Philosophy Study, Jan.-Feb. 2026, Vol. 16, No. 1, 30-44
doi: 10.17265/2159-5313/2026.01.003

A New Proposal for Scientific Standards: Modified Hilbert Axiomatic Standards and Dynamic Demarcation Criteria

CUI Weicheng
Zhejiang Engineering Research Center of Micro/Nano-Photonic/Electronic System Integration, Hangzhou, China
Westlake University, Hangzhou, China

Defining science and demarcating it from pseudoscience are longstanding core issues in the philosophy of science. Hilbert's traditional axiomatic standards (consistency, completeness, independence) struggle with modern complex systems, while existing demarcation criteria like replicability and Popper's falsificationism have practical limitations. Analyzing incompatibilities in complex systems, this paper proposes a revised framework: modified Hilbert axiomatic standards (clearly defined concepts, logical consistency, unrefuted axioms) and dynamic demarcation criteria. Abandoning unattainable traditional requirements of completeness and independence, it emphasizes conceptual clarity, openness, and progressiveness, offering an operational, self-consistent basis for judging scientificity.

Keywords: scientific standards, axiomatic system, demarcation problem, logical consistency, dynamic criteria

Introduction

We live in an era dominated by science, where the scientific nature of research directly affects the recognition and support it receives from society (Walach, 2019). However, there is still no universally accepted definition of science or a unified criterion for demarcating science from pseudoscience (Blachowicz, 2009; Johansson, 2016). This "demarcation problem" proposed by Popper (1963) has long plagued philosophers of science and scientists alike.

The definition of science is the premise for constructing scientific standards. Existing definitions from dictionaries, academic institutions, and Wikipedia only cover partial connotations of science and fail to fully reflect its core characteristics (e.g., the finiteness of research scope, the rigor of logical structure, the clear definitions of important concepts). In terms of scientific criteria, Hilbert's axiomatic standards (consistency, completeness, independence) have been regarded as the classic framework for evaluating the rigor of mathematical and scientific theories since his proposal (Hilbert, 1902). However, Gödel's incompleteness theorems demonstrated the existence of theories that are inherently incomplete, thereby challenging the necessity of imposing completeness as a mandatory requirement for scientific theories (Gödel, 1931). In terms of the question whether quantum mechanics is a complete theory or not, Einstein, Bohr, and later physicists had a long debate and no agreement has reached yet (Einstein, Podolsky, & Rosen, 1935; Bohr, 1935; Whitaker, 2006). With the development of disciplines such as physics and cosmology, especially the in-depth discussion of the Theory of Everything (TOE) and Meta-Theory of Everything (MToE), the limitations of traditional standards have

CUI Weicheng, Ph.D., Professor, Zhejiang Engineering Research Center of Micro/Nano-Photonic/Electronic System Integration; Department of Electronic and Information Engineering, School of Engineering, Westlake University, Hangzhou, China.

become increasingly prominent. Faizal, Krauss, Shabir, and Marino (2025) proposed the MToE aiming for a "complete and consistent" effective theory, but Cui (2026) pointed out that completeness and consistency are sometimes mutually exclusive, making the pursuit of completeness unrealistic for MToE. Da Costa (2011) further reinforced this view by showing that modern physical theories often overstate their ability to achieve Hilbert's traditional completeness requirement, and prioritizing pragmatic consistency (alignment with empirical evidence) over absolute completeness is a more viable path for scientific progress.

In addition, existing demarcation criteria have their own shortcomings: Popper's falsificationism faces difficulties in dealing with "relative falsifiability" of statements (Hansson, 2006); replicability is criticized for being inapplicable to many complex scientific fields (Braude, 2018; Frigg, 2023); Ellis's (2014) systematic criteria are too complex and lack operability; Maudlin's (2018) canonical presentation requirements focus more on the form of theory expression than on the core criteria of scientificity.

To address these issues, this paper first redefines science based on integrating existing definitions and then modifies Hilbert's axiomatic standards according to the characteristics of complex open systems. On this basis, a dynamic demarcation criterion for scientific theories is constructed, aiming to provide a new, practical, and logically rigorous framework for scientific standards.

Re-defining Science: Core Connotations and Boundaries

The ambiguity of the definition of science is an important reason for the difficulty in formulating scientific standards. By sorting out and integrating existing definitions, the present author proposes the following definition: Science is a set of clearly defined and logically consistent knowledge about the structure and behavior of natural and social systems. It is obtained through observation, measurement, and experimentation, and is presented in the form of testable explanations and predictions for the observable systems within the world we live in.

This definition emphasizes three core connotations:

First, the finiteness of the research scope. Science can only study observable systems, which are finite in time and space. This means that any claim of a "Theory of Everything" covering the entire universe (which is beyond human observation and untestable) is beyond the scope of science (Weinberg, 2011; Holman & Wilholt, 2022). The universe's infinity and openness make it impossible for human beings to obtain complete knowledge about it, so science must abandon the illusory pursuit of "universal completeness".

