized tasks like the TTCT, and even outperforms highly creative human groups on some dimensions (Cropley, 2023; Guzik et al., 2023). This result does not arise because AI possesses truly high creative ability, but mainly stems from its structural advantages on assessment dimensions. For example, in fluency, AI can output large numbers of candidate solutions in extremely short time through massive parallel generation mechanisms, easily satisfying the quantitative standard of "number of ideas." In this task, humans are limited by working memory capacity and information retrieval speed, making it difficult even for highly creative individuals to match AI model performance. In flexibility, AI's training corpus covers numerous knowledge domains, enabling AI models to rapidly switch between different semantic clusters, showing advantages in cross-category generation. In originality (or rarity) scoring, due to high-speed information retrieval capabilities and massive training sets, AI models can often generate combinations rare in human samples within limited time, thus obtaining high scores on this dimension. In elaboration, AI model algorithms are completely rational reasoning; this determines they almost never make logical errors and are not disturbed by irrational factors (such as emotions) in generalization reasoning, thus generating logically rigorous, structurally clear, and rationally sound outputs.

In divergent thinking or "remote association" tasks, generative AI similarly relies on its advantages in information processing capacity and retrieval efficiency to obtain higher scores than humans. When humans perform such tasks, they are limited by working memory capacity. As is well known, human working memory' s instantaneous retention capacity is about 7$±$2 information units (Miller, 1956). This means that when facing the need to generate large numbers, cross-domain, and structurally diverse novel results, human individuals often struggle to maintain simultaneous activation and transformation of multiple parallel ideas. AI, based on nearly unlimited data pools and computing power, can perform large-scale information retrieval and comparison in extremely short time, not only discovering numerous information combinations difficult for humans to find or already forgotten, but also completely preserving records of all generation paths and semantic mappings ever executed. In summary, AI' s generation behavior is essentially an efficient coverage and compressed expression of large-scale combination spaces, and well matches the scoring orientation of creativity assessments for tasks like "diversity," "flexibility," "originality," "elaboration," and "remote association." Therefore, AI does not obtain "high creativity scores" by effectively simulating human creative minds, but by technically satisfying the formal requirements of creativity assessment indicators under completely different information processing conditions.

Multiple assessment results show that GPT-4' s scores on TTCT tasks reach the top 1% of human samples in originality and fluency dimensions, with flexibility also in the 93rd-99th percentile range (Colton et al., 2012; Guzik et al., 2023). This performance is "near perfect" in the scale scoring system. However, comparative studies reveal a noteworthy trend: in the past 18-24 months, mainstream large language models' scores on tasks like DAT (Divergent Association Task) and AUT (Alternative Uses Task) have not continued to rise, and have even stagnated (ScienceDirect, 2023; UMWestern, 2023; Technology Networks, 2023; Tykierko et al., 2025). This phenomenon suggests that current AI models' reliance on training corpora and combination strategies may have approached the boundary limits of their expressive space. We can call this boundary of expressive space limited by corpora the AI model' s creativity bottleneck. In other words, although models like GPT-4 can obtain high creativity scores within existing scoring frameworks, these scores do not reflect that AI models possess true creative expansion or cognitive breakthrough capabilities.

More critically, creativity measurement tools themselves have structural biases. A typical case is the widely discussed "fluency-originality coupling problem." Specifically, because scoring rules often use rarity as the main indicator of originality, the more results participants generate (i.e., the higher the fluency score), the higher the probability of rare answers appearing, thus systematically increasing originality scores. This phenomenon is particularly significant in TTCT Figural tasks. Research shows the correlation coefficient between the two scores reaches r 0.79-0.86 (typical value 0.844) (Kim, 2004; Kim, 2006; Kim et al., 2006). Studies on reliability and equivalence found that when excluding subtests with possible overlapping responses, their correlation significantly decreased; this indicates that high test scores only reflect part of subjects' creative processing ability, while another part of the score may be statistical bias caused by task structure itself (Acar et al., 2023; Forthmann et al., 2020; Said-Metwaly et al., 2018). For AI models with big data support and high computing power, this structural bias-induced scoring advantage is even more significant.

This means that traditional creativity test scales amplify the importance of fluency and flexibility, and approximate originality through a low-threshold rarity indicator, thus lacking sufficient consideration and effective differentiation for creativity dimensions like remote association strength, problem reconstruction, concept introduction, and theoretical integration. In this evaluation structure, AI models can naturally achieve high scores on the "fluency-flexibility-uniqueness combination" tasks through advantages in massive training corpora and computing power, thus being endowed with "high creativity" evaluation. However, this score performance is not equivalent to originality in the strong sense; especially when tasks involve value judgment and deep meaning integration, problem space reconstruction, and paradigm leaps across semantic scenarios, the creative ability demonstrated by AI models significantly decreases (Koestler, 1964; Mednick, 1962; Mumford et al., 2012).

2.4 Re-examining the Meaning and Value of Human Creativity

AI models' high scores on creativity tests give us reason to deconstruct classical creativity assessment schemes and reconstruct the meaning of human creativity. Using fluency, flexibility, originality, and elaboration as core evaluation indicators, this assessment scheme inevitably systematically overestimates AI models' creativity and correspondingly ignores humans' unique advantages in high-level originality and paradigm leap capabilities. The structural difficulties deep learning systems exhibit when handling out-of-distribution (OOD) tasks are a powerful challenge to this evaluation orientation (Lake & Baroni, 2018; Liu et al., 2021). Due to biased evaluations of AI models' creativity, two worrisome public opinion tendencies have emerged: on one hand, some people, amazed by AI models' generative capabilities, have begun to doubt the level and value of human creativity; on the other hand, others, overemphasizing AI' s difficulties in OOD tasks, make misjudgments or denials of its potential auxiliary creative functions.

Focusing on AI' s creativity is essentially to more clearly understand the characteristics and value of human creativity. Deconstructing classical creativity assessment systems through AI' s creativity performance is fundamentally aimed at reconstructing the definition and evaluation strategy of human creativity in the context of the AI era. Starting from this point, we can further reflect on what the meaning and value of human existence and development are when facing AI' s step-by-step approach; how humans should rationally view AI; and how to establish cooperative and symbiotic relationships with AI.

This reconstruction must be based on the fundamental differences between humans and AI at the ontological and epistemological levels. Human creative activities are adaptive expressions of groups or individuals in natural selection survival competition; they are embodied; they aim at meaning generation and problem solving; their cognitive processes are limited by humans' limited computing power and incomplete experience accumulation, representing "small computing power, small data pool" cognitive expression. Correspondingly, AI is a typical disembodied cognitive system whose generative ability depends on preset program architecture, massive datasets, and high-density computing power support; it is essentially a system optimization process based on symbol processing and pattern mapping, rather than meaning generation and construction. Simply put, human creativity fundamentally serves the subject' s survival and species continuation; while AI' s operation does not have purposefulness associated with its own existence needs.

Therefore, we advocate redefining creativity within an evolutionary dynamics framework centered on survival and development, understanding it as an efficacy characteristic oriented toward "effect-to-cost ratio." This definitional scheme for creativity will provide a theoretical foundation for constructing creativity measurement models applicable to human cognitive systems, and further provide a quantifiable, computable representation system for creativity. It can also incorporate human creativity and AI' s creative outputs into a unified epistemological framework for comparable evaluation.

3 Formal Analysis of Creative Activities

Before discussing the meaning of creativity, we must first distinguish between creative activities and the meaning and value of creativity. Creative activities describe a type of organism activity that can produce novel and effective outcomes; creativity is an evaluation of an individual' s ability to engage in creative activities. As previously mentioned, when comparing AI and humans, people often ignore the conceptual distinction between creative activities and creativity. That is, we may see the stunning performance of AI-created works and take the works themselves as the evaluation standard for "creativity" ; we may also ignore that AI and humans may each have their own strengths in different types of creative activities.

Based on the phenomenological expression of creative outcomes, we categorize creative activities into three forms: **interpolation**, **extrapolation**, and **leap**; corresponding respectively to the recombination of existing concepts, expansion of knowledge boundaries, and hierarchical enhancement and generation of cognitive schemas.

3.1 Interpolation: Refinement of Knowledge Within a System

Interpolation refers to the structural refinement and systematic enhancement of knowledge within existing knowledge structures through deductive reasoning. This type of creative activity is characterized by logical consistency and systematic completeness. Although it does not break through the boundaries of the original problem framework, it achieves high refinement and organizational reconstruction within the framework. Euclidean geometry is a typical representative of this type of creative activity. Starting from limited axioms, Euclidean geometry constructs a broad and rigorously structured geometric system through deductive reasoning systems.

This type of creative activity is "system-internal optimizing creativity," that is, under the premise of stable existing cognitive schemas, it enhances the organization degree and internal consistency of knowledge systems through intra-system element recombination and logical deduction (Lenk, 2009).

The psychological representation of interpolation creativity usually manifests as element recombination supported by high fluency and flexibility. In the classic definition framework of "novelty × effectiveness," interpolation creativity meets the compliance and innovation requirements of evaluation standards (Runco & Jaeger, 2012; Torrance, 1966/1974; Warne et al., 2022).

The logical meaning of "interpolation" can be expressed as the "interpolation theorem,"that is, if a proposition A C can hold, then in some systems there exists an intermediate proposition B such that A B and B C are both valid inferences, and B contains only symbols common to A and C. The "interpolation theorem" is the logical foundation for automatic proof and program verifiability within formal systems. In modern computer technology, this "interpolation deductive reasoning" is widely applied in formal verification, theorem provers, and logic programming languages (like Prolog), with core operations built on enumerable derivation and optimal selection mechanisms within rule spaces. The "creative solutions"AI demonstrates in well-structured tasks (like board games, formalized decision-making, and logic puzzles) can also be understood as a complex form of interpolation process, where models execute search and recombination within rule spaces to generate formally compliant but combinatorially novel solution strategies.

In information theory terms, interpolation creativity can be understood as achieving effective encoding of facts within a larger range through shorter description lengths. Therefore, interpolation creativity not only manifests as optimization and recombination of system knowledge according to rules, but can also achieve regulation of rule systems. For example, in *Foundations of Geome-*

try (1899), Hilbert reconstructed Euclidean geometry with about 20 axioms and introduced formal discussions on axiom independence and system consistency. This example demonstrates how "interpolation creativity" can expand from theorem generation to reflection on and clarification of rule systems themselves (Blanchette, 2024; Hilbert, 1899; Wikipedia, 2025b; Zach, 2023).

Generative AI often demonstrates extremely high creativity in "interpolation creativity" tasks. For example, the award-winning work *Théâtre d' opéra Spatial* (commonly known as "Space Opera House") at the 2022 Colorado State Fair' s digital art competition. This work conducted coordinated and stylistically unified fusion around image vocabulary related to "space" and "opera house," presenting "exquisite image craftsmanship" and "highly compatible aesthetic style" (Smithsonian, 2022; VICE, 2022; Hyperallergic, 2022).

3.2 Extrapolation: Expanding Knowledge Boundaries

Extrapolation refers to the process of applying existing structures to new domains, typically relying on analogy and inductive reasoning. The essence of extrapolation creativity is expanding existing cognitive boundaries, thereby increasing the total information volume of the knowledge system and reducing the uncertainty of the cognitive system.

From the perspective of organism survival and adaptation, extrapolation creativity enables organisms to break through original cognitive system boundaries, reduce system uncertainty by introducing new elements (conditions, concepts, and other variables), and achieve broader adaptability. This ability to expand knowledge makes extrapolation the "incremental" engine in knowledge system evolution.

Compared with interpolation representing structural optimization within knowledge systems, extrapolation represents functional adaptation of bio-cognitive systems to new environments. Extrapolation creativity mainly relies on inductive reasoning. However, the philosophical meaning and epistemological foundation of inductive reasoning itself are controversial. The famous "Hume' s questions" made later philosophers realize that inductive reasoning cannot obtain logical necessity from premises to conclusions (Hume, 1748/2000). Although philosophers after Hume all tried to argue, they could never prove the logical legitimacy of inductive reasoning (Russell, 1948/2004). However, in actual knowledge construction and behavioral decision-making, humans and other organisms universally show dependence on inductive reasoning. Experimental studies show that even small insects with limited cognitive abilities, such as fruit flies, can form experience-based inductive extrapolation in environments for path selection and resource anticipation (Guo, 1990). This indicates that inductive extrapolation is not unique to complex intelligent agents, but a universal adaptive strategy possessed by biological individuals.

From an epistemological perspective, inductive reasoning is not only an important cognitive process for generating new concepts and forming categorical judgments, but also a key channel for abstracting more general knowledge from partial experience. It often appears in the form of analogical reasoning, manifesting as migration from old knowledge domains to new knowledge domains, and is an important method in human cognitive systems for "re-encoding problem spaces" (Gentner, 1983). Therefore, extrapolation ability is the instinct of living organisms to adapt to environments, predict risks, and optimize resource allocation. Organisms construct generalized response tendencies through limited experience to make optimal decisions under uncertainty; this is an information compression and entropy reduction process of "foreseeing the whole from parts."

Currently, AI' s ability improvement in extrapolation tasks (like generalization and reasoning) mainly relies on algorithm architecture iteration and optimization. For example, enhancing context modeling capabilities by expanding model scale (such as Transformer depth/width evolution), using chain-of-thought prompting and code interpreter techniques to simulate multi-step Bayesian reasoning, and improving strategy migration efficiency through reinforcement learning and memory mechanisms. However, these advances essentially belong to probability theory computational efficiency improvements and do not touch upon transformation of the "problem space" itself. Moreover, this extrapolation ability improvement often comes at the cost of extremely high energy consumption. Large language model training processes often consume hundreds of gigawatt-hours of electricity resources. For example, GPT-3' s training process is estimated to consume about 1,287 MWh (Patterson et al., 2024). Some cutting-edge institutions have begun building gigawatt-level computing and energy infrastructure to support AI training and inference tasks (Gao et al., 2025; Reuters, 2025; Tom' s Hardware, 2025). The self-play learning of reinforcement learning systems represented by AlphaGo/AlphaZero similarly accompanies huge computational overhead and long-term operation requirements (Silver et al., 2017). In contrast, the human brain achieves efficient cognitive processing with extremely low energy consumption. The power required for nervous system operation is only about 10-20 W, and the energy utilization efficiency in gray matter activity has achieved high optimization through ion channel dynamics and glucose metabolic pathways in the evolutionary process (David Attwell & Simon B. Laughlin, 2001; Raichle & Gusnard, 2002). Brain networks far outperform existing general computing architectures in reasoning and abstract value generated per unit of energy (Sengupta, 2013; Sengupta, 2014).

This comparison shows that AI mainly relies on computing power-driven "brute force enumeration + probability optimization" strategies for extrapolative knowledge creation, lacking the energy efficiency trade-offs of biological intelligence agents.

3.3 Leap: The Power Operation of Cognitive Schema Elevation

Interpolation and extrapolation constitute recombination and generalization of cognitive activities in a two-dimensional problem space. **Leap** enables creative activities to break through the two-dimensional plane, achieving hierarchical elevation and self-reflective reconstruction of three-dimensional cognitive structures. The so-called "leap"refers to the cognitive enhancement from the experiential level to the relational level, then to higher structural levels, ultimately forming meta-structure reconstruction of original cognitive structures. For example, Newton abstracted concepts like "motion" and "mass" from the phenomenon of "apple falling," and further reconstructed a unified physical framework for celestial and object motion, forming unified mathematical expressions that can describe motion from apples on the ground to celestial bodies. This series of cognitive leaps is not simple data generalization, but meta-structure reconstruction of problem spaces, explanatory structures, and evaluation indicators.

Among the three creative activities, leap creativity has the highest characteristics of compressibility, theoretical unity, and structural reconstruction. Therefore, "leap creativity" holds a particularly important position in cognitive evolution processes (Hanson, 1958; Kuhn, 1962; Lakatos, 1978).

3.3.1 From "Phenomenon-Relation" to "Structure-Meta-Structure"

The higher form of creativity manifests as the extraction and construction of cognitive structures and meta-structures (Piaget, 1950, 1970). Piaget' s genetic epistemology describes the cognitive development process of "equilibrium-disequilibrium-re-equilibrium."This process reflects the "phenomenon → relation → structure → meta-structure" cognitive level elevation, generating logical relations from perception-motion coordination, generating cognitive structures and operational rules from logical relation coordination, and ultimately completing formal construction of meta-rule systems, i.e., generating "cognitive schemas."

Current large language models (LLMs) have demonstrated high combinatorial ability at the "phenomenon-relation" level (such as pattern recognition and semantic matching), which is essentially recombination and fitting of explicit patterns. However, they have not yet possessed the ability to generate deep structures from heterogeneous phenomena. In contrast, human cognitive creation often occurs at the structural abstraction level. For example, Newton' s unification of terrestrial and celestial motion through "universal gravitation," and Darwin' s integration of species variation through "natural selection," both reflect the process of constructing unifying meta-theories through cognitive identification reaching structural abstraction.

Furthermore, structural elevation is not merely local adjustment of the "relation level," but lies in constructing new "structure levels" to reshape problem spaces. This transformation goes beyond identifying or combining relationships between existing inputs and outputs. That is, the key to leap creativity is not discovering more correlations, but proposing a structural framework that can subsume or replace old relational systems, enabling originally dispersed or incompatible information to be integrated into new explanatory systems.

Further elevation based on the "structure level" manifests as construction and recombination of "structures of structures" (meta-structures). This process is no longer limited to rule systems within specific domains, but introduces meta-rules and formal grammars that can subsume multiple cognitive domains and multi-level representations, thereby achieving cross-domain unified representation and problem transformation. For example, calculus abstracts continuous change into operable limit processes; group theory formalizes symmetry into transferable operational structures; information theory measures uncertainty and information volume changes with entropy at its core. Such meta-structure level abstractions not only enhance the compressibility and generality of knowledge organization, but also provide logical and rule foundations for the development of new theoretical frameworks.

3.3.2 Classic Example: The "Unity-Compressibility-Testability" of the Law of Universal Gravitation One of the most representative classic cases of leap creativity is Newton' s construction of a unified theoretical framework for terrestrial objects and celestial motion laws. Phenomena like apple falling and moon orbiting the Earth belong to different types of motion events at the perceptual experience level; the former is daily near-Earth observation, the latter belongs to astronomical phenomena; there is no direct association between the two phenomena. Newton' s core insight was placing these two types of phenomena into the unified framework of the law of universal gravitation and the three laws of motion (Newton, 1687/1999). This leap reflects three aspects of elevation in cognitive structure:

1. **Unity**: Through the "law of universal gravitation" representing the common mechanism of terrestrial and celestial motion, it achieved mechanical integration of "all things connected as one," breaking the traditional cosmological dichotomy between celestial and terrestrial realms (Janiak, 2008).
2. **Compressibility**: Through a few basic laws (formulas like $\mathrm{F} = \mathrm{ma}$, $\mathrm{F} = \mathrm{Gm_1m_2/r^2}$), it can derive quantitative predictions for numerous complex phenomena, including tidal changes, comet trajectories, planetary motion, etc. (Cohen, 1985). This embodies the principle of parsimony of "covering broader phenomenon domains with shorter description lengths" in information theory (Li & Vitányi, 2008).
3. **Testability**: The theory not only explains known phenomena but also generates quantifiable new predictions, and allows measurement of deviations and model corrections (such as combining motion theory with observed anomalies in Uranus' s orbit to predict Neptune' s existence), thus constituting an "adaptive feedback loop" in evolutionary terms (Lakatos, 1978).

From the perspective of cognitive creation, this process includes the elevation from "phenomenon-relation" observation patterns to "structure" unified representation, and further elevation to "meta-structure" power operations. This power operation manifests as Kuhnian paradigm shifts at the collective knowledge system level. Paradigm shift is the update of the "worldview" adopted by scientific communities (Kuhn & Hacking, 1970). Lakatos (1978) also pointed out that scientific theory progress is not incremental accumulation, but evolution of "research programs," whose hallmark is new theories surpassing old paradigms in unity and predictive power (Lakatos, 1978).

Furthermore, Newton's theory also provided a paradigmatic template for "structural compression" for subsequent scientific development. For example, Maxwell's equations unified electricity and magnetism into electromagnetic theory; Einstein's general relativity replaced Newton's gravitational and motion models with geometrized forms; both can essentially be seen as higher-order "structure → meta-structure" leaps (Piaget, 1950, 1970). Classic creations in scientific history are precisely the presentation of such leaps at the collective knowledge level (Kuhn, 2004).

3.3.3 Embodied Experience: The Difference Between Humans and AI
Human knowledge does not originate from isolated abstract computation, but is rooted in triple embedding of embodied experience, specific contexts, and social practice. First, the body is not only the executor of perception and action, but also forms initial cognitive frameworks through movement coordination, spatial navigation, and body boundary perception (Gallagher, 2005; Thelen & Smith, 1994). Second, cognitive activities always occur in specific contexts, constrained by task goals, tool availability, and environmental constraints, i.e., so-called "situatedness"(Suchman, 1987). Third, social norms, language systems, and cultural practices further constitute the interactive context of knowledge formation, making knowledge not only functionally effective but also possess group-shared understandability and public evaluation criteria (Tomasello, 1999; Vygotsky, 1978).

Within this series of nested embeddings, abstract structures of individual knowledge are not top-down endowed, but gradually constructed through multi-channel coordination and feedback during embodied operations and social interactions. This process has historical and structural characteristics, i.e., experience accumulation is not linear growth, but achieves leap from operational structures to abstract formats through representation recombination and cross-layer integration (Barsalou, 2008; Clark, 1997). With support from embodied-contextual-social embedding, human cognition demonstrates highly hierarchical, parallel, and cross-channel coordination capabilities (Barsalou, 2008). One of its core features is that the mind can jump and align between multiple "cognitive windows" like an operating system (Baars, 1993; Nelson, 1990). These cognitive windows may include: perception-action layer (such as audio-visual feedback and spatial operation), relation-symbol layer (such as understanding metaphor, identity, and power relations), and mathematical-logic layer (such as abstract reasoning and model construction). In complex tasks, humans can rapidly switch between different windows and complete integration of multi-level information through metacognitive regulation mechanisms (Clark, 2015).

Neuroscience research conclusions support this multi-layer coordination model. The cerebral cortex presents a progressive hierarchical structure from primary sensorimotor areas, to limbic systems (such as amygdala, cingulate gyrus and other emotion processing areas), to abstract planning and control systems supported by prefrontal cortex (Fuster, 2000; Mesulam, 1998). These three layers roughly correspond to peripheral-surrounding-central information processing levels: the central system forms goals and plans, the surrounding system integrates cross-modal information, and the peripheral system implements strategies through body-environment interaction and feeds back to adjust plans. Cognitive leaps unfold within this hierarchical nested system; coordination between different levels enables knowledge to not only form structures, but also achieve power operations from structure to "meta-structure."

Although generative AI systems (like large language models) have shown remarkable success in simulating human surface language structures and identifying statistical patterns, and mainstream AI' s probabilistic modeling algorithms can perform conditional fitting and sampling optimization in multi-layer parameter spaces, with deep neural networks possessing multi-level transformation algorithm structures that seemingly can approximate power operations of arbitrarily complex functions, this "algorithmic hierarchy" does not correspond to the "meaning hierarchy" in human embodied cognition, but is linear computational mapping chain superposition and optimization (Bengio et al., 2021; Chollet, 2019).

More fundamentally, AI' s "knowledge" mainly comes from statistical integration of massive corpora, rather than multi-layer meaning nested structures formed through embodied action, social practice, and contextual reflection. Due to lack of perception-motion coordination and emotion-intention regulation, AI cannot perform contextual modulation, norm internalization, or value alignment on its generated results, thus performing poorly in cross-level self-reflective control required for cognitive leaps (Bisk et al., 2020; Lake et al., 2017).

Additionally, meta-structure level creation requires systems to possess reflective adjustment capabilities for structures, i.e., understanding, modifying, and reconstructing generation rules. Current AI systems have not yet demonstrated meta-control over their own representation mechanisms; they can simulate rule results but cannot propose, evaluate, or reconstruct rules themselves. This deficiency makes it difficult for AI to complete cognitive leaps from "structure" to "meta-structure," i.e., it lacks cognitive paths toward theoretical elevation and paradigm transformation. Therefore, although AI performs excellently in paradigm-internal optimization and combinatorial innovation, its creativity in hierarchical leaps remains constrained by innate limitations of computational architecture.

3.3.4 Individual and Group Cognitive Leaps Leap creativity not only manifests as knowledge content updates, but more importantly involves cognitive architecture reorganization and paradigm changes. This process can be understood as a cognitive system jumping from one stable structure to a higher-power paradigm structure, thereby generating new organizational logic and meaning formats (schemas). Leaps are not continuous deductive or inductive reasoning, but nonlinear structural mutations. They are usually accompanied by intensified cognitive conflicts, deconstruction of existing structures, and reconstruction of higher order, ultimately leading to the formation of a new stable structure. This mechanism manifests as a dynamic sequence of "conflict-disorganization-reconstruction-re-stabilization."

Piaget' s cognitive development theory describes that individual cognitive structures continuously experience cycles of "equilibrium-disequilibrium-new equilibrium" during development. This dynamic process is driven by two core operations: **assimilation**, which integrates new information into existing schemas; and **accommodation**, which adjusts original schemas to restore structural consistency when facing heterogeneous information. When the intensity of information input or cognitive conflict exceeds the system' s regulation threshold, the original cognitive schema faces disintegration, thereby triggering systematic leaps. A new, higher-order cognitive organization level is generated, and the cognitive system completes structural reconstruction and paradigm updating.

Kuhn (1962) pointed out that scientific paradigms have several key characteristics: first, **framework nature**—paradigms define standards for what constitutes "valid questions" and "reasonable answers," determining the cognitive boundaries of research activities; second, **exclusivity**—old and new paradigms are incommensurable, and paradigm shifts often accompany reconstruction of entire worldviews; third, **nonlinear jump nature**—paradigm changes are not continuous evolution, but emergent occurrences along the chain of "anomaly accumulation-crisis outbreak-structural leap." This model structurally echoes Piaget' s "equilibrium-disequilibrium-re-equilibrium" process, but Kuhn extended this cognitive occurrence process from the individual level to the evolution dimension of scientific community knowledge structures, emphasizing the group-level cognitive rupture and worldview reshaping carried by paradigm shifts.

3.4 Mathematical Expression of the Three Creative Forms

To more precisely characterize the three manifestations of creative activities, we further employ information theory, thermodynamics, and the free energy principle to quantitatively model the three forms of creativity.

3.4.1 Interpolation Creativity is System Entropy Reduction Interpolation creativity usually does not involve expansion of knowledge boundaries, but seeks optimal paths, more refined structures, or more efficient organization methods within original structures. The core of this creativity lies in enhancing internal system order, and thus can be characterized as a process of system "entropy reduction." From thermodynamic principles, internal system order reorganization and structural optimization represent a transition from high-entropy states to low-entropy states. Internal system entropy reduction does not change system boundary conditions, but enhances system operational efficiency and explanatory power through enhanced internal coupling and refined structural processing.

First, **thermodynamic entropy** (S) is used to measure a system' s degree of disorder at the microscopic level or the breadth of its state space distribution. Entropy reduction means the system' s states tend to concentrate, i.e., its structure becomes more ordered and constraints increase. For example, when a physical system receives energy input and undergoes internal reorganization, the system' s entropy value decreases, meaning the system experiences a transformation from chaos to order. This mechanism can be analogized to "interpolation creativity" in knowledge structures, i.e., reorganizing and organizing elements within the system without expanding boundaries, thereby enhancing overall order.

Second, from an information theory perspective, the definition of **information entropy** (Shannon Entropy, H) is (Formula 1):

$$H(X) = -\sum_{x} p(x) \log p(x)$$

This formula measures the uncertainty degree of discrete random variable X; it can also characterize the distribution breadth of information in symbolic systems. When a system undergoes processing and its output probability distribution tends to concentrate, i.e., the certainty of certain symbol combinations increases (p(x) increases), information entropy (H) decreases. Thus, rule refinement, category compression, and expression constraints in interpolation creativity are precisely "entropy reduction" processes in information theory terms. In other words, disorder compression achieved through deduction and integration in logical structures manifests as decreased uncertainty of representation systems. Georgiou examined dynamic change characteristics of information entropy in cognitive processing (Georgiou, 2005). He found that in deductive reasoning operations, cognitive systems continuously narrow possibility spaces to move toward more deterministic structures, manifesting as information entropy reduction. Organic systems need to continuously introduce external energy or information to maintain or enhance their internal structural order. Correspondingly, creative cognitive activities can be seen as an active entropy reduction process, aiming to transform uncertain knowledge states into clearer, more refined, and more structured expression systems.

Overall, generating new structures by enhancing internal system order without expanding problem spaces constitutes the core characteristic of "interpolation creativity." Therefore, we can model it as an entropy reduction process of a knowledge system.

Let the system reach final state S_1 from initial state S_0 through interpolation reasoning and structural optimization, with corresponding entropy values S_0 and S_1. Then entropy change is (Formula 2):

$$\Delta S = S_1 - S_0$$

Accordingly, the "interpolation creativity value" can be expressed as (Formula 3):

$$C_{\text{interpolation}} = -\Delta S$$

That is, the amount of entropy reduction a system achieves through structural reorganization is equivalent to the "interpolation creativity value."This definition characterizes the refinement process of knowledge systems from "disorder" to "order."

3.4.2 Extrapolation Creativity Generates Information Gain and Uncertainty Control Extrapolation creativity is when individuals, when dealing with unknown environments or new tasks, compress the uncertainty space of knowledge systems and improve decision accuracy by constructing hypothetical models, analogical migration to new domains, and other means. This process reflects dual goals of system information gain and environmental uncertainty reduction, i.e., it both expands the application scope of original knowledge systems and enhances the cognitive system' s adaptability and robustness to atypical inputs.

From an information theory perspective, extrapolation can be understood as extracting generalizable low-dimensional structures from high-entropy states in input spaces, thereby achieving "prior commitment"to potential models, and continuously correcting and updating them through verification processing. This makes extrapolation creativity not only a process of rational cognitive resource scheduling, but also an "order output" oriented toward the unknown (Friston, 2010; Jaynes, 2003).