Second, the clarity of concepts. Drawing on Hua et al. (2024), "clear definition" is explicitly defined as "operationalizable, consistent across the theory, and distinguishable from competing definitions in the literature". A concept is clear if it can be used to derive testable predictions without internal contradiction—a heuristic that aligns with Maudlin's (2018) "canonical presentation" requirement for ontological clarity. All core concepts, research objects, and scope of a scientific theory must meet this standard of clarity. However, nowadays, the fundamental concepts of universe, world, time, space, matter, consciousness, mass, energy, information, entropy are found to be defined very differently in different theories, which makes communication and unification of different theories hard (Cui, Li, & Pan, 2024; 2025). Undefined concepts (i.e., concepts universally recognized in dictionaries) should also be reasonably selected to avoid ambiguity and confusion in theoretical reasoning (Maudlin, 2018), and these primitive terms must be grounded in shared empirical intuitions to avoid circularity (Hua et al., 2024; Cui, Li, Pan, & Zeng, 2025). It is the requirement of logical consistency that forces us to give up the completeness of defining all concepts. Let us take a dictionary as an example to illustrate the mutually exclusive nature of completeness and logical consistency in concept definitions in a theory. When defining Concept A, one might inevitably resort to the use of Concepts B and C. If the dictionary's editors aim to meet the requirement of completeness by claiming that "all concepts in this dictionary have explicit definitions", then defining Concept Z will inevitably involve reliance on previously defined concepts, resulting in logical circularity. So to maintain logical consistency in concept definitions, certain concepts must be left undefined and treated as primitive terms—terms that are universally understood and require no further explanation. For example, Newton made such a mistake in the definitions of mass and density (Newton, 1846). He defined mass = density × volume and density = mass / volume. It is found that this problem of logical self-circulation is still very popular in modern scientific theories (Cui, Li, Pan, & Zeng, 2025). For contested concepts like "consciousness" in cognitive science, operationalizability becomes the key benchmark: A definition of consciousness is scientifically valid if it links the concept to measurable neural correlates or behavioral outcomes, even if debates about its ultimate nature persist (Hua et al., 2024). In contrast, pseudoscientific concepts like "qi" lack such operationalization and consistency, rendering them unscientific.

Third, the rigor of logical structure. Scientific knowledge must be logically consistent, and there should be no contradictions between axioms, laws, and derived conclusions. This is the fundamental guarantee for the reliability of scientific theories.

This definition clarifies the boundary of science: It is not an attempt to pursue absolute truth or complete explanation of the universe, but a systematic exploration of observable systems based on empirical evidence and logical reasoning. This boundary setting lays the foundation for the subsequent revision of scientific standards.

Revision of Hilbert's Axiomatic Standards: From Traditional to Modified Framework

Limitations of Traditional Hilbert Axiomatic Standards

Hilbert (1902) proposed three core requirements for axiomatic systems: consistency (no contradictory propositions can be derived from the axioms), completeness (all true propositions within the system can be derived from the axioms), and logical independence (no axiom can be derived from other axioms). These standards have achieved remarkable success in simple mathematical systems, but they face insurmountable challenges when applied to complex scientific theories, especially in the field of modern physics.

First, the incompatibility between completeness and consistency. Gödel's incompleteness theorem (Gödel, 1931) has proved that for any sufficiently complex axiomatic system, if it is consistent, it must be incomplete—there exists at least one true proposition within the system that cannot be derived from the axioms (Robertson, 2000). For open and complex natural systems (such as the universe, ecological systems, etc.), human observation is limited, and the axioms of scientific theories are inductively derived from finite empirical evidence. This makes it impossible to guarantee the completeness of the axiom system, and forcing the pursuit of completeness will inevitably conflict with logical consistency (Cui, 2021). Da Costa (2011) supported this through their analysis of quantum field theory, which maintains scientific rigor despite its axiomatic incompleteness by prioritizing pragmatic consistency with empirical evidence. For example, Faizal et al.'s (2025) pursuit of "complete and consistent" MToE ignores this inherent contradiction, so Cui (2026) advocates abandoning completeness and prioritizing consistency.

Second, the unattainability of logical independence. It is crucial to distinguish Hilbert's "logical independence" (a property where no axiom can be logically derived from others) from the statistical concept of "mutually exclusive" (where the occurrence of one event precludes the other), a conflation that has previously muddied discussions (Giovannini & Schiemer, 2025). For complex scientific theories, the axiom system is constructed based on empirical induction, and the interconnection between axioms is often difficult to completely separate because they draw on overlapping empirical observations (Giovannini & Schiemer, 2025). Giovannini and Schiemer (2025) used climate science as an example: Axioms about atmospheric pressure and temperature are not logically independent, yet the theory remains scientifically rigorous. Moreover, due to the openness of the research system, it is impossible to prove the absolute logical independence of axioms—new observational evidence may reveal potential logical connections between seemingly independent axioms. Therefore, logical independence is not a necessary condition for scientific theories, and "empirical coherence" (alignment between axioms and real-world data) is a better measure of rigor (Giovannini & Schiemer, 2025).

Modified Hilbert Axiomatic Standards

Based on the above analysis, combined with the characteristics of the accuracy of scientific language and the limitations of traditional standards, this paper proposes the modified Hilbert axiomatic standards for scientific theories, including three core criteria:

Clearly defined concepts. All key concepts of a scientific theory must be clearly defined in line with the standard of "operationalizable, consistent across the theory, and distinguishable from competing definitions" (Hua et al., 2024), including research objects, research scope, core concepts, and basic categories. For a general theory for complex systems, these concepts should be included as core concepts: universe, world, system, environment, support, constraints, physical objects, non-physical objects, vacuum, time, space, coordinate system, matter, non-matter, mass, position, momentum, energy, information, consciousness, mind, ether, entropy, closed system, isolated system, and open system. The definitions for these core concepts should avoid ambiguity and ensure that the theoretical framework has a unified understanding basis. For undefined concepts (such as "good" and "bad", "high" and "low" in natural language), they should be reasonable and universally recognized to avoid arbitrary assumptions, and grounded in shared empirical intuitions (Hua et al., 2024). At the same time, the ontology behind the research object and scope should be clearly specified to help readers understand the theoretical foundation (Maudlin, 2018). For example, in quantum mechanics, the definition of "wave function" (operationalized as a mathematical representation that predicts measurement outcomes for microscopic systems) and the scope of its application (microscopic systems) must be clearly defined to avoid misapplication to macroscopic systems (Hua et al., 2024).

To enhance the universality of the "clearly defined concepts" criterion across disciplines, we provide discipline-specific operational guidelines and examples for contested foundational concepts:

- Physics (e.g., "spacetime" in quantum gravity): Operationalization requires linking the concept to measurable physical effects or mathematical formalisms that generate testable predictions. For instance, in loop quantum gravity, "spacetime" is operationalized as a discrete network of quantum spin foams, with observable implications such as corrections to the speed of light at high energies (Rovelli, 2018). Consistency demands that this definition aligns with general relativity in the classical limit (e.g., reproducing Einstein's field equations for large-scale structures). Distinguishability involves clarifying how it differs from competing definitions (e.g., string theory's

11-dimensional spacetime, which predicts different gravitational wave signatures). This operationalization avoids metaphysical ambiguity while retaining predictive power.

- Biology (e.g., "species"): Operationalization adopts either the biological species concept (reproductive isolation, measurable via interbreeding success rates in controlled environments) or the phylogenetic species concept (shared derived traits, quantifiable through DNA sequence similarity thresholds) (De Queiroz, 2007). Consistency requires adhering to one framework within a theory (e.g., not switching between reproductive and morphological criteria without justification). Distinguishability involves explicitly contrasting with alternative concepts (e.g., explaining why reproductive isolation is prioritized over ecological niche overlap in a given taxonomic study). For example, a theory on speciation in birds would operationalize "species" as groups with <1% interbreeding success in the wild, consistent across all studied populations, and distinguishable from morphological species concepts by referencing genetic divergence data.
- Psychology (e.g., "intelligence"): Operationalization relies on standardized, replicable measures such as fluid intelligence (assessed via matrix reasoning tasks) or crystallized intelligence (measured through vocabulary tests), with clear scoring rubrics (Case, 1985). Consistency mandates that the same operational definition is used across studies (e.g., not conflating IQ scores with emotional intelligence without explicit bridging). Distinguishability involves differentiating from competing constructs (e.g., "creativity" or "wisdom") by specifying unique behavioral or neural correlates (e.g., fluid intelligence correlates with prefrontal cortex activation during problem-solving, distinct from the temporal lobe engagement linked to creativity).
- Social science (e.g., "power" in political theory): Operationalization ties the concept to observable actions or resource distributions—e.g., "decision-making power" measured by the frequency of a stakeholder's votes determining policy outcomes, or "structural power" quantified via control over economic resources (Weber, 2002). Consistency requires limiting the scope to a specific type of power (e.g., not conflating coercive power with persuasive power) within a theory. Distinguishability involves contrasting with alternative definitions (e.g., Lukács's "ideological power" vs. Bourdieu's "symbolic capital") and specifying empirical contexts where the operationalization applies (e.g., liberal democratic institutions vs. authoritarian regimes).

Based on the Relativity of Simultaneity Axiom (Cui, 2021), Cui found a pair of Concepts "A" and "Non-A" occur or disappear at the same time since two-valued logic is the minimal logical system for human beings. This implies that monist philosophy such as materialism and idealism is ontologically incomplete. In the Unified Complex System Theory (UCST) proposed by Cui's group (Cui et al., 2024; Cui, Li, & Pan, 2025), ether is defined as the essence of material objects (bodies) while non-matter mind is defined as the essence of a living agent, so a dualist solution to the famous mind-body problem is adopted as the ontology of UCST.