Delisle (2009) proposed a "selective information sampling" model, providing a biological analogy model for extrapolation creativity from a cognitive constructivist perspective. According to this model, when biological intelligent agents face complex and dynamic environments, they do not adopt random information intake strategies, but actively sample information segments that have predictive efficacy for future actions, characterized by goal-directedness and structural sensitivity. This sampling mechanism shows preference and exploration tendency for potential structural patterns in the environment, enabling agents to extract key patterns with extrapolation potential under highly uncertain input conditions. By continuously strengthening perception of "local integrability" in uncertain regions, individuals gradually form structural understanding and strategy presets for unfamiliar situations.

Therefore, extrapolation creativity can be seen as an optimization activity based on Expected Information Gain (EIG). In the Bayesian reasoning and active inference framework, the optimization activities taken by intelligent agents aim to maximize future information gain and actively reduce environmental uncertainty. Its mathematical definition is as follows (Formula 4):

$$\mathrm{EIG}(a) = \mathbb{E}_x[H(p(y|x)) - H(p(y|x,a))]$$

where x represents observed variables, a represents possible action strategies, and y is the system' s true state. Extrapolation creativity is thus manifested as a selective exploration process with the goal of maximizing EIG, possessing high structural sensitivity and prediction orientation[4].

[4] **Explanation**: Taking a medical examination as an example: How much "uncertainty" can an action a (like doing a check/running an experiment) on average reduce about what we want to know (y)?

We use ΔI to represent system information gain (EIG); use H_{prior} (prior entropy) to represent the agent' s uncertainty about environmental states before creative activity (H(p(y|x))); H_{posterior} (posterior entropy) represents the remaining uncertainty after extrapolation operation (H(p(y|x,a))). Thus: ΔH = H_{posterior} - H_{prior} < 0. Obviously, this is an entropy reduction process. System entropy reduction means information increase, i.e., ΔI = -ΔH. Note that the entropy reduction meaning generated by extrapolation creativity differs from closed-system order improvement in interpolation creativity; extrapolation creativity' s entropy reduction refers to uncertainty reduction in "unknown domains" and overall information theory gain of the system, i.e., ΔI > 0.

Thus, we define extrapolation creativity value as system Information Gain (IG) (Formula 5):

$$C_{\text{extrapolation}} = \Delta I$$

For example, migrating the strategy of "crossfire" in military combat to enhance strike effectiveness to tumor radiotherapy, inventing the "Gamma Knife" technique for stereotactic radiotherapy, is a typical extrapolation creativity. The essence of this process lies in data migration from known domains to target modeling in another domain. This migration significantly reduces uncertainty in the target system, and the resulting Information Gain is thus regarded as an indicator for quantifying such extrapolation creativity.

3.4.3 Leap Creativity is Free Energy Reduction in System-Level Transformation The cognitive structure leaps embodied in human creative activities often exhibit characteristics of nonlinear mutation, critical triggering, and probabilistic activation. This process is difficult to characterize through traditional continuous function modeling methods. To better describe its dynamic mechanisms, we introduce the physics electron energy level transition model as an analogical framework to formally characterize cross-level transition processes of cognitive structures under specific conditions.

In the Bohr model, the necessary condition for electrons to be excited from low energy level n to high energy level n' is absorbing photon energy and satisfying h ΔE = E_n' - E_n. If excited through particle collision, incident particles must carry at least threshold energy ΔE. Whether transition occurs is determined by transition probabilities in quantum mechanics (such as Einstein coefficients or Fermi' s golden rule), thus presenting statistical characteristics.

Analogizing to cognitive systems, we can formalize the conditions for cognitive leaps.

First, let: - E_{avail}: Cognitive resources available to the subject, such as attention, motivation, knowledge schema invocation ability, etc. - E_{th}: Existing resource level of the subject' s current cognitive level. - ΔE: The abstract span and integration difficulty required for cognitive structure leap, i.e., the "energy level difference" between old and potential new structures. ΔE = E_{avail} - E_{th}. - h : Frequency or intensity of anomalous evidence, cognitive conflict, or 反常 input per unit time, i.e., "excitation intensity."

When h ≥ ΔE is satisfied, the probability of structural leap (P_{jump}) increases significantly, thus leap probability equals (Formula 6):

$$P_{\text{jump}} = \frac{1}{1 + e^{-\kappa(h\nu - \Delta E)}}$$

where is the coupling parameter, reflecting coupling strength of factors like individual background schema organization ability, motivation intensity, and social/situational support. This formula indicates that leap probability has nonlinear sensitivity to excitation input and energy budget. When the difference between energy level difference (ΔE) and excitation intensity (h) is not zero, the system is in a non-steady state and may trigger cognitive leaps, i.e., high-level creative behaviors like insight and paradigm shifts.

The occurrence of leaps will make the difference between system energy level difference (ΔE) and excitation intensity (h) approach zero. At this point, the cognitive system returns to a steady state.

Next, we adopt Friston's (2010) "Free Energy Principle"to describe changes in the difference between system energy level difference (ΔE) and excitation intensity (h). According to predictive processing theory, cognitive activities of biological intelligent agents can be seen as continuous regulation processes to minimize free energy, ultimately achieving better environmental predictions and stable internal system states (Friston, 2010). Thus, the difference between system energy level difference (ΔE) and excitation intensity (h) can be transformed into system free energy (F), expressed as (Formula 7):

$$F = E - TS$$

where E represents predictive energy error, representing the prediction difference between the agent' s internal model and external input (i.e., the difference between ΔE and h in Formula 6); S represents system entropy, i.e., uncertainty about input information; T is system activation temperature, reflecting cognitive mobilization degree or resource input level. TS describes the system' s thermodynamic activity level, i.e., "enthalpy." When cognitive systems undergo leap creativity, we assume system enthalpy remains constant, so free energy (F) changes only relate to expected energy error (E).

If new paradigms can effectively explain original anomalous information (E) with more concise structures, predictive errors decrease accordingly, information entropy is compressed; when system activation can be mobilized through more optimized paths, overall cognitive load can also be reduced. The result of paradigm leap is minimizing free energy (F) (Friston, 2010).

Therefore, leap creativity leads to the difference between new free energy level (F_{new}) and old free energy level (F_{old}) (Formula 8):

$$\Delta F = F_{\text{new}} - F_{\text{old}} < 0$$

Leap creativity involves the synergistic action of three mechanisms: cognitive psychology, physics, and neurodynamics (see Table 1).

Table 1 Three Mechanisms Explaining Leap Creativity

Mechanism Type	Core Process	Analogous Physical Model	Neurocomputational Description
Cognitive Psychological Mechanism	Equilibrium-disequilibrium-new equilibrium	Electron energy level transition analogy	Free energy minimization principle
	Cognitive schema rupture due to input disturbance, generating new schema	Information excitation causes cognitive system to leap to new energy level	Achieving system steady state through predictive error minimization

Through the unification of the above three coupled mechanisms, leap creativity can be redefined as a systematic cognitive leap process centered on paradigm reconstruction. Its basic dynamics originate from external disturbances causing original cognitive schema disintegration and new cognitive schema reorganization, driving the cognitive system to leap from a high free energy state to a new low free energy steady state. Therefore, leap creativity can be simplified as system free energy reduction (see Formula 9):

$$C_{\text{leap}} = -\Delta F$$

Leap creativity not only represents optimization at the information organization level of cognitive structures, but also reflects improvement in system predictive ability. In complex systems, creative leaps often manifest as the reshaping process of "cognitive free energy valley bottoms," i.e., the system leaps from a set of locally optimal predictive structures to globally optimal structures with greater overall stability and adaptability, thereby achieving hierarchical updates in the paradigm sense (Spuzic, 2024).

4 Creativity

In the previous sections, we analyzed the three basic forms of creative activities —interpolation, extrapolation, and leap—and mathematically characterized the absolute value of corresponding creative activities, such as the absolute amount of increased internal order of knowledge systems, characterized as ΔS; the absolute value of total system information gain, ΔI; and the reduction of system prediction error and establishment of new steady states, reflected as system free energy reduction ΔF. Although these indicators can describe the absolute value of outputs of various creative activities, to measure the efficiency of creative agents "maximizing value gains under limited resource conditions," we need to calculate the output-cost ratio of creative activities by dividing the creative output (i.e., entropy reduction, information gain, free energy reduction) by the cost paid for creation (time, computation, cognitive load). **Creativity** is the description of creative activity efficiency. Therefore, the definition of creativity cannot stop at the level of phenomenon description or outcome evaluation. Creativity is not the ability to produce novel results; it describes a system attribute, i.e., to what extent the subject can cause structural changes in cognitive systems under given resource budgets.

4.1 Evolutionary Constraints on Creativity: Adaptive Optimization in Open Systems

From a system evolution perspective, creativity is the ability to achieve local adaptive optimization under dual constraints of energy and structure. Life systems can maintain their own order against the background of overall entropy increase in the universe not because they violate the second law of thermodynamics, but because as open systems, they can continuously absorb energy and information input from the environment, thereby maintaining dynamic low-entropy states within the system (Schrödinger, 1944). That is, entropy reduction patterns within systems come at the cost of entropy increase in the external environment. System adaptive activities convert environmental entropy increase into entropy reduction within the system. The efficiency of this entropy conversion reflects the system' s adaptability level. Because systems with high entropy conversion efficiency can achieve greater internal entropy reduction based on relatively lower entropy increase costs, such systems gain greater survival and reproduction probabilities among multiple systems competing for survival and development opportunities, thus maintaining stability and development within the system.

According to Darwin' s evolutionary theory, natural selection means "individual variations that better match environmental change characteristics will gain more survival and reproduction opportunities"(Darwin, 2005). The matching between individual variation and environmental change characteristics is "adaptation." However, each individual variation has a cost; while gaining survival advantages in one aspect, it may also increase survival threat risks in another. For example, the variation of human bipedal walking gave humans huge survival advantages; the physiological-anatomical structure of upright posture also increased risks of cardiovascular disease and spinal injury. But weighing the two, advantages outweigh risks; thus the variation of bipedal walking reflects optimized adaptive expression of humans in natural selection processes; the physiological structure of upright posture also became a structure of "maximizing fitness under specific costs" in human evolutionary processes. Ashby (1956) expanded on biological evolution mechanisms within a systems theory framework, proposing that highly adaptive structures usually have a key characteristic: obtaining higher levels of prediction and control ability with lower energy consumption. This characteristic specifically manifests as cognitive systems maintaining structural complexity and behavioral flexibility under resource-limited conditions, and obtaining broad understanding of the environment through less information or energy. That is, the optimization goal of mental systems is not only to process information accurately, but to maximize adaptability under energy consumption constraints.

All biological intelligence evolution processes are constrained by "energy budgets" (i.e., cognitive resources mobilizable per unit time) and "structural constraints" (i.e., processing capacity, representation modes, and storage patterns of cognitive systems). Well-adapted systems do not unlimitedly expand their response spaces, but achieve optimal regulation of disturbances under limited resource conditions (Ashby, 1956). This idea is deepened in Prigogine and Nicolis' s (1977) research on dissipative structures: all structural orders far from equilibrium are built upon nonlinear stable mechanisms driven by continuous energy flow; "order generation" in information processing never separates from dependence on energy consumption (Nikolis & Prigogine, 1977).

4.2 Free Energy Formulation of Creativity

According to the free energy principle, creative activities manifest as overall system free energy decrease ($-\Delta$F), with leaps in cognitive efficiency and structural stability. Clark (2016) further expanded this explanation into the "predictive processing framework," emphasizing that the core function of human intelligence is to maximize environmental prediction ability while minimizing cognitive energy consumption (Clark, 2016). Under continuous entropy increase pressure, creative activities manifest as reconstruction of existing generative models, thereby achieving regeneration of local order and re-optimization of system energy distribution. A cognitive system leaping from an initial state to a new steady state reflects system free energy reduction; it also requires energy consumption costs. If a system consumes less energy to achieve similar steady-state transitions, it indicates higher transition efficiency.

Therefore, when we adopt the free energy principle to define creativity, we need to consider both free energy decrease ($-\Delta$F) and the energy budget paid for it, to fit thermodynamic efficacy evaluation methods: "efficiency = effective output / energy consumption." Accordingly, creativity defined from the free energy principle can be formally expressed as (Formula 10):

$$C = \frac{-\Delta F}{E}$$

where $-\Delta$F represents the amount of system free energy reduction, and E is the energy resources invested in cognitive or information processing processes. Formula 10 characterizes the degree of cognitive system optimization achieved per unit of energy input. Creativity level is related not only to system structural leaps and generative model reorganization outcomes, but also to the resource consumption required by the system.

In 1953, information theory founder Léon Brillouin proposed the concept of "negentropy" (Brillouin, 1953), thereby connecting information theory with entropy concepts in thermodynamics. He pointed out that information is essentially entropy reduction, or "increase in orderliness." In this framework, each bit of information can be regarded as corresponding to one unit of entropy reduction in the system: 1 bit = $-\Delta$S. This view provides a physical foundation for measuring information volume, and also reveals the energy cost of cognitive creation processes. Any new knowledge generation necessarily marks system information increase and uncertainty reduction, i.e., information entropy decrease. Therefore, from an information-thermodynamics perspective, "creation" is not only a conceptual activity, but also a thermodynamic process–information gain = entropy reduction = physical realization of new knowledge.

In 1961, physicist Rolf Landauer proposed the physical limit principle of information processing, pointing out that information manipulation is inseparable from the physical world (Landauer, 1961). Its core proposition is: "Information processing is not cost-free." In any physically realized computing system, especially during irreversible operations (like information erasure), there must be accompanied minimum energy dissipation. Landauer' s lower bound for minimum energy consumption is expressed as (Formula 11):

$$E_{\min} = kT \ln 2$$

where E_{min} is the minimum energy required to erase one bit of information; k is the Boltzmann constant; T is system temperature, representing physical temperature or cognitive activation level; ln2 originates from the entropy value of two states (0 or 1) possessed by one bit.

This formula indicates that information creation, maintenance, and erasure all require energy consumption, and this energy has its insurmountable theoretical lower bound.

Combining Landauer' s principle with Brillouin' s "information is negentropy" definition, we can derive an important corollary: any creative information output theoretically must cross a minimum energy consumption threshold. This threshold not only defines the physical foundation of intelligent systems when processing new information, but also provides a basic constraint condition of "information-energy consumption ratio" for creativity quantification. For example, if we compare a cognitive system (like the brain) or artificial system (like a computer) to a room storing information, with numerous documents in the room corresponding to the system' s current knowledge state and structure, when the system attempts creative activities, it often needs to clear outdated information, break original frameworks, and construct new logical structures. This process inevitably involves information "deletion," "rewriting," and even "reorganization" of entire cognitive models. Landauer' s principle points out that each such operation accompanies energy dissipation. In computers, this energy consumption manifests as electrical energy conversion; in brains, it manifests as energy released by glucose metabolism. Information erasure and structural reorganization are not only logical processes, but also physiological or physical processes. Therefore, creative activities are not "free thinking," but costly order reconstruction.

In 1973, Charles Bennett further developed Landauer' s principle. He pointed out that if a computing process is logically reversible—i.e., each operation step preserves complete information about the previous state—then this process theoretically requires almost no energy dissipation (Bennett, 1973). This concept is called "logical reversibility." We can vividly compare it to a person walking in a maze: if he accurately records his path at each step, then even deep inside the maze, he always has the ability to return along the original path. In this process, since no information is "forgotten" or "erased," the system does not need to pay additional thermodynamic costs for lost states.