Logical consistency. Logical consistency is the core requirement of scientific theories and the fundamental guarantee for their reliability, aligning with da Costa's (2011) concept of "pragmatic consistency" (here defined as the alignment between theoretical propositions and empirical evidence, where discrepancies prompt theoretical revision rather than dismissal). A scientific theory includes four basic components: axioms, laws, logical analysis methods, and empirical phenomena. Axioms are inductively derived from finite observational phenomena; laws are logically deduced from axioms; empirical phenomena should be explained and predicted by laws without contradiction. There should be no logical paradoxes between axioms, between axioms and laws, or between laws and empirical phenomena. For example, the contradiction between general relativity (describing macroscopic gravity) and quantum mechanics (describing microscopic interactions) indicates that at least one of the two theories has room for revision in terms of logical consistency in the unified scope, which also promotes the research of quantum gravity theory (e.g. Wu, Zhang, & Li, 2024). In Wu's Gravitational Quantum Field Theory (GQFT, Wu et al, 2024), the Einstein Equivalence Principle (EEP) is the key axiom that has been given up, specifically its strong form which assumes the complete equivalence between gravitational and inertial mass in all physical contexts and the universality of free fall without quantum corrections. In social sciences, logical consistency may require consistency within a specific cultural or contextual scope rather than universal consistency (Frigg, 2023). For instance, a theory of economic decision-making may be logically consistent within the context of individualistic cultures but require adjustment for collectivist cultures, as the underlying assumptions about human behavior are context-dependent.

We also clarify "empirical coherence" (Giovannini & Schiemer, 2025) as the degree to which a theory's axioms and laws align with diverse sets of empirical data, including both supportive evidence and non-contradictory observations. Unlike strict consistency, empirical coherence allows for minor discrepancies that are tentatively attributed to measurement error or unaccounted variables, as long as the core predictions hold. For example, climate models exhibit empirical coherence by aligning with temperature records, ice core data, and ocean current measurements, even if regional precipitation predictions have minor inaccuracies (Giovannini & Schiemer, 2025).

Axioms not refuted yet. Karl Popper's falsibility criterion (Popper, 1963), a foundational concept in the philosophy of science, argues that the defining feature of a scientific theory is its capacity to be falsified (i.e., tested and potentially proven false by empirical evidence)—since no amount of confirming observations can prove a theory absolutely true, but a single counterexample can disprove it. However, this criterion faces inherent flaws which is based on individual's judgement, as illustrated by the logical breakdown in Figure 1: The criterion hinges on whether a theory has a "way to falsify it", but this framework collapses into inconsistency. If a theory has already been falsified (e.g., via counterexample or paradox), it is rightly excluded from science—but if it has not yet been falsified, the criterion incorrectly assumes it remains scientific only if falsification is possible. Worse, if I believe there is "no way to falsify it", this is just my current belief and my future judgement could be different. Also my belief does not imply that others have the same belief and others' current beliefs may not represent their future beliefs. Therefore, each individual cannot confirm a theory to be unfalsifiable. So the criterion forces an unresolvable split: It incorrectly dismisses claims that might be falsifiable in the future, while also allowing no definitive judgment of "unfalsifiable" (since future methods could enable falsification). This invalidates Popper's rigid falsibility standard—instead, the more coherent and dynamic alternative is the "unrefuted yet" criterion: A theory stays within science as long as no counterexamples or logical contradictions have actually been found (regardless of hypothetical falsifiability), and is revised or excluded only when such refutation occurs. This shifts the focus from abstract "falsifiability potential" to evidence-based, provisional acceptance—aligning with science's progressive, open-ended nature and Hansson's (2006) "dynamic falsifiability" framework.

Scientific theories are not absolute truths but tentative explanations of natural and social systems. Therefore, the axioms and laws of scientific theories only need to be "not refuted yet" rather than "proven to be unfalsifiable". This criterion includes three aspects: (1) No counterexamples have been found to contradict the axioms and laws; (2) no logical paradoxes have been derived from the axioms and laws; (3) the logical deduction process from axioms to laws is correct. It should be emphasized that "not refuted yet" is a temporary state—with the expansion of observation scope and the deepening of research, the axioms and laws of existing theories may be refuted by new evidence, which is the driving force for the progress of science. For example, Newtonian mechanics was once considered universally applicable, but it was refuted in the fields of high speed and strong gravity, and was revised to be applicable only in the scope of low speed, weak gravity, and macroscopic systems, thus maintaining its scientificity within a specific scope (Hansson, 2006).

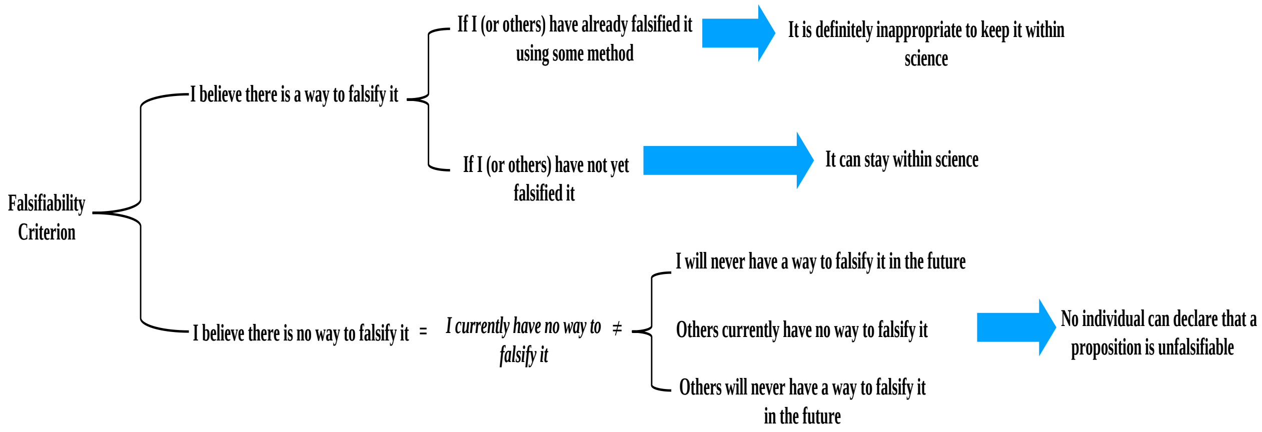

Figure 1. Logical inconsistencies in Popper's falsifiability criterion (in the first category of falsifiable, it contains unscientific part which should be removed from science; in the second category of non-falsifiable, Popper's criterion relies on individual judgments of "no way to falsify" (e.g., "I currently see no way to test this"). Yet such judgments are subjective: They do not represent consensus (others may disagree) or permanence (future methods could enable falsification). As no individual can declare a claim "ultimately unfalsifiable", this category collapses into incoherence).

Addressing Objections to the "Unrefuted Yet" Criterion

Two key counterarguments to the "unrefuted yet" criterion require explicit engagement:

1. Does "not refuted yet" collapse into verificationism? Popper (1963) rejected verificationism for its inability to account for the tentativeness of scientific knowledge, and this tension has been extensively debated (Frigg, 2023). To avoid this pitfall, "not refuted yet" explicitly requires testability in principle within the theory's defined scope. A theory is only scientific if its axioms/laws could, in principle, be refuted by observable evidence—even if current technology limits testing. For example, a theory about dark matter is scientific because it makes predictions about gravitational lensing that could be refuted by future observational data, even if dark matter has not yet been directly detected. In contrast, a theory that makes no predictions about observable phenomena (e.g., a metaphysical claim about the "purpose of the universe") is untestable in principle and thus unscientific, regardless of whether it has been refuted.

2. How to distinguish between "not refuted due to insufficient evidence" and "not refuted due to inherent untestability"? Drawing on Frigg (2023) and Holman and Wilholt (2022), the distinction lies in the theory's observational scope and testability potential. A theory is "not refuted due to insufficient evidence" if its scope is restricted to observable systems and it makes testable predictions that current methods cannot yet verify. For example, early quantum gravity models are not refuted because current particle accelerators lack the energy to test their predictions, but they are scientific because they could be refuted by future technological advancements. In contrast, a theory about the unobservable universe (e.g., claims about regions beyond the cosmic horizon) is inherently untestable because no observable evidence could, in principle, refute it—thus, it falls outside the scope of science (Holman & Wilholt, 2022). Robertson (2000) reinforced this by linking testability to Gödel's theorems: Scientific theories must remain within the bounds of observable systems to maintain their potential for refutation.

The modified Hilbert axiomatic standards abandon the unattainable completeness and unnecessary logical independence requirements of traditional standards, and adapt to the openness and progressive nature of scientific research. It emphasizes that scientific theories are dynamic systems that are continuously revised and improved based on empirical evidence and logical reasoning, which is consistent with the actual development of science.

Dynamic Demarcation Criteria: A New Basis for Distinguishing Science From Pseudoscience

Limitations of Existing Demarcation Criteria

The demarcation problem is to find a clear criterion to distinguish scientific theories from pseudoscience. Existing representative criteria have obvious limitations:

- Falsificationism (Popper, 1963): Popper holds that the core of scientificity is falsifiability—scientific theories should be able to be falsified by empirical evidence, while pseudoscience is unfalsifiable. However, this criterion faces two challenges: First, Popper claims that there are "absolute non-falsifiable statements", but in practice, most statements are "relatively falsifiable"—a statement that cannot be falsified under current conditions may be falsified with the progress of technology and research (Hansson, 2006); second, falsificationism fails to consider the revisability of scientific theories—when a theory encounters counterexamples, it is not necessarily directly rejected, but may be revised by narrowing the scope of application, which makes falsificationism too absolute (Hansson, 2006).
- Replicability: Some scholars regard replicability as the demarcation criterion of science, believing that scientific results should be repeatable by other researchers. However, Braude (2018) and Frigg (2023) pointed out that replicability is not applicable to many scientific fields, such as cosmology (the Big Bang cannot be replicated) and social science (the uniqueness of historical events), so it cannot be used as a universal demarcation criterion.
- Ellis's systematic criteria (Ellis, 2014): Ellis divides scientific criteria into four categories: satisfactory structure, intrinsic explanatory power, extrinsic explanatory power, and observational/experimental support, each including multiple sub-items. Although this criterion is comprehensive, it is too complex and many sub-items (such as simplicity, beauty, connectedness to other sciences) are subjective and difficult to operate, making it difficult to apply in practical research.