But the problem is that completely realizing computational process reversibility requires preserving all intermediate computational steps, input states, and transformation paths, thereby significantly increasing storage space and model complexity required by the system. Regarding this, Rissanen further proposed: between limited cognitive resources and modeling efficiency, how to achieve optimal balance between "model simplicity" and "expressive power" ? This is the **Minimum Description Length (MDL)** principle proposed by Finnish mathematician Jorma Rissanen in 1978, aiming to evaluate model quality using information theory methods. This principle holds that a good model should achieve minimum encoding length for data while maintaining high explanatory power, i.e., reaching optimal balance between expressive power and structural simplicity (Rissanen, 1978). Overly complex models, although they may improve prediction accuracy, are not conducive to generalization due to high redundancy and encoding costs; while overly simplified models may not sufficiently capture key differences in data structures. Therefore, the MDL principle provides a trade-off standard between information compression and model fitting, i.e., the optimal modeling process should effectively represent the widest range of observed patterns with minimal information complexity.

Traditional creativity measurement indicators are positively correlated with information reserves, i.e., believing that individuals with richer information reserves have greater creative potential. However, in AI-centered digital environments, information availability no longer constitutes a major limiting factor. Instead, the key issue shifts to information quality, degree of structuring, and ability to explain complex systems or adapt to new situations. Thus, the essence of creativity lies not in the absolute amount of information, but in information organization efficiency and knowledge structure representation ability.

Taking "description of climate change" as an example, one descriptive method is to exhaustively list historical data, which has huge information volume but lacks structural compression, making it difficult to provide effective explanation; another method is highly simplified statements, such as "the climate is changing," which have extremely high compression but almost no predictive or explanatory power. In contrast, high creativity manifests as constructing a set of models that are structurally compact, have limited variables, and possess predictive ability, capable of retaining explanatory and predictive power for phenomena while maximizing information length compression.

This is precisely the theoretical significance of the Minimum Description Length (MDL) principle in creativity research. That is, efficient creative behavior does not rely on information accumulation, but on expressing the highest information structural complexity with the lowest representational cost, achieving optimal "energy-information" efficiency ratio. Therefore, we can understand creative activities as information structure construction processes with energy costs, i.e., an "energy investment behavior." In this framework, creativity can be defined as effective information output per unit of energy input, defined as (Formula 12):

$$C = \frac{\Delta I}{E}$$

where ΔI represents Information Gain, and E represents the energy input required to achieve this information gain.

According to Brillouin' s definition of "information is negentropy" (Information = −Entropy), information gain can be regarded as system entropy reduction, i.e. (Formula 13):

$$\Delta I = -\Delta S$$

Thus, we obtain the formal definition of creativity (Formula 14):

$$C = \frac{-\Delta S}{E}$$

That is, creativity equals the ratio of system orderliness improvement (entropy reduction) or information gain amount to energy consumption. The greater the system information gain and entropy reduction, and the lower the energy consumption, the higher the creativity level.

4.3 Formal Definition of Creativity

Based on the above arguments, this paper proposes the following propositions:

Proposition 1: The essence of creativity is adaptive optimization. As previously argued, in resource-limited ecological niches, mental and social mechanisms that bring higher fitness gains at lower costs are more likely to be preserved by natural selection and continued through culture (Ashby, 1956; Darwin, 1859/1964). In this framework, the meaning of creativity couples with system steady-state principles and free energy minimization mechanisms, while also maintaining formal isomorphism with learning criteria like the Minimum Description Length (MDL) principle and Information Bottleneck theory that "promote generalization through compression" (Rissanen, 1978; Tishby et al., 2000). Therefore, we define creativity as: achieving adaptability improvement under constraints.

Proposition 2: Creativity manifests as maintaining the quantity and quality of creative output under limited energy consumption conditions. High creativity does not equal "producing more content," but refers to the system' s ability to form more generalizable and explanatory output capabilities for unknown task domains under energy or resource constraints. In other words, the essence of creativity is task domain coverage gain per unit energy consumption. It can be formally expressed as (Formula 15):

$$C = \frac{\Delta CA}{E}$$

where C represents creativity, ΔCA represents the expansion amplitude of explainable/controllable state spaces or task domains (Cognitive/Controllable Area), representing the increment of new content or new structures the system can create under current conditions; E is the energy consumed in the corresponding process. This formula characterizes the energy efficiency ratio of the system generating "meaningful structures" under specific resource constraints.

Proposition 3: The gain of creative activities in corresponding domains (interpolation, extrapolation, and leap), i.e., the improvement of controllability/explainability of task domains or state spaces, is equivalent to system entropy reduction, information increase, and free energy reduction within those domains, i.e. (Formula 16):

$$\Delta CA \equiv -\Delta S \equiv \Delta I \equiv -\Delta F$$

Based on the above three propositions, this paper proposes a formal definition of creativity in three domains: within a given time window Δt and task domain D, the ratio of entropy reduction (−ΔS), information gain (ΔI), and free energy decrease (−ΔF) achieved by the subject on the "internal representation-external environment" joint system in the corresponding domain to the energy consumption E paid for it. That is (Formula 17):

$$C = \frac{-\Delta S}{E} = \frac{\Delta I}{E} = \frac{-\Delta F}{E}$$

This definition describes the measurement method for creativity in interpolation, extrapolation, and leap domains, and stipulates that measurement values in the three domains are equivalent. Accordingly, creativity in the three domains can be evaluated and measured separately, while measurement values in each domain also have additivity. Further explanations:

1. **Creativity is a manifestation in human-environment joint systems**: When mentioning creativity, we should not examine the "brain" as an isolated system, but should consider individuals and their external environments together as a cognitive-operation joint system. In actual creative activities, individuals often rely on multiple external support factors such as tools, funding, materials, time and space conditions, and collaborative teams. Therefore, creativity performance should be measured embedded in the "subject-environment" interactive system. To construct a systematic creativity measurement framework, the following two basic constraints need to be clarified: **Time window (Δt)**: Refers to the time interval of creative activity duration, such as 30 minutes or one day; **Task domain (D)**: Refers to the specific task scope the system needs to complete within that time period, such as writing a text or designing a prototype.

2. **Creative activity measurement indicators are transformations from chaos to structure, from unknown to controllable**: From the "joint system" perspective, the "benefits" brought by creative activities

are not only manifested as increased output quantity, but should be understood as optimization of system cognitive structure and expansion of functional domains. Specifically, these benefits can manifest as: Information ordering: Enhanced ability to organize originally chaotic information, manifested as posterior entropy decrease; Improved prediction ability: More accurate expectations for future states, manifested as surprise reduction; Enhanced structural compressibility: Expressing broader phenomena with more concise models, reflected as reduced model complexity or shortened minimum description length; Enhanced explanation and control ability: Expanded range of explainable, operable, and predictable state spaces, manifested as enhanced task domain coverage ability.

3. **The value of creative benefits is constrained by input costs**: When measuring the actual value of creative benefits, the following cost items must be considered simultaneously: Time and physiological energy consumption: Including energy consumption from physical activity, cognitive resource allocation, and metabolic energy consumed by brain neural activity; External resources: Such as tools, data, computing platforms, and economic costs; Cognitive load: The psychological-cognitive pressure individuals bear when processing complex problems, including subjective "cognitive fatigue" and emotional experiences in operational practice; Algorithm operation costs (for artificial intelligence systems): Including computation amount (such as FLOPs), model parameter scale, memory occupation, and server energy consumption, etc.

4. **Creativity expression can be reflected in two dimensions of certainty and transcendence**: In a further abstract analytical framework, creativity can also be measured through two-dimensional indicators (see Table 3):

Table 3 Two Types of Creativity

Type	Description	Example
C_{cert} (Certainty)	Making problems clear, making things simple	Understanding a formula, summarizing a document
C_{trans} (Transcendence)	Introducing new concepts, breaking old frameworks	Inventing a completely new concept, discovering a new problem

C_{cert} reflects information ordering and compression ability, an indicator of high-efficiency expression within existing frameworks; while C_{trans} corresponds to problem space expansion and structural leaps, emphasizing transcendence and rewriting of original explanatory systems. Together they constitute the functional quality of creativity performance.

5. **Domain specificity of creativity performance**: Multiple studies also show that in human-AI collaborative completion of complex tasks, creative behavior demonstrates obvious structural division of labor and regularity. Specifically, when facing tasks with unclear problem definitions and high path openness (i.e., "ill-structured tasks"), humans are more inclined to perform conceptual leaps, generate new frameworks, or reconstruct problem settings, showing higher C_{trans}; while in problems with clear steps and explicit rules (well-structured tasks), AI is better at optimizing expression and advancing solution processes within established structures, showing higher C_{cert} (Li, Huang, Liu, & Zheng, 2022; Singh, Lyu, Deterding, et al. 2025).

Through the above explanations, we can see that creativity improvement does not have a simple linear relationship with energy input. In routine tasks (such as calculation, retrieval, structural reorganization), increasing resources (such as extending working hours, improving computing power) does help improve processing efficiency and task completion; but in high-level creative activities involving paradigm shifts or concept reconstruction, systems show obvious leap threshold characteristics. This means that leap innovation often depends on triggering a certain "energy critical point." Before reaching this threshold, continuously investing more energy cannot effectively drive mutational reorganization of cognitive structures; but once excitation conditions are satisfied, the system may undergo nonlinear leap reactions, generating new conceptual frameworks or explanatory paths.

Furthermore, research shows that external disturbances and "abnormal inputs" can often stimulate creative leaps. When task environments contain some seemingly "anomalous" but key meaningful elements, individuals are more likely to jump out of existing frameworks and enter creative processing states. Synergetics' principle of disturbance-induced reorganization explains this phenomenon: high-order ordered states of systems often originate from destruction and challenge to steady states; in other words, innovation often originates from being "disturbed" and "challenged" (Haken, 1976; Kelso, 1995).

In terms of group creativity, loosely structured small teams are also often more likely to stimulate higher "per-unit energy consumption" creativity than resource-sufficient but path-dependent large teams. A 2019 study by Wu et al. showed that although small teams do not have advantages in resources and stability, because their collaborative structures are more flexible and knowledge distribution is more diverse, they are more likely to achieve "outside-paradigm" breakthrough results per unit energy consumption, showing higher disruptiveness index (Wu et al., 2019).

4.4 Parameter Value Constraints and Verification for Creativity Definition

To ensure robustness and operability of the creativity definition (Formula 17) in cross-subject and cross-situation measurements, we need to systematically specify constraints for its core parameters. Specifically, reasonable value constraints and boundary conditions should be proposed for variables involved in the definition (such as entropy reduction −ΔS, energy consumption E, information gain ΔI, free energy decrease −ΔF, etc.) from four aspects, and a set of falsifiable evaluation indicators should be established.

1. **Computational Consistency Verification: Boundaries, Scaling, and Additivity**
   To ensure creativity indicator C = −ΔS/E = ΔI/E = −ΔF/E has stable computational consistency across different systems and domain tasks, we propose the following constraints and technical processing suggestions for parameter values:
   Boundary conditions: When entropy reduction ΔS → 0 or output lacks effectiveness (i.e., appropriateness factor A 0), creativity indicator should approach zero (C → 0). If apparent ΔS > 0 but energy consumption E → 0 occurs, this usually reflects that costs of external resources (such as prior knowledge, group experience libraries, external memory systems) are not accounted for. In such cases, total E should be recalculated based on the "subject-environment joint system" (Cover & Thomas, 2006; MacKay, 2003).
   Composition/Additivity: For multiple independent sub-domain tasks D_i, total creativity C_{tot} does not equal simple summation of indicators for each sub-domain task. It is recommended to calculate overall C_{tot} using weighted average referencing "proportion of domain creativity value in total creativity value," i.e., using harmonic mean to highlight contributions of "bottleneck tasks" (Jaynes, 1957).
   Scale invariance: When time window Δt scales proportionally and corresponding benefits and costs change proportionally, C should remain unchanged. This constraint ensures the indicator' s transferability and comparability across different time scales.
   Effectiveness filtering mechanism: Introduce appropriateness/verifiability factor A(D) [0,1] to weight and correct C, defined as (Formula 18):

$$C_{\mathrm{adj}} = A(D) \cdot \frac{-\Delta S}{E}$$

   where A(D) can be estimated based on external indicators such as prediction accuracy improvement, control effectiveness improvement, or expert evaluation scores (Amabile, 1983; Runco & Jaeger, 2012). This correction formula excludes low-quality outputs that are formally entropy-reducing but actually ineffective[5].

2. **Physical and Computational Lower Bound Verification: Observable Energy Constraints**
   To ensure the realizability and physical consistency of creativity indicators, computational logic must be limited within observable minimum energy consumption floors. We propose three constraint evidences:
   Landauer-Bennett lower bound: Set minimum dissipation for any irreversible operation as kTln2, where k is Boltzmann constant and T is thermodynamic temperature (Bennett, 1973; Landauer, 1961). This bound applies to basic algorithm operations like information erasure, normalization, and rejection-acceptance steps in random sampling, as well as basic energy consumption for cognitive updates and perception-motion coordination in embodied cognition.
   Neural energy budget constraint evidence: The human brain's average power is about 10-20W; over 90% of energy consumption concentrates on synaptic transmission, action potential generation, and transmembrane ion pump operations (David Attwell & Simon B. Laughlin, 2001; Niven & Laughlin, 2008; Raichle & Gusnard, 2002). This budget supports the energy efficiency optimization strategy of "minimum input → high order output," and is also an adaptive strategy that natural systems inevitably form "low-consumption creativity" through long-term evolution.
   Empirical counterexample constraint evidence: If a system simply increases energy consumption (such as expanding model scale, extending training time) without obtaining sufficient entropy reduction (ΔS), overall creativity C_{tot} will decrease. This corresponds to the "high computing power-low output plateau phenomenon" in practice, where marginal computing resources cannot effectively convert into structural benefits (Belkin et al., 2019; Goyal & Bengio, 2022).

3. **Information Theory and Learning Theory Verification: Compression-Generalization Consistency**
   From learning theory and information theory perspectives, the unit information gain ΔI/E involved in creativity definition can be examined for correspondence and consistency with model compressibility, generalization ability, and statistical verifiability from four aspects:
   MDL and Information Bottleneck (IB) principle verification: In MDL framework, models need to explain data with shortest description length (Rissanen, 1978); while IB methods require maximizing compression of intermediate representations while retaining sufficient task-related information to avoid redundancy (Achille & Soatto, 2017; Tishby et al., 2000). Both point to the same core principle: "compression is generalization," i.e., ΔI/E can be regarded as generalization return rate brought by compression.
   PAC-Bayes and compression bound verification: There is a clear theoretical connection between compressibility and generalization error upper bounds. The PAC-Bayes framework and its derived compression bounds can map system compression ability to statistical learning gener-

alization indicators, thus providing verifiable paths for ΔS (Dziugaite & Roy, 2017; McAllester, 1999; Neyshabur et al., 2018).

Maximum entropy principle and prior predictive rationality verification: Jaynes (1957/2003) proposed that selecting the prior with maximum entropy under given constraints can avoid introducing unwarranted assumption biases. In this definition framework, creativity can be seen as the ability to achieve greater posterior entropy reduction after introducing new constraints, and this entropy reduction is achieved with minimum energy consumption (Jaynes, 1957).