Construction of Dynamic Demarcation Criteria

Table 1

Mapping Between Modified Hilbert Axiomatic Standards and Dynamic Demarcation Criteria

Modified Hilbert axiom	Dynamic demarcation criterion	Operational test
Clearly defined concepts	Compliance with clearly defined requirements	Check that core concepts are operationalizable, consistent across the theory, distinguishable from competing definitions, and that the research scope is explicitly limited to observable systems (Hua et al., 2024; Holman & Wilholt, 2022). Discipline-specific guidelines apply (see Section 3.2.1 for examples).
Logical consistency	Compliance with logical consistency requirements	Verify that there are no contradictions between axioms, laws, and conclusions; in social sciences, confirm consistency within the specified cultural/contextual scope (Frigg, 2023; Giovannini & Schiemer, 2025). For edge cases with marginal inconsistencies, apply the hierarchy of resolution (see Section 4.2.4).
Axioms not refuted yet	Compliance with the "Not Refuted Yet" Requirement	Ensure that no counterexamples or paradoxes exist within the defined scope; confirm that the theory is testable in principle (even if current technology limits testing) (Frigg, 2023; Hansson, 2006). Rely on specific empirical evidence to support the "not refuted" status (see Section 4.3 for detailed case studies).

Combined with the modified Hilbert axiomatic standards, the redefined concept of science, and the limitations of existing criteria, this paper proposes dynamic demarcation criteria for scientific theories. The core idea is: The scientificity of a theory is not an absolute attribute but a dynamic state related to time, observation scope, and empirical evidence (Holman & Wilholt, 2022). The specific criteria include three aspects, which correspond to the modified Hilbert axiomatic standards one by one, as shown in Table 1.

This table explicitly links each core axiom of the modified framework to its corresponding demarcation criterion and provides an operational test for practical application. The mapping ensures that internal theoretical rigor (axiomatic standards) is directly translated into external demarcation power (criteria for distinguishing science from pseudoscience).

Compliance With Clearly Defined Requirements

The theory must clearly define its research object, scope, core concepts, and undefined concepts in line with the "operationalizable, consistent, distinguishable" standard (Hua et al., 2024), with discipline-specific adaptations as outlined in Section 3.2.1. If a theory uses vague, ambiguous, or arbitrarily defined concepts, or fails to clarify its applicable scope, it cannot be regarded as a scientific theory. For example, some pseudoscientific theories use ambiguous concepts such as "qi" (without clear operationalization or consistency) to explain natural phenomena, which violates the requirement of clear definition and thus is unscientific (Hua et al., 2024). In contrast, a theory of consciousness that defines the concept as "neural correlates of subjective experience measurable via functional Magnetic Resonance Imaging (fMRI)" meets the clarity requirement, even if debates about consciousness's ultimate nature persist (Hua et al., 2024).

Compliance With Logical Consistency Requirements

The axioms, laws, and conclusions of the theory must be logically consistent, without contradictions or paradoxes. In natural sciences, this requires universal consistency with empirical evidence (da Costa, 2011), while in social sciences, it may require consistency within a specific cultural or contextual scope (Frigg, 2023). If a theory contains logical contradictions (such as simultaneously affirming and denying a proposition) or can derive paradoxes from its axioms, it is unscientific. For example, some "perpetual motion machine" theories violate the law of conservation of energy (a logically consistent axiom in physics) and thus are pseudoscientific. In social science, a theory of economic inequality that claims "all individuals have equal access to resources" while also asserting "resources are distributed based on inherited wealth" contains a logical contradiction and is unscientific.

Compliance With the "Not Refuted Yet" Requirement

Within the defined scope of application, no counterexamples or paradoxes have been found to refute the theory's axioms and laws. This criterion has three characteristics:

- Dynamicity: A theory that is "not refuted yet" may be refuted by new evidence in the future, and thus its scientificity may change (Hansson, 2006). For example, the geocentric theory was once a scientific theory (not refuted under the observation conditions at that time), but it was refuted after the heliocentric theory was supported by sufficient evidence (e.g., Galileo's observations of Jupiter's moons, Kepler's planetary motion laws), and thus became unscientific.
- Scope limitation: When a theory encounters counterexamples, it can maintain its scientificity by narrowing the scope of application (Hansson, 2006). For example, Newtonian mechanics is refuted in the fields of high speed and strong gravity, but it is still scientific within the scope of low speed, weak gravity, and macroscopic systems.

- Relative falsifiability: All statements of scientific theories are relatively falsifiable—there is no absolute non-falsifiable scientific statement (Hansson, 2006). The fact that no counterexamples have been found so far does not mean that there will be no counterexamples in the future. This avoids the absolutism of Popper's falsificationism and reflects the progressive nature of scientific development.

Addressing Counterarguments to Dynamic Demarcation

To strengthen the framework's robustness, we explicitly engage with key objections:

- Relativism concern: Critics may argue that dynamicity collapses into relativism, as past theories like phlogiston could be deemed "scientific at the time" while lacking a robust distinction from pseudoscience. We resolve this by distinguishing between contextual scientificity and pseudoscience: A theory is contextually scientific if it meets the three criteria (clear concepts, logical consistency, not refuted) within the observational and technological constraints of its era, but it is pseudoscientific if it fails to revise or narrow its scope when new refuting evidence emerges. For example, phlogiston theory was contextually scientific in the 18th century (it had operationalized concepts, logical consistency, and no known counterexamples), but it became pseudoscientific when Lavoisier's oxygen theory provided conclusive counterevidence (e.g., mass gain in combustion) and phlogiston advocates refused to revise the theory (Kuhn, 1996). Pseudoscience, by contrast, inherently violates one or more criteria (e.g., vague concepts, unresolvable contradictions) regardless of context, rather than merely being superseded by newer science. This preserves objectivity by anchoring scientificity in both contextual compliance and responsiveness to evidence.
- Pragmatic implementation: Resolving criterion conflicts: The three-step application process may encounter conflicts (e.g., a theory with clear concepts but marginal logical inconsistencies, common in interdisciplinary research). We propose a hierarchy of resolution to adjudicate edge cases:

1. Priority to logical consistency: Severe contradictions (e.g., violating fundamental axioms like conservation laws) render a theory unscientific, as consistency is the foundational guarantee of reliability.

2. Tentative acceptance for marginal inconsistencies: For minor, unresolved inconsistencies (e.g., a social science theory with consistent core axioms but conflicting results in a small subset of contexts), the theory remains provisionally scientific if: (a) The inconsistencies are attributed to unaccounted variables (not inherent logical flaws); (b) the theory makes testable predictions to resolve the inconsistencies (e.g., adjusting for contextual confounders); and (c) the core concepts and scope remain clear. For example, early versions of evolutionary game theory had marginal inconsistencies in predicting altruistic behavior in small populations, but it remained scientific because researchers proposed testable modifications (e.g., inclusive fitness models) to resolve the discrepancies (Nowak, 2006).

3. Rejection for unresolvable conflicts: If a theory's inconsistencies persist despite repeated attempts at revision, or if resolving them requires abandoning clear concepts or testability, it is deemed unscientific.

Application of Dynamic Demarcation Criteria With Strengthened Empirical Grounding

The dynamic demarcation criteria provide a clear and operable method for distinguishing science from pseudoscience:

1. First, check whether the theory has clearly defined research objects, scope, and concepts (operationalizable, consistent, distinguishable) (Hua et al., 2024). If not, it is pseudoscience.

2. Second, check whether the theory is logically consistent (universally in natural sciences, contextually in social sciences) (Frigg, 2023; Giovannini & Schiemer, 2025). If there are contradictions or paradoxes, it is pseudoscience.

3. Finally, check whether the theory's axioms and laws have been refuted within the defined scope, and whether the theory is testable in principle (Frigg, 2023). If there are confirmed counterexamples and the theory cannot be revised by narrowing the scope, it is pseudoscience; if there are no counterexamples or the counterexamples can be eliminated by revising the scope, it is scientific.

This set of criteria avoids the subjectivity and complexity of existing criteria, and can effectively handle the demarcation problems in different fields. Below are detailed case studies with strengthened empirical grounding:

- Case 1: Big Bang Theory (Cosmology): To qualify as scientific, the theory's scope must be restricted to the "observable world" (finite spacetime accessible to human observation, distinct from the infinite "universe"). Key concepts are operationalized: "Cosmic Microwave Background (CMB) radiation" is defined as the residual thermal radiation from the Big Bang, measurable via satellite detectors (e.g., Planck Satellite) as a nearly uniform 2.725 K blackbody spectrum (Planck Collaboration, 2020). Logical consistency is demonstrated by aligning with other cosmological observations (e.g., redshift of distant galaxies, abundance of light elements like hydrogen and helium, which match Big Bang nucleosynthesis predictions). The "not refuted yet" criterion is supported by empirical evidence: CMB anisotropies (fluctuations in temperature) precisely match the predictions of inflationary Big Bang models, with no confirmed counterexamples (e.g., no observations of CMB radiation that cannot be explained by the theory within its observable scope). While the Big Bang cannot be replicated, its scientificity is preserved by clear concepts, logical consistency, and robust empirical support—fulfilling the dynamic demarcation criteria.
- Case 2: Economic Inequality Theory (Social Science): A theory defining its scope as "post-industrial Western societies" (e.g., EU and North American countries, 1990-2023) operationalizes "inequality" as the "ratio of top 10% income to bottom 10% income" (measurable via national tax records and household surveys). Logical consistency is maintained through axioms linking labor market polarization (e.g., demand for high-skill vs. low-skill labor) to income distribution, with no internal contradictions. Empirical support for the "not refuted yet" criterion comes from studies such as Piketty, Saez, and Zucman (2018), which find that the top 10% income share in post-industrial Western societies has increased from ~30% in 1990 to ~40% in 2023, aligning with the theory's predictions. The theory is testable in principle (e.g., future data could refute it if the income ratio declines without changes in labor market conditions) and maintains scientificity despite not applying to pre-industrial or non-Western societies.
- Case 3: Pseudoscience Contrast—Supernatural Healing: This theory uses vague concepts (e.g., "spiritual energy" with no operational definition: It cannot be measured, quantified, or linked to observable phenomena) and contains logical contradictions (e.g., claiming to "cure all diseases" while acknowledging failures without revising the theory's scope or axioms). It is untestable in principle: There is no way to refute the claim that "spiritual energy heals" because failures are attributed to "lack of faith" (an unobservable, unfalsifiable variable). Thus, it violates all three dynamic demarcation criteria and is clearly pseudoscience (Hines, 2003; Mukerji & Ernst, 2022).