Inductive bias and Solomonoff/MDL theory verification: Solomonoff induction and AIXI framework emphasize "shortest program is optimal predictor," i.e., shortest description brings maximum generalization ability, providing theoretical limit arguments for the equivalence relationship between "compression → creativity" (Hutter, 2005; Solomonoff, 1964)[6].

4. **Evolutionary and Cybernetic Verification: Adaptability and System Steady-State Consistency**
From evolutionary psychology and cybernetic perspectives, creativity indicator C can be seen as system adaptability improvement obtained per unit energy consumption, while maintaining relative system steady-state:
According to natural selection principles, in resource-limited ecological niches, cognitive and social mechanisms that can achieve higher fitness gains with less energy consumption are more likely to be preserved and inherited through cultural mechanisms (Darwin, 1859/1964; Kauffman, 1992; Szathmáry & Smith, 1995). Therefore, high C values can be understood as advantageous strategy selections in system evolution.
Cybernetic steady-state mechanism verification: According to Ashby' s (1956) "law of requisite variety," regulator systems must possess response diversity matching environmental disturbances to maintain system steady-state. Creativity indicator C can be understood as an efficiency measure of maintaining requisite variety at minimum cost, i.e., the cost-effectiveness between system regulation ability and energy consumption (Ashby, 1956).
Evidence from group-level organizational structures: Wu et al. (2019) research shows that under unit resource conditions, small-scale teams are more likely to produce results with high "disruptive innovation" characteristics (Wu et al., 2019). This indicates that high C values often appear in small systems with flexible structure, low coordination costs, and high exploration efficiency, suggesting "creative efficiency" can serve as an important parameter for optimizing group innovation strategies.

5. **Comprehensive Falsifiability Verification**
To verify the empirical operability of creativity definition and its parameter system, this paper proposes five hypothetical propositions (P1-P5) for subsequent falsifiability testing. The hypotheses cover human-machine comparison, energy thresholds, information density regulation, learning

theory indicators, and group structure aspects:

P1: Cost-matched human-machine comparison: Under energy consumption normalization conditions, human individuals or small-scale teams will show higher C_{trans} levels than AI systems in problem reconstruction and concept introduction tasks.

P2: Energy increment threshold effect: Under controlled "anomalous evidence density" conditions, the occurrence frequency of extrapolation/leap events shows threshold-mutation characteristics with energy consumption, accompanied by variance amplification and critical slowing phenomena.

P3: Anomalous evidence density regulation: By increasing "anomalous evidence density" in task inputs, the energy threshold E_{th} required for the system to achieve extrapolation or leap can be effectively reduced. This mechanism is consistent with the "disturbance-subordination" performance described by synergetics theory.

P4: Compression-generalization correlation: Under equal energy consumption conditions, models with higher MDL compression gains and lower PAC-Bayes generalization bounds can predict higher C_{cert} and stronger OOD (out-of-distribution) generalization performance.

P5: Team structure optimization effect: Under controlled total energy consumption and time budget, small-scale or weakly-connected cross-domain teams show significantly stronger "disruptive innovation" ability per unit energy than large-scale and strongly-connected team structures, supporting the optimization value of C_{tot} at the group level and organizational aspects.

4.5 The Brain is a "Thermodynamically Efficient" Creativity Engine

Compared with computers, the human brain is an extremely efficient creativity engine in thermodynamic terms. Neuroscience research shows the brain's basal metabolic energy consumption is about 20W. In contrast, contemporary supercomputers typically consume thousands of watts to megawatts when performing tasks like model generation, problem solving, and logical reasoning. Taking OpenAI's training of GPT-3 as an example, the process used thousands of GPUs, with cumulative training energy consumption estimated at hundreds of megawatt-hours. Even during inference, running a medium-scale language model for real-time generation often requires tens to hundreds of watts of power (see comparison in Table 2). Hinton also pointed out at the 2025 World Artificial Intelligence Conference that the most powerful AI systems may face the real risk of "power-off termination" [7]. This reflects from the side that current AI systems are highly dependent on continuous energy input.

This shows that although computers or AI systems can generate large amounts of content per unit time, their high energy consumption mechanisms make their "creativity" far lower than that of the human brain.

Table 2 Power Consumption and Time Dimension Description for Each Device

Device	Power	Duration	Description
Human Brain	20W	24 hours continuous (resting/thinking state)	The brain almost never "shuts down." Even during rest (sleep, meditation), the default mode network (DMN) remains active, so it' s 24-hour continuous power consumption.
Ordinary Laptop	60W	In-use operation (active state)	Refers to average power consumption during daily tasks like opening browsers, document editing, light programming; hibernation/standby significantly reduces power; full-day use requires multiplying by operating hours.
Large AI Server (e.g., GPT inference)	1000–5000W+ instantaneous power (seconds to minutes level)	Model inference peak power, i.e., instantaneous power demand for one generation response (like answering a question); not continuous all-day operation, but cluster long-term use can accumulate huge energy consumption.	

Device	Power	Duration	Description
Training Large AI Models (e.g., GPT-3)	Millions of watts total consumption	Days to weeks continuous	Total energy consumption calculated as "power × total training time." Training phase requires high-parallel GPU/TPU operations, consuming large amounts of electricity per second, usually measured in MWh (megawatt-hours).

[7] See domestic formal media reports: https://m.163.com/dy/article/K5DI8VN80532N2UB.html

5 The "Three-Dimensional" Model of Creativity Measurement

To achieve creativity measurement, we need to construct corresponding quantitative representation models for creative cognitive activities in three dimensions: interpolation, extrapolation, and leap. Therefore, this paper proposes a three-dimensional creativity measurement model (3D Model of Creativity), corresponding respectively to the three basic forms of creative activities: the X-axis represents **Interpolation Depth**, measuring the ability for element recombination and optimization within established paradigms and the same structural level; the Y-axis represents **Extrapolative Breadth**, characterizing the ability for structural extension and problem migration near existing knowledge boundaries; the Z-axis represents **Abstractive Altitude / Structural Transition**, used to measure cross-level structural recombination ability, especially the cognitive potential to construct unified explanatory frameworks at higher abstraction levels to subsume heterogeneous phenomena.

5.1 Interpolation Depth

The X-axis represents Interpolation Depth, referring to the ability of individuals to enhance system complexity and internal consistency through deductive reasoning and structural recombination within established cognitive frameworks.

Creative activities do not necessarily depend on paradigm breakthroughs or introduction of new concepts. In many knowledge-intensive fields, such as mathematical proof, program design, engineering system modeling, and even artistic creation, deep mining and formal refinement of existing rule systems and structural configurations constitute creative processes. In the interpolation dimension, creativity manifests as the ability of cognitive systems to achieve optimal organization in the tension between structural compressibility and combinatorial complexity within original paradigm structures. That is, if individuals can achieve higher density, tighter, and higher-resolution knowledge structure configurations while maintaining logical consistency and functional effectiveness, their interpolation creativity is higher.

Creativity in the interpolation dimension usually has an approximately linear functional relationship with the following variables: First, the size of basic knowledge reserves (for AI, this is training dataset size). Introducing more and more heterogeneous knowledge elements can expand potential space dimensions for structural recombination and configuration; second, cognitive ability or computing power. Only by enhancing system cognitive processing ability or computing power can complex tasks like multi-dimensional structural search, combination path optimization, and error inversion be realized; third, model depth and parameter complexity. Deeper models and more complex parameter training help capture finer-grained dependency patterns and coupling features in multi-layer nested structures; finally, training. Accumulating more training experience can gradually optimize semantic mapping and structural compression, improving inference chain efficiency and shortening path length to achieve optimal structural recombination. For humans, this manifests as automated processing or skilled operations formed through training.

Current large language models (like ChatGPT) and deep learning systems often show significant creativity in the interpolation dimension; especially in tasks like automated mathematical proof generation, code structure optimization, and artistic style fitting, AI has excellent performance in detail-level combination ability and paradigm-internal structural recombination efficiency.

Measurement of creativity in the interpolation dimension can adopt multiple indicators as shown in Table 4 :

Table 4 Creativity Evaluation Indicators for Interpolation Dimension

Indicator	Formal Definition	Typical Measurement Method (Example)	Interpretation and Application Scenarios
Reasoning Depth	L_{depth} = average inference chain length	Graph structure analysis, tree level of deductive paths	The longer the effective reasoning chain to complete target tasks, the deeper the structural processing
Combinatorial Density	D_{density} = effective connections between elements	Hypergraph node average degree, combinatorial scale estimation of state space	The more connection and combination path numbers between effective constituent elements, the higher the combinatorial complexity

Indicator	Formal Definition	Typical Measurement Method (Example)	Interpretation and Application Scenarios
Structural Compression	C_{comp} = (L_{base} - L_{model}) / L_{base}	Minimum Description Length (MDL) ratio, symbol redundancy compression rate	Focus on trade-off between information compression and generation quality; suitable for high-structure tasks like code rewriting, proof simplification, structured writing
Consistency Robustness	R_{cons} = output stability under single-point perturbation	Output structure variance under single-point perturbation conditions, error propagation sensitivity analysis	Ability to maintain logical/functional output stability under local perturbation or replacement conditions
Energy Efficiency Ratio	C_E,X = Q_{out} / (E_{comp} / T)	Record effective output quality Q_{out} generated per unit computation/time/neural metabolism, normalized by computational energy consumption E_{comp} and duration T	Directly binding creative effectiveness with resource input, facilitating cross-subject (human-machine) benchmarking and efficiency comparison

The above indicator system can be used to evaluate AI systems (like language models or symbolic reasoning programs) in high-density tasks, and is also applicable for horizontal comparison and longitudinal tracking analysis of human individuals or collaborative teams in dimensions like knowledge compression, deductive efficiency, and logical robustness.

It should be emphasized that improvement in interpolation depth should not be misunderstood as "shallow processing" or "low-level innovation." In many fields with highly rigorous cognitive structures, truly meaningful innovation often originates from deep mining of existing rule spaces and limit recombination within boundaries. Therefore, creativity in the interpolation dimension should be systematically evaluated based on depth of structural processing, efficiency of element combination, and compressibility and precision of model expression, rather than simply being classified into the "low-level segment" of the creativity spectrum.

5.2 Extrapolative Breadth

The Y-axis represents Extrapolative Breadth, reflecting the ability of individuals or systems to explore and extend near existing knowledge paradigm boundaries. It focuses on how creators migrate, transform, or reconstruct existing knowledge representations to effectively apply them to new domains, contexts, or functional goals.

In the extrapolation dimension, the value of creativity lies not only in "executability," but more fundamentally in "whether existing knowledge can be cross-contextually migrated to problem spaces that have not yet been clearly defined." The level of creativity depends on whether individuals or systems possess the following abilities: Breaking through internal constraint conditions set by current paradigms; Identifying weak connection relationships and potential mapping paths between heterogeneous structures; Migrating and applying familiar rule systems to unfamiliar problems or domains; and Still achieving structural adaptation and functional generation in contexts lacking clear guidance or with changed objective functions.

Creative activities in the extrapolation dimension usually present a series of characteristics. First, **gradual breakthrough**: Usually begins with fine-tuning, reconstruction, or parameter modification of original structures, gradually deviating from established patterns to explore potential extension paths. Second, **boundary sensitivity**: The size of creative space is constrained by proximity to original paradigm boundaries, and boundary zones often contain high variation potential. Third, **structural mapping ability**: Subjects need to identify high-level analogical relationships or function-level mapping paths between original structures and target domains to achieve structural coupling for cross-domain migration. Fourth, **purpose-driven**: Extrapolation activities are often oriented toward "goal alignment," manifesting as technology migration (such as cross-industry applications), knowledge transformation (such as from basic theory to productization), and other cognitive reconstruction processes with clear adaptation goals.

Compared with the interpolation dimension' s emphasis on reasoning path rigor and expression compressibility, extrapolation creativity relies more on adaptability and robustness of structural migration, reflecting the system' s "cognitive migration ability" to maintain functional integration and flexible response in unfamiliar environments.

To systematically evaluate the creative migration ability involved in extrapolative breadth, we can construct an evaluation system from indicators like knowledge migration depth, structural mapping strength, and functional generalization breadth (see Table 5):

Table 5 Creativity Evaluation Indicators for Extrapolation Dimension

Indicator	Formal Definition	Typical Measurement Method (Example)	Interpretation
Domain Distance	D_{dist} = semantic distance between original and target domains	Semantic embedding space center vector distance (like Word2Vec similarity, graph structure hop count)	Original knowledge structure and target application domain' s conceptual similarity or functional difference
Mapping Path Diversity	N_{map} = number of cross-domain structural mapping paths	Cross-domain path count in heterogeneous knowledge graphs, or cross-modal counterpart logarithm	Number of cross-domain structural mapping paths the system can identify and activate
Reconstruction Robustness	R_{rec} = performance retention rate after migration	Performance retention rate, generalization error, or upper bound disturbance in migration tasks	Ability to maintain core functional pattern stability during migration and reconstruction
Functional Expansion Score	S_{func} = cross-task adaptation ability	Migration task count × average performance improvement amplitude, or cross-task adaptation surface gradient	Applicability and expansion ability of the same knowledge structure in multiple tasks or domains

Indicator	Formal Definition	Typical Measurement Method (Example)	Interpretation
Energy-Performance Curve Slope	k_E,Y = Perf_{OOD}/ E	Fix tasks, gradually adjust computing power/sample size, fit performance-energy curve	High slope indicates "cheap extrapolation" ; often approximates linear in log-log coordinates

Compared with the interpolation dimension' s focus on structural depth and deductive optimization, extrapolative breadth emphasizes more on collaborative performance in areas like structural generalization, goal adaptation, and concept expansion. Creativity in this dimension requires the system to not only master and apply existing rule systems, but also possess the ability to flexibly "jump out of rules" and perform structural migration and functional generalization under new tasks, new contexts, or new goal-driven conditions. When the system can effectively break through knowledge boundaries and expand structural adaptation spaces under limited cognitive resource conditions.

5.3 Leap Height

The Z-axis represents **Abstractive Altitude / Structural Transition**, i.e., the ability of cognitive systems to leap from existing knowledge levels to explanatory frameworks with higher subsuming power and expression density through paradigm reconstruction, structural abstraction, and model elevation. Creative behavior in this dimension is not simple knowledge accumulation or flexible scheduling of rules, but reconstruction of the internal organizational logic of original knowledge structures, with core characteristics being the leap from cognitive structure to cognitive meta-structure.

Different from the interpolation dimension' s fine processing and extrapolation dimension' s boundary extension, leap creativity points to systematic doubt and reorganization of existing knowledge organization logic by cognitive systems, and through structural abstraction and symbol compression, extracting new generative mechanisms or explanatory principles with subsuming power from low-level experience or operational modes.

Such creative activities usually occur when original paradigms cannot adequately explain anomalous experiences or handle structural conflicts. Their result is constructing new cognitive formats, conceptual spaces, or problem frameworks, thereby achieving overall leaps in knowledge systems at structural, hierarchical, and expression levels. Creativity is not only reflected in producing "different content" or "new forms," but more fundamentally in creating a new way of understanding the world. This leap marks the power operation conversion of knowledge systems from the structural level to the meta-structure level.

Leap creativity usually manifests as a continuous process from paradigm disequilibrium to format updating. First, when original knowledge systems frequently experience prediction failures and explanatory gaps when dealing with anomalous experiences, cognitive systems enter paradigm disequilibrium states, with internal tension and free energy levels significantly increasing. Second, continuous structural surprises prompt subjects to accumulate cognitive conflicts, existing formats struggle to accommodate input information, and the system shows high uncertainty and tendency to seek structural reorganization. Third, during this stage, subjects often enter a period of multiple format coexistence, simultaneously mobilizing multiple explanatory frameworks for cognitive operations. Although overall stability decreases, exploration flexibility and structural allocation ability are significantly enhanced. Fourth, with abstraction and reconstruction of high-level common structures, the system may form a new cognitive format where original formats are embedded as local structures or completely reorganized. Finally, if the new paradigm significantly outperforms the original paradigm in dimensions like prediction accuracy, energy efficiency, and explanatory compressibility, it may gain dominant position, thereby completing the leap transformation of cognitive systems.

Creativity in the leap dimension usually does not increase correspondingly with simple increases in input data volume or computing resources, but shows nonlinear and mutational change characteristics. First, leaps have obvious threshold effects; cognitive systems need to accumulate certain intensity of conflict or anomalous input (such as anomalous case density) to a critical point before possibly triggering structural reorganization. Second, before leaps occur, systems often exhibit critical slowing phenomena, i.e., decreased reaction speed, structural hysteresis, and increased instability, similar to activity characteristics of physical systems approaching phase transition points. Third, the leap process is essentially discontinuous, i.e., structural transformation occurs not through gradual evolution but through sudden, discrete emergence. Fourth, systems often accompany significantly amplified output variance during leap periods, i.e., increased volatility of creative outputs in form, effect, and adaptability, reflecting uncertainty expansion of potential cognitive spaces.

Based on the above characteristics, creativity evaluation in the leap dimension needs to introduce dynamic indicators that can capture hierarchical reconstruction and system instability; focus on identifying whether cognitive systems have achieved paradigm-level transformation and modeling method updates; and reflect key characteristics like abstract elevation, structural compression, and energy optimization.

Table 6 Measurement Indicators and Cognitive Meanings for "Leap Height"

Indicator	Formal Definition	Example Measurement Method	Cognitive Meaning
Structural Compression Leap Gain	SCJ = (L_{old} - L_{new}) / E	Given old theory/model MDL; construct new structure and re-evaluate MDL; record energy consumption	Description length decrease brought by unit energy consumption; higher SCJ indicates more significant structural order improvement
Surprise Collapse Rate	v_{FE} = - F(t)/ t / Δt	Monitor free energy/cross-entropy time series, take maximum slope in sudden drop interval	Larger slope indicates more "abrupt" leap; corresponds to mutation/phase transition signals
Paradigm Conflict Density	D_{conflict} = anomaly accumulation density	Anomaly ratio, prediction error distribution	Accumulation density of anomalous inputs or "unexplainable events" before triggering leap
Predictive Power Leap	ΔP_{predictive} = prediction improvement amplitude	AUC gain, explanation degree improvement curve slope	New model' s prediction improvement amplitude for old problems
Energy-Normalized Leap Value	C_Z = ΔSCJ_{cum} / E_{tot} or C_Z = -ΔF/E	Free energy estimated by combining prediction error + model complexity	Leap efficiency from organizational layer perspective, supporting investment decision-making

Note: Such indicators usually require long time-span evolution trajectory data and energy consumption records during model operation. At the implementation level, they can be collaboratively extracted using citation graphs, model logging systems, knowledge structure visualization tools (like CiteSpace), and computing power measurement mechanisms to support automated monitoring and longitudinal comparison of leap creative behavior.

5.4 Functional Description of the Three-Dimensional Creativity Model

Under the same energy budget or resource input, creative performance in interpolation and extrapolation dimensions can show linear or power function growth with data scale, computing power, and experience accumulation; while creative performance in the leap dimension shows typical discontinuity, usually requiring accumulation of cognitive tension and structural conflicts to a certain threshold before mutation occurs, with significant mutational characteristics.

The three-dimensional creativity assessment model shows that creativity should not be regarded as a single continuous indicator, but should be differentiated on functional dimensions. It provides a design framework for evaluation tasks or questions for creativity measurement and comparison between heterogeneous subjects (such as AI, human individuals, group systems).

According to this model, AI systems may have advantages over humans in X/Y axis dimensions; but due to limitations in structural reorganization and meta-cognitive regulation functions, there are capability bottlenecks in Z-axis leaps; relatively speaking, humans are more likely to activate Z-axis leaps under resource-limited, task-uncertain, or cross-domain fusion conditions.

The three-dimensional model can also serve as a unified framework for exploring structural characteristics and generation conditions of creative activities.

6 Three-Dimensional Measurement System and Operational Indicators

Based on the three-dimensional creativity evaluation model, this paper further constructs a quantitative measurement system for three creativity dimensions: X-axis (Interpolation Depth), Y-axis (Extrapolative Breadth), and Z-axis (Leap Height), and proposes corresponding indicator design and calculation implementation points. (Relevant variables and symbol definitions are detailed in Appendix A)

6.1 X-axis: Interpolation Depth (Interpolation)

To quantify creativity performance in interpolation depth, this paper introduces four types of indicators, measuring the interpolation creative process from four aspects: spatial coverage, solution diversity, compression expression ability, and resource utilization efficiency. Formal definitions, calculation paths, and applicable explanations for each indicator are shown in Table 7 .

Table 7 Operational Evaluation Indicator System for X-axis (Combinatorial Breadth) Dimension

Indicator	Formal Definition	Calculation Path	Interpretation and Application Scenario
Combinatorial Coverage	C_{cov} = Vol(H(S)) / Vol(H(U))	1) Project candidate solutions into unified semantic/structural vector space; 2) Calculate convex hull volume Vol(H(S)) of candidate solution set; 3) Normalize by baseline universe volume Vol(H(U))	Higher coverage indicates broader exploration range in established semantic space; suitable for text generation, melody arrangement, design space exploration tasks
Diversity/Decorrelation	D_{div} = E[		s_i - s_j
MDL/IB Compression Gain	ΔL_{MDL} = (L_{base} - L_{model}) / L_{base}	Use language model to estimate description length (like n-gram/GPT-2, etc.); 2) Estimate shortest description length MDL; 3) Calculate relative compression gain	Focuses on trade-off between information compression and generation quality; suitable for high-structure tasks like code rewriting, proof simplification, structured writing

Indicator	Formal Definition	Calculation Path	Interpretation and Application Scenario
Energy-Normalized Score	$C_{E,X} = Q_{out} / (E_{comp} / T)$	Record effective output quality Q_{out} generated per unit computation/time/neural metabolism, normalized by computational energy consumption E_{comp} and duration T	Directly binds creative effectiveness with resource input, facilitating cross-subject (human-machine) benchmarking and efficiency comparison

6.2 Y-axis: Extrapolative Breadth (OOD Generalization)

Extrapolative breadth mainly measures the system' s ability for knowledge migration and adaptation outside existing paradigm boundaries; it focuses on how the system applies existing knowledge to unfamiliar environments or task objectives, manifesting as adaptability to changing conditions and generalization ability to unknown inputs.

Table 8 Evaluation Indicator System for Y-axis (Extrapolation Depth) Dimension

Indicator	Formal Definition	Calculation Path	Interpretation
Extrapolation Radius	$R_{OOD} = \sup\{\delta\}$ {		δ
Causal/Structural Migration Score	$S_{str} = (1/N)\ \Sigma\ 1[f_{src}\ \ f_{tgt}]$	Compare reasoning consistency on structurally equivalent (but surface-transformed) task pairs	Measures whether truly learning structure rather than surface patterns (systematicity)

Indicator	Formal Definition	Calculation Path	Interpretation
Counterfactual Robustness	R_{cf} = 1 - Perf_{cf} / Perf_{orig}	Inject "anomalous/contradictory" perturbations into inputs, measure performance damage	Smaller value indicates more robust; reflects stability for rare/anomalous situations
Energy-Performance Curve Slope	k_E,Y = Perf_{OOD} / E	Fix tasks, gradually adjust computing power/sample size, fit performance-energy curve	High slope indicates "cheap extrapolation"; often approximately linear in log-log coordinates

6.3 Z-axis: Leap Height (Structural Transition)

The Z-axis reflects the structural level leap characteristics of creativity. Compared with the previous two dimensions, leap creativity not only relies on knowledge content accumulation or rule generalization, but more importantly involves reconstruction and elevation of cognitive structures themselves, often showing mutational, nonlinear, and high energy consumption characteristics.

Table 9 Evaluation Indicators for Z-axis (Leap Height) Dimension

Indicator	Formal Definition	Calculation Path	Interpretation
Structural Compression Leap Gain	SCJ = (L_{old} - L_{new}) / E	Given old theory/model MDL; construct new structure and re-evaluate MDL; record energy consumption	Description length decrease brought by unit energy consumption; higher SCJ indicates more significant structural order improvement
Surprise Collapse Rate	v_{FE} = - F(t)/ t / Δt	Monitor free energy/cross-entropy time series, take maximum slope in sudden drop interval	Larger slope indicates more "abrupt" leap; corresponds to mutation/phase transition signals

Indicator	Formal Definition	Calculation Path	Interpretation
Disruption Index	$D = (N_{later} - N_{prior}) / N_{total}$	Construct citation network, compare whether successor citations bypass predecessor core	>0 indicates replacement effect, significantly changing research trajectory
Deep Citation Depth	R_{deep} = citations $\geq$ T years / total citations	Set threshold T (like 15 years), count proportion of references	High deep citation ratio often indicates “cross-paradigm integration” tendency
Z-axis Energy Threshold	E_{th} = minimum energy for first significant SCJ	Record energy consumption and SCJ across multiple iterations, find first SCJ >	Empirical “resource $\rightarrow$ leap” threshold from organizational perspective
Leap Efficiency	$C_Z = \Delta SCJ_{cum} / E_{tot}$ or $C_Z = -\Delta F/E$	Aggregate leap counts or cumulative SCJ across projects/time periods with total energy consumption	Leap efficiency from organizational perspective, supporting investment decision-making

Note: Such indicators usually require long time-span evolution trajectory data and energy consumption records during model operation. At the implementation level, they can be collaboratively extracted using citation graphs, model logging systems, knowledge structure visualization tools (like CiteSpace), and computing power measurement mechanisms to support automated monitoring and longitudinal comparison of leap creative behavior.

6.4 Comprehensive Creativity Evaluation and Visualization

To achieve comprehensive presentation and dynamic tracking of the three-dimensional creativity measurement system, structured evaluation spaces and graphical representation methods need to be constructed after indicator extraction. This paper proposes three key visualization strategies:

1. **Three-Dimensional Coordinate Projection (3D Creativity Embedding)**
   Normalize scores of each analysis subject (can be individual, human team, AI system, or hybrid collaborative unit) in three dimensions X (Interpo-

lation Depth), Y (Extrapolation Breadth), and Z (Leap Height) (such as Z-score standardization, principal component analysis, etc.), and embed them into three-dimensional coordinate space. Combined with energy consumption constraints, different creative activities' creativity characteristics can be positioned in this space to identify specialized creativity expression methods for different task types (as shown in Figure 1 [Figure 1: see original paper]).

Figure 1 Three-dimensional projection of multiple creative activities

2. **Energy-Creativity Contour Mapping**
Construct a two-dimensional area on the E-Z plane with energy input (E) as the horizontal axis and leap height (Z) as the vertical axis, and overlay heatmaps or contour maps of X and Y dimensions on it. This map can reveal the collaborative performance ability of different systems in interpolation and extrapolation dimensions under different energy budgets, thereby identifying "multi-dimensional benefit maximization zones" (example shown in Figure 2 [Figure 2: see original paper]).

Figure 2 Energy-creativity contour map
Note: Different colors 1-5 in the legend represent heat values of X-axis or Y-axis on E-Z plane coordinates.

3. **Temporal Trajectory Analysis (Creativity Dynamics Tracking)**
For subjects with continuous operation or long-term accumulation characteristics (such as large-scale language models, research institutions, innovative enterprises, etc.), time series trajectories (X_t, Y_t, Z_t) can be constructed to observe the evolution paths of their creativity structures at different time points. By identifying nonlinear mutations, speed reduction segments, and Z-axis surge points in trajectories, prediction and positioning of "critical slowing," "leap points," and "paradigm replacement periods" can be achieved (as shown in Figure 3 [Figure 3: see original paper]).

Figure 3 Temporal trajectory analysis of creativity expression

The above three visualization methods not only support multi-dimensional analysis of single system creativity, but also facilitate horizontal comparison and strategy optimization at organizational, task, or disciplinary levels.

6.5 Implementation Elements

To ensure consistency, reproducibility, and fairness of cross-subject comparison in practical application of the three-dimensional creativity evaluation system, indicator extraction and calculation processes must satisfy the following key implementation elements:

1. **Data Homogenization**
Before indicator calculation, it must be ensured that output results from different subjects (individuals, human teams, or AI systems) can be embed-

ded into unified measurement spaces. Common strategies include using shared semantic embedding models (like Sentence-BERT), unified structural representations (like code abstract syntax trees AST), or standardized knowledge graph entity embedding, to ensure indicators like combinatorial coverage and diversity are comparable on the same basis.

2. **Energy Accounting Consistency**
   All energy efficiency indicators (like C_E,X, C_Z, etc.) depend on accurate recording and unified conversion of energy consumption. It is recommended to use GPU-hours, kWh, or person-hours as basic units, and convert them into equivalent free energy through standard models to support normalization across dimensions and systems.
3. **Transparent Threshold Specification**
   When indicators involve leap determination, robustness evaluation, or historical citation depth, critical parameters like (performance baseline), T (deep citation year threshold), and δ (controlled drift amplitude) should be clearly set during experimental pre-registration, accompanied by task-specific explanations to prevent reproducibility bias from "post-hoc parameter tuning."
4. **Open-Source Toolchain and Pipeline**
   To promote reproducibility of this measurement framework across different research teams and systems, it is recommended to build full-process automated indicator extraction and visualization pipelines based on mainstream technical platforms (like Python + PyTorch / HuggingFace Transformers + NetworkX), and publicly release configuration files, example scripts, and evaluation results through platforms like GitHub for reproducibility verification and version collaboration.

7 Formal Bridging of Information Theory-Thermodynamics-Cognitive Science

Next, this paper continues to demonstrate how to embed the three-dimensional creativity evaluation system into a unified physics-information-cognitive framework, establishing computable mapping relationships between energy consumption, structural orderliness, and cognitive certainty through three core concepts: entropy, information gain, and free energy.

7.1 Three Equivalent Quantities: -ΔS ΔI -ΔF

In energy-structure-cognitive coupling modeling of creative activities, entropy reduction ($-\Delta$S), information gain (ΔI), and free energy decrease ($-\Delta$F) come from thermodynamics, information theory, and cognitive science frameworks respectively. Under certain assumptions, they can be interconverted to uniformly characterize cognitive systems' "order generation," "information gain," and "uncertainty reduction" during creative processes. Table 10 summarizes the definitions, interchange conditions, and cognitive meanings of the three:

Table 10 Three Equivalent Relationships

Concept	Definition	Interchange Condition	Cognitive Meaning
Entropy Reduction	−ΔS = −k_B ln P (macro probability P contraction)	System near equilibrium, state space enumerable	Cognitive process converges to few high-probability explanations by eliminating uncertainty
Information Gain	ΔI = H_{prior} - H_{posterior}	Using same encoding scheme; Bayes update converges	Prediction error absorbed by explanatory model, knowledge base entropy reduction
Free Energy Decrease	−ΔF = F(q) - F(p)	Under variational free energy framework, q can sufficiently approximate true posterior distribution p	In active inference, "perception-action" minimizes expected surprise

When the following conditions are simultaneously satisfied: (i) constant temperature; (ii) observation noise can be modeled as Gaussian distribution; (iii) system internal representation can be stably mapped to external states, equivalence conversion relationships can be established between the three (Formula 19):

$$-\Delta S \equiv \Delta I \equiv -\Delta F$$

That is, entropy reduction is equivalent to information gain is equivalent to free energy dissipation; all three can be uniformly regarded as equivalent representations of cognitive systems "reducing uncertainty." This conversion relationship is the computational basis for normalization and cross-subject comparison of creativity indicators across X-Y-Z dimensions.

7.2 C (Creativity) Index Family

To make the three-dimensional structural measurement of creativity more operable and comparable, this paper introduces a set of efficiency indicators under the information-energy unified framework, i.e., the **Creativity Efficiency Index Family** (C index family).

Under assumptions of constant energy, modelable Gaussian noise, and existing structural mapping, information gain (ΔI), entropy reduction (−ΔS), and free energy dissipation ($-\Delta$F) can be equivalently interchanged. This equivalence provides a unified measurement foundation for "structural compression-prediction optimization-energy cost" in cognitive processes, offering a physics-information theory bridge for cross-system creativity indicator normalization. The C index family (C_X, C_Y, C_Z, C_{tot}) measures creative output efficiency through energy normalization, suitable for horizontal comparison and dynamic evaluation among multiple subjects and tasks.

In the leap dimension, indicators like critical slowing and fluctuation amplification in synergetics theory can be applied to dynamic feature modeling before structural mutations, having potential early warning significance.

1. **Single-Axis Efficiency Index**
   To quantify "unit energy creativity" of individuals or systems in each dimension, the following single-axis efficiency indicators are introduced (Formula 20):

   $$C_X = \frac{S_X}{E_X}, \quad C_Y = \frac{S_Y}{E_Y}, \quad C_Z = \frac{S_Z}{E_Z}$$

   where S_X, S_Y, S_Z represent creative output scores achieved in interpolation, extrapolation, and leap dimensions respectively (like indicators listed in previous sections); while E_X, E_Y, E_Z correspond to equivalent energy consumption in the three dimensions, including resources like GPU-hours, brain metabolism indicators, and experimental investment funds. The core meaning of this indicator family is to measure "creative output per unit energy consumption."

2. **Overall Creativity Index**
   Further, three-axis scores can be merged with weights to form a unified overall creativity indicator (Formula 21):

   $$C_{\text{tot}} = \alpha C_X + \beta C_Y + \gamma C_Z$$

   where weight coefficients (α, β, γ) are calculated as: absolute value of creative output in corresponding dimension divided by absolute value of total creative output, e.g., α = C_{Ax} / (C_{Ax} + C_{Ay} + C_{Az}); β = C_{Ay} / (C_{Ax} + C_{Ay} + C_{Az}); γ = C_{Az} / (C_{Ax} + C_{Ay} + C_{Az}). For basic science exploration and theoretical paradigm reconstruction tasks, this may manifest as γ $\beta \geq \alpha$, because Z-axis leaps dominate cognitive innovation; for engineering system integration or product prototype design tasks, this may manifest as α β $> \gamma$, because tasks emphasize internal reorganization and extrapolation abilities; for large-scale model training and multi-task generalization ability testing, this may manifest as $\beta \geq \alpha$ γ, highlighting extrapolation boundary transcendence ability on Y-axis.

3. **Calibration and Comparison**
   To support horizontal comparison and temporal tracking across multiple subjects and tasks, this paper proposes the following standardization and confidence evaluation methods:

   (1) **Same-task benchmarking**: Set the historically optimal C_{tot} performance for the task as 1.0, and relatively process other subjects' scores using it as the unit.

   (2) **Cross-task transfer**: Use Z-score normalized values of each dimension indicator in task-specific spaces, then aggregate according to task weighting to ensure comparability across sub-tasks.

   (3) **Uncertainty modeling**: Introduce perturbation models for each dimension' s score based on Monte Carlo repeated sampling (or bootstrap), calculate 95% confidence intervals for the total index to provide risk boundaries for strategy judgment or decision-making.

Through the above index system, "creativity" can be modeled and cross-subject or cross-domain comparative evaluation can be made, predictions can be made in diachronic analysis, and empirical foundations can be provided for collaborative optimization of human cognitive behavior and artificial systems.

7.3 Threshold-Emergence "Foreshocks" and Leaps

With the help of synergetics and nonlinear dynamics theory, we can regard structural creative leaps as a "phase transition" process triggered by control parameters, i.e., the system enters a new cognitive order state through self-organization. By identifying precursor features of phase transitions, we may establish predictive signals for leaps before they occur.

1. **Typical "Foreshock" Signals**
   As shown in Table 11 , system evolution before leaps is often accompanied by four key dynamic features that can serve as early warning signals. These signals reflect that when the system approaches some critical state, its internal regulation mechanisms experience hysteresis and instability, which are important signs that cognitive systems or AI models are "about to reconstruct structure."

   Table 11 Four Key Dynamic Features Accompanying System Evolution Before Leaps

Dynamic Feature	Observable Indicator	Early Prediction Window
Critical Slowing Down	Recovery rate $\lambda \to 0$, perturbations require longer recovery	10^1-10^2 iterations before AC(Collaborators) > 0.8

Dynamic Feature	Observable Indicator	Early Prediction Window
Variance Amplification	State fluctuation increases, variance σ^2 expands, gradient norm surges	Concurrent with critical slowing
Mode Collapse	Multi-peak coexistence gradually collapses, latent coding structure/cluster count decreases	Dominated by main peak

2. **Synergetics Modeling and Critical Distance Estimation**
   In the synergetics framework, the leap process can be formalized as a coupled dynamic system of control parameter g(t) (like learning rate, exploration temperature, external support) and order parameter m(t) (like SCJ, C_Z) (Formula 22):

$$\frac{dm}{dt} = a(g - g_c)m - bm^3 + \xi(t)$$

   where g_c represents the system' s mutation bifurcation point; (t) represents perturbation or noise terms. The key to system leap lies in judging the relative distance |g - g_c| between current g(t) and critical point g_c. This model suggests we can estimate parameters like a, b, to judge whether the system is near the critical point, and combine the aforementioned critical slowing signals to more accurately estimate the distance from the current state to the leap critical point.

3. **Implementation Strategy and Pipeline Configuration Recommendations**
   To apply leap identification theory to AI system training or cognitive behavior observation practice, we recommend building a complete online monitoring process: based on streaming frameworks (like Apache Kafka) to dynamically record key indicators, including three-dimensional creativity indices (C_X, C_Y, C_Z), free energy, unit energy consumption, and error changes; using sliding window technology to calculate AC(Collaborators), volatility, and gradient norm to characterize system variables.

   For critical state identification, we recommend using dual-indicator linkage as trigger conditions for prediction: first, when AC(Collaborators) rises and remains above 0.8, indicating the system' s recovery ability to perturbations is weakening; second, if the variance growth rate of indicator fluctuations exceeds five times the historical benchmark standard deviation (5σ), it indicates the system is in an unstable interval of dynamic fluctuation amplification. Once either condition is met, high-frequency recording of structural logs, snapshot archiving of model states, and safe regulation of key hyperparameters should be initiated immediately. It is also recommended to build a visual "leap proximity" evaluation system to provide real-time intervention basis for decision-makers.

4. **Intervention Strategies for Pre-Leap "Foreshocks"**
Let' s use an analogy framework of physical systems to illustrate dynamic characteristics of pre-leap states.

First, the critical slowing phenomenon is similar to the process of spring relaxation releasing potential energy. When spring elasticity is normal, it can quickly return to original state after appropriate compression or stretching; but when spring stress exceeds elastic limits, spring rebound is destroyed. Similarly, when system structure is in relatively stable state, slight perturbations can be quickly recovered; but near critical points, system recovery after perturbations becomes slow, manifesting as prolonged state stabilization time and decreased system "rebound force." These changes are reflected in data as significantly increased first-order autocorrelation coefficients (like AC(Collaborators) > 0.8), reflecting enhanced system inertia and weakened regulation ability.

Second, fluctuation amplification corresponds to situations where macro fluctuations suddenly increase. Like water surface transforming from slight ripples to violent rolling waves, indicator fluctuation variance or gradient norm rapidly amplifies in short time, indicating enhanced instability of system dynamics.

Finally, the appearance of power-effect inflection points indicates the system has approached functional bottlenecks. Even with more computing power invested (like GPU hours or neural metabolism resources), performance indicators no longer improve or even decline. This transformation can be judged through derivative changes of function against energy consumption (Err/ E changing from negative to positive), marking that original structures are no longer suitable for new tasks and require higher-level reorganization or leaps to solve functional bottlenecks.

When systems approach leap critical points, intervention strategies vary according to specific task objectives. If research or engineering goals aim to stimulate structural changes or promote paradigm shifts, exploration can be enhanced by increasing heterogeneity and structural diversity of input data, introducing more complex or unusual information stimuli. Additionally, exploration temperature can be appropriately increased (like increasing learning rate, introducing input perturbations or parameter mutations) to improve system response sensitivity to new structures. At the same time, task scenarios with cognitive conflict characteristics can be set to activate potential reorganization.

Conversely, if the system is already in deployment or clinical application stage where stability takes priority over innovation, system convergence should be achieved by limiting degrees of freedom. This includes reducing learning rate to suppress violent parameter updates; freezing high-level representation layers to maintain structural continuity; or actively constraining model variation and energy injection amplitude to avoid performance oscillations caused by disordered activation.

8 Practical Implications: Blueprint for a New Science of Innovation

This paper proposes the "entropy reduction-information gain-certainty" creativity model and the XYZ measurement framework for creativity. This is not only a theoretical model, but also an action blueprint guiding practice. It has profound implications for human-machine collaboration models, reflection and reform of education systems, and research strategies.

8.1 New Cognitive Division of Labor: Human-Machine Symbiosis and Extended Mind

The XYZ model clearly reveals the relative advantages of humans and current AI, foreshadowing a new cognitive division-of-labor cooperation model: AI is responsible for innovation on X-axis and part of Y-axis. With its powerful computing power and massive data processing capabilities, AI systems can be positioned as efficient "combination and search engines," responsible for exploring vast possibility spaces, generating candidate solutions, and completing intensive computation and data analysis. Humans focus on Z-axis creativity, concentrating limited cognitive resources on tasks like conceptual leaps, problem reconstruction, value judgment, and ethical supervision.

This division-of-labor cooperation model goes beyond simple "human-machine interaction" and enters the stage of "human-machine symbiosis." We can draw on philosophers Andy Clark and David Chalmers' "Extended Mind" theory to understand this relationship. In the extended mind model, AI is not just an external tool, but an "external cognitive module" deeply integrated with human cognitive processes.

For example, the "Zidong Taichu" (ScienceOne) scientific large model launched by the Chinese Academy of Sciences is designed with the concept of using AI as a research assistant to automatically handle heavy X-axis tasks like literature review, data analysis, and simulation computation, thereby liberating scientists from repetitive labor to focus on proposing original hypotheses (Z-axis activities). Top research teams from Tsinghua University, Peking University, and other institutions are exploring "human-in-the-loop" and human-AI teaming projects. The innovations of these studies in theory and practice share the common point of not treating AI as a passive tool, but as a "teammate" with certain agency, emphasizing the establishment of shared cognitive models between humans and AI and achieving value alignment.

In addition, industrial ecological layout characteristics show that domestic tech giants like Huawei, Alibaba, Tencent, and Baidu are striving to build powerful AI infrastructure (providing computing power for X-axis) while also investing in basic research, forming an industrial ecology that supports this new cognitive division-of-labor cooperation model.

8.2 Educational Reform' s "Z-axis Turn"

Creativity development is one of the important tasks of 21st-century human development. The most valuable manifestation of human creativity should be reflected on the Z-axis. In view of this, current education systems and practices must undergo profound reflection and transformation. Education models that overemphasize standardized answers and knowledge memory essentially train X-axis abilities, which are precisely the areas most easily replaced by AI. Therefore, the core of educational innovation reform should focus on the "Z-axis turn."

Educational goals should shift from knowledge transmission to cultivation of core competencies, especially reflective abstraction, structured extrapolation, cross-domain integration, and ability to pose new questions. This requires curriculum design to shift from "problem solving" to "problem construction," encouraging students to define problems, introduce new variables, and construct new explanatory frameworks when facing complex, uncertain situations.

The "Z-axis turn" is consistent with the strategic deployment of China' s current education reform. In recent years, the Chinese government has vigorously advocated the transformation from "exam-oriented education" to "quality education," and in 2016 released the "Framework for Chinese Students' Development of Core Competencies," explicitly proposing to cultivate students' "critical thinking" and "innovation ability." The national "New Engineering" construction plan also emphasizes cultivating composite innovative talents who can respond to future challenges. Our theoretical framework provides specific cognitive scientific foundations for these macro policy objectives and proposes operable and evaluable implementation paths.

8.3 Structural Optimization of Research Funding

The research funding system is the "command baton" guiding scientific and technological innovation. Traditional funding models often swing between two extremes: either preferring incremental research with high success rates and low risks (X-axis and Y-axis), or vaguely calling for "disruptive innovation" without effective identification mechanisms.

Our creativity framework suggests adopting a more refined investment evaluation strategy. On one hand, stable support is needed for "exploitative" research led by large teams aimed at consolidating and expanding existing paradigms; on the other hand, a specific proportion of resources must be allocated to support exploratory research conducted by small, flexible teams with high risk and high potential disruptiveness (high Z-axis potential). Scientometric research has already shown that small teams are more efficient than large teams in producing disruptive results.

China' s research funding system reform provides a highly valuable case study. While maintaining conventional funding channels like general and key projects (supporting X-axis and Y-axis), the NSFC has established the "Original Exploration Program" in recent years. The official guidelines of this program explicitly state its goal is to support high-risk frontier exploration with "disruptive and non-consensus characteristics." This can be seen as an institutionalized funding tool specifically designed for Z-axis activities.

Comprehensively, China' s strategic layout in three key areas—AI infrastructure construction, education reform, and research funding—can be understood as a large-scale practice at the national level that coincides with the human-machine collaborative innovation model with Z-axis as the core proposed in this paper. This shows that our theoretical framework not only has philosophical and cognitive scientific rationality, but also reflects the internal logic and evolution direction of national-level innovation strategies in the AI era.

9 Possible Objections and Responses

Any new theoretical framework must face scrutiny and challenges. This paper preemptively responds to four possible objections to further clarify the boundaries and applicability of this model.

Objection 1: The traditional definition of creativity as "novelty × effectiveness" is sufficient; there is no need to propose new creativity definitions and evaluation models.

Response: This traditional definition was effective before the AI era, but now has exposed insufficient discriminative power. It ignores two fundamental physical constraints: energy cost and information structure. Because it cannot distinguish between two fundamentally different types of "novelty" : one is "combinatorial novelty" generated by consuming massive energy and performing high-dimensional interpolation in existing data spaces (like AI-generated massive images); the other is "structural novelty" achieved through low-energy cognitive reorganization and invention of completely new conceptual frameworks (like Einstein' s theory of relativity). Our model compensates for the insufficient discriminative power of traditional definitions in key domains by introducing core constructs like the C index ($\Delta S/E$) and Z-axis (structural leap).

Objection 2: In open cognitive systems, entropy reduction (ΔS) is difficult to define and measure precisely.

Response: Directly measuring total entropy change of an open system is indeed difficult. However, we can adopt a series of effective proxy variables for operational measurement. According to the free energy principle, entropy reduction is mathematically equivalent to reduction in prediction error (or surprise) and decrease in variational free energy. Therefore, we can quantify ΔS by tracking the rate of prediction accuracy improvement in a cognitive agent' s (human or AI) predictive model during learning, or the magnitude of surprise reduction when facing new data. At the information theory level, ΔS can also be measured through Minimum Description Length (MDL) compression gains. These are all operational measurement methods already widely applied in modern computational neuroscience and machine learning.

Objection 3: Analogizing cognitive leaps to "electron energy level transitions" is far-fetched.

Response: We clearly state that this analogy is not an ontological equivalence, but a dynamical heuristic analogy. Its core value lies in capturing several key dynamical features of insight processes: (a) **threshold nature**: Insight occurrence usually requires accumulation of cognitive conflict or anomalous evidence to a certain degree; (b) **nonlinearity/mutation**: State transitions are rapid and sudden, not smooth linear processes; (c) **rate dependence**: The rate of evidence accumulation affects leap probability. Phase transition and bifurcation models in complex systems theory universally exhibit these common dynamical structures. The electron energy level model simply uses a concise, well-known physical image to characterize this universal nonlinear dynamical feature, thereby providing an intuitive starting point for building testable mathematical models and experimental paradigms.

Objection 4: Energy consumption (E) measurement across different substrates (biological brain vs. silicon chip) is not comparable.

Response: Directly comparing joules of energy consumption is indeed challenging. Therefore, our theoretical framework emphasizes relative efficiency and normalized comparison. We compare not absolute energy consumption, but entropy reduction benefits per unit energy consumption (i.e., improvement in C index). When making cross-system comparisons, multiple normalization strategies can be adopted: for example, converting both AI' s energy consumption (electrical) and brain' s energy consumption (metabolic) into standard thermodynamic units or equivalent economic costs. More importantly, we focus on the functional relationship between energy consumption and creativity output. For example, we can examine how much output on X, Y, Z axes increases for different systems when energy consumption increases by an order of magnitude. This analysis of "scaling laws" can reveal fundamental differences in creativity generation mechanisms between different systems without requiring precise equivalence of absolute energy consumption values.

10 Conclusion: Reorienting Human Meaning with Creativity as Anchor

The rise of AI has not diminished the value of human creativity, but rather forced us to examine creativity more profoundly and essentially. The core argument of this paper is that the misalignment between the "high-score creativity" currently exhibited by AI and historic paradigm breakthroughs of humanity stems from our 沿用至今的创造力定义未能充分考虑创造性活动的物理学和信息论约束。

By defining creativity as entropy reduction efficiency per unit energy consumption (C index) and constructing a three-dimensional evaluation model comprising "Interpolation Depth (X-axis), Extrapolative Breadth (Y-axis), and Leap Height (Z-axis)," we find that AI creativity originates from high-energy-consuming, large-scale combinatorial computing power on the X-axis, while the unique core of human creativity lies in achieving leaps on the Z-axis with extremely low energy consumption, i.e., completing fundamental reorganization and unification of knowledge structures through reflective abstraction and cognitive phase transitions.

The new framework is grounded in bottom-level constraints from physics, information theory, and cognitive science (energy conservation, information entropy, certainty). It uses leap dynamics from synergetics and complex systems theory as a bridge to explain and "gain insight into" developmental characteristics of creative mutational processes. It also uses Kuhn' s paradigm theory and Piaget' s genetic epistemology as frameworks to clarify cognitive generation paths for new knowledge structures "from nothing to something."

In the new era of human-machine symbiosis, human-machine competition is not a good strategy; forming cognitive division-of-labor cooperation alliances based on respective relative advantages is the ideal model for human-AI coexistence. AI can become a powerful X-axis and Y-axis engine, processing massive information combination and optimization; while humans focus on the Z-axis, becoming value setters, problem discoverers, and ultimate meaning interpreters.

Future work will 致力于将这个理论模型置于实证研究中做可重复性检验，调节参数以实现模型的精细化修订。这包括但不限于：构建标准化的 XYZ 三维创造力基准测试集和公开数据集；确立可操作性指标（如破坏度指数、结构性压缩增益）的计算方法；在神经科学、认知心理学和人工智能领域开展跨学科的实验验证。

Finally, we hope that through the definition and value reconstruction of creativity, we can compare and distinguish creative activities of humans and AI under a unified theoretical framework; through mutual illumination of humans and AI, we can reflect on ourselves and reconsider the fundamental meaning of human mind and consciousness.

References

Darwin. (2005). *On the Origin of Species* (Trans. by Zhou Jianren, Ye Duzhuang, Fang Zongxi). Beijing: The Commercial Press.

Jiang, K. (2018-02-27). Whose cheese did AlphaZero move? *Chinese Social Sciences Today*, (006).

Kuhn. (2004). *The Structure of Scientific Revolutions* (Trans. by Jin Wulun, Hu Xinhe). Beijing: Peking University Press.

Russell. (1948/2001). *Human Knowledge*. Beijing: The Commercial Press.

Acar, S. (2023). A reliability generalization of the TTCT-Figural & the fluency confound. *Psychological Test and Assessment Modeling*.

Amabile, T. M. (1983). *The social psychology of creativity*. Springer.

Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman & Hall.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, 21(10), 1133-1145.

Baars, B. J. (1993). *A cognitive theory of consciousness*. Cambridge University Press.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645.

Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine learning practice and the classical bias–variance trade-off. *Proceedings of the National Academy of Sciences*, 116(32), 15849-15854.

Bender, E. M., & Koller, A. (2020, July). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 5185-5198).

Bennett, C. H. (1973). Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6), 525-532.

Bisk, Y., Holtzman, A., Thomason, J., et al. (2020). Experience grounds language. *Proceedings of EMNLP*.

Brillouin, L. (1953). The negentropy principle of information. *Journal of Applied Physics*, 24(9), 1152-1163.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). The Creative Achievement Questionnaire: A measure of creative achievement and creative potential. *Creativity Research Journal*, 17(1), 37-50.

Chen, Z., & Cao, Y. (2024). A polishing model for machine-generated ancient Chinese poetry. *Neural Processing Letters*, 56(2), 77.

Clark, A. (2015). Embodied prediction. In *Open mind*. Open MIND. Frankfurt am Main: MIND Group.

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

Cohen, I. B. (1985). *Revolution in science*. Harvard University Press.

Cropley, D. (2023). *Creativity and artificial intelligence*. Springer.

Diaz, J. A., Nelson, S. M., Beaujean, A. A., Green, A. E., & Scullin, M. K. (2024). The impact of adding a fourth item to the traditional three-word Remote Associates Test. *Creativity Research Journal*.

Dollinger, S. J., & Shafran, M. (2005). Note on consensual assessment technique in creativity research. *Perceptual and Motor Skills*, 100(3), 592-598.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127-138.

Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology*, 37(8), 1398-1405.

Hadas, E., & Hershkovitz, A. (2025). Assessing creativity across multi-step intervention using generative AI models. *Journal of Learning Analytics*, 12(1), 15-32.

Haken, H. (1983). *Synergetics: An introduction*. Springer.

Hongdizi, J., Cui, K., & Zhou, X. (2023). Influence of analytic processing on divergent and convergent thinking tasks: The role of rational and experiential thinking styles. *Journal of Intelligence*, 11(2), 23.

Janiak, A. (2008). *Newton as philosopher*. Cambridge University Press.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.

Jiang, L., & Zhou, M. (2008, August). Generating Chinese couplets using a statistical MT approach. In *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 377-384).

Kapoor, H., Zheng, M., Reiter-Palmon, R., & Kaufman, J. C. (2023). Toward equitable creativity self-assessment: Measurement invariance and mean differences across ethnicity in the Kaufman Domains of Creativity Scale. *Psychology of Aesthetics, Creativity, and the Arts*.

Kim, K. H. (2004). The creativity crisis: The decrease in creative thinking scores on the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 16(4), 347-354.

Koestler, A. (1964). *The act of creation*. Hutchinson.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge University Press.

Lake, B., & Baroni, M. (2018). Generalization without systematicity: On the compositional skills of sequence-to-sequence recurrent networks. *Proceedings of ICML*.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183-191.

Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP. *NeurIPS*.

Li, J., Huang, J., Liu, J., & Zheng, T. (2022). Human-AI cooperation: modes and their effects on attitudes. *Telematics and Informatics*, 73, 101862.

Li, M., & Vitányi, P. (2008). *An introduction to Kolmogorov complexity and its applications* (3rd ed.). Springer.

MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. Cambridge University Press.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220-232.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.

Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. In *Psychology of learning and motivation* (Vol. 26, pp. 125-173). Academic Press.

Ogurlu, U., Acar, S., & Ozbey, A. (2023). Does word frequency impact ideational fluency in divergent thinking? A meta-analytic exploration with the Alternate Uses Test. *Thinking Skills and Creativity*, 49, 101139.

Ontañón, S., et al. (2022). An empirical study on compositional generalization of in-context learning. *EMNLP Workshop*.

Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), 138-144.

Piaget, J. (1950). *The psychology of intelligence*. Routledge.

Piaget, J. (1970). *Genetic epistemology*. Columbia University Press.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465-471.

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential: A meta-analytic review. *Creativity Research Journal*, 24(1), 66-75.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96.

Said-Metwaly, S., Fernández-Castro, J., & Kyndt, E. (2018). Measuring creativity: The state of the art. *Thinking Skills and Creativity*, 27, 23-37.

Schrödinger, E. (1944). *What is life? The physical aspect of the living cell*. Cambridge University Press.

Silver, D., Huang, A., Maddison, C. J., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484-489.

Silver, D., Hubert, T., Schrittwieser, J., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), 1140-1144.

Silver, D., Schrittwieser, J., Simonyan, K., et al. (2017). Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv:1712.01815*.

Singh, S., Lyu, M., Deterding, S., et al. (2025). A systematic review of human-AI co-creativity (Tech. Rep. arXiv:2506.21333). ACM.

Torrance, E. P. (1974). *Torrance Tests of Creative Thinking: Norms-technical manual*. Personnel Press.

Warne, R. T. (2022). An evaluation (and defense) of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, 34(2), 111-127.

Wu, C.-L., Huang, S.-Y., Chen, P.-Z., & Chen, H.-C. (2020). A systematic review of creativity-related studies applying the Remote Associates Test from 2000 to 2019. *Frontiers in Psychology*, 11, 573432.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378-382.

Yi, X., Li, R., & Sun, M. (2018). Chinese poetry generation with a salient-clue mechanism. *arXiv preprint arXiv:1809.04313*.

Zhang, S., Zhong, L., Zhang, Y., Ren, J., & Zhang, J. (2023). Development and validation of the Kaufman Domains of Creativity Scale–Chinese short version (K-DOCS-CSV). *Thinking Skills and Creativity*, 49, 101268.

Appendix: Symbol and Abbreviation 对照表

Symbol/Abbreviation	Definition	Typical Unit/Dimension	Notes/Usage
S, ΔS	Entropy, entropy change	bits or dimensionless (k_B=1)	Measure of disorder; change from “disorder” to “order”

Symbol/Abbreviation	Definition	Typical Unit/Dimension	Notes/Usage
I, ΔI	Information, information gain	bits or dimensionless	Difference between prior and posterior entropy; information absorption
F, ΔF	Free energy, free energy change	J or equivalent to kT	Measures "surprise" in variational free energy framework
k_B	Boltzmann constant	$J \cdot K^{-1}$	Often set to 1 for dimensionless treatment in theory
Vol(·)	Convex hull volume operator	$(\text{unit})^3$	Used for volume calculation in combinatorial coverage C_{cov}
H(S), H(U)	Candidate solution embedding set, baseline universe set	-	Used for Vol(H(S))/Vol(H(U)) in X-axis space coverage/comparison
C_{cov}	Combinatorial coverage	-	Vol(H(S))/Vol(H(U))
D_{div}	Diversity/decorrelation	-	Mean embedding distance or JSD/CLIP diversity
L_{base}, L_{model}	Baseline/current model description length	bits	Required for MDL calculation
ΔL_{MDL}	MDL compression gain	-	(L_{base} - L_{model})/L_{base}

Symbol/Abbreviation	Definition	Typical Unit/Dimension	Notes/Usage
Q_{out}	Effective output quality	-	Such as BLEU, ROUGE, music/design scores
T	Observation/computation duration	s	Can be converted to unified free energy equivalent
C_E,X	Energy-normalized score (X)	-	Q_{out}/(E_{comp}/T)
δ	Distribution drift scale	Same dimension as δ	Used for extrapolation radius R_{OOD}
S_{str}	Causal/structural migration score	-	(1/N)Σ1[f_{src} f_{tgt}]
R_{cf}	Counterfactual robustness	0-1 (smaller is more stable)	1 - Perf_{cf}/Perf_{orig}
k_E,Y	Energy-performance curve slope	Performance/Energy	Perf_{OOD}/ E
L_{old}, L_{new}	Description length before/after leap	bits	Used for SCJ calculation
SCJ	Structural compression leap gain	-	(L_{old} - L_{new})/E
F(t)	Time series free energy	J	Used for characterizing surprise over time
v_{FE}	Surprise collapse rate	$J \cdot s^{-1}$	- F(t)/ t/Δt
D	Disruption Index	-	(N_{later} - N_{prior})/N_{total}

Symbol/Abbreviation	Definition	Typical Unit/Dimension	Notes/Usage
R_{deep}	Deep citation depth	-	Proportion of references $\geq$ T years old
E_{th}	Z-axis energy threshold	J	Minimum energy for first significant SCJ
C_X, C_Y, C_Z	Single-axis efficiency indices	-	Output efficiency per unit free energy
α, β, γ	Weight coefficients	-	$\alpha+\beta+\gamma=1$; set according to tasks
g, g_c	Control parameter, critical value	-	Such as learning rate, exploration temperature, funding
(t)	Random perturbation term	-	In bifurcation equation
AC(1)	Autocorrelation coefficient (lag 1)	-	Critical slowing: λ$→$0; >0.8 often signals critical slowing
T (year threshold)	Deep citation year threshold	years	Fluctuation amplification monitoring; often set to 10-20 years

Contact Author:
Ke Jiang, Professor, School of Teacher Education, Lishui University
E-mail: jiangke200@126.com
Phone: 13558681429

Note: Figure translations are in progress. See original paper for figures.