For the clarity of concepts, the finite observable part of the universe can be called a world. So it is very important to distinguish the two concepts of universe which is the largest spacetime imaginable by human beings and world which is the largest spacetime observable by human beings (Cui, 2024; Holman & Wilholt, 2022).

Discussion: The Significance of New Scientific Standards

Responding to the Challenges of Modern Scientific Research

The new scientific standards proposed in this paper effectively respond to the challenges faced by modern scientific research. In the field of fundamental physics, the discussion of MToE has been troubled by the conflict between completeness and consistency (Cui, 2026). The modified Hilbert axiomatic standards abandon the pursuit of completeness and prioritize consistency, providing a reasonable framework for the construction of an effective MToE (da Costa, 2011). In cosmology, the openness and infinity of the universe make it impossible to construct a testable scientific theory, so the new standards emphasize the importance to distinguish the two fundamental concepts of universe and world. A scientific cosmological theory can only study the origin and the operation laws of the world under the condition that the universe already exists (Bunge, 1985; Cui, 2024; Holman & Wilholt, 2022).

In addition, the new standards also adapt to the characteristics of interdisciplinary and complex systems research. For example, in social science and ecological science, research objects are open and complex, and it is impossible to meet the traditional requirements of completeness and logical independence (Giovannini & Schiemer, 2025). The modified axiomatic standards and dynamic demarcation criteria provide a scientific evaluation basis for these fields, avoiding the problem of being regarded as “unscientific” due to the inability to meet overly strict traditional standards. Giovannini and Schiemer (2025) supported this by showing that complex sciences like climate science maintain rigor through empirical coherence rather than logical independence.

Promoting the Progress of Science and the Philosophy of Science

The new scientific standards emphasize the dynamic and progressive nature of science, which helps to correct the misunderstanding that “scientific theories are absolute truths”. It reminds scientists to maintain a critical spirit and continuously test and revise existing theories through empirical evidence, which is the core of the scientific spirit (Popper, 1963; Hansson, 2006). At the same time, the new standards clarify the boundary between science and pseudoscience, which helps to avoid the confusion of scientific research and guide social resources to be invested in truly scientific research.

In the field of the philosophy of science, the new standards integrate the advantages of traditional axiomatic systems and modern demarcation theories, and propose an operable framework, which helps to promote the solution of the demarcation problem. It also provides a new perspective for the discussion of issues such as the nature of scientific theories, the progress of science, and the relationship between science and truth. By engaging with recent scholarship (e.g., da Costa, 2011; Hansson, 2006; Hua et al., 2024), the framework contributes to ongoing debates about scientific methodology and demarcation.

Limitations and Future Research Directions

The new scientific standards proposed in this paper still have certain limitations. First, the “clearly defined concepts” criterion involves the problem of the degree of clarity, and there may be differences in the judgment of the clarity of concepts in different fields, which needs to be further refined combined with specific disciplines (Hua et al., 2024). For example, the operationalization of concepts in biology (e.g., “species”) may differ from that in psychology (e.g., “intelligence”), and future research should develop discipline-specific guidelines for concept clarity (building on the examples provided in Section 3.2.1). Second, the “not refuted yet” criterion relies on empirical evidence, and the reliability of empirical evidence itself may be affected by observation methods and technical conditions, which requires the support of standardized empirical research methods.

Future research can focus on two aspects: first, combining the characteristics of specific disciplines (such as physics, biology, social science) to formulate detailed implementation rules of the new standards, enhancing their operability; second, verifying and revising the new standards through case studies of scientific history and modern scientific research, and continuously improving their universality and rationality. For example, case studies of quantum gravity research (da Costa, 2011) or social science theories of inequality (Frigg, 2023) could further test and refine the framework.

Conclusion

This paper redefines science by integrating existing definitions and points out that science is a set of knowledge about observable systems based on empirical evidence and logical reasoning. Aiming at the limitations of traditional Hilbert axiomatic standards and existing demarcation criteria, this paper proposes a new framework of scientific standards, including modified Hilbert axiomatic standards (clearly defined concepts, logical consistency, unrefuted axioms) and dynamic demarcation criteria. The new scientific standards abandon the unattainable completeness and unnecessary logical independence requirements, emphasize the dynamic and progressive nature of scientific theories, and provide a clear, operable, and logically consistent basis for evaluating the scientificity of theories.

By engaging with recent scholarship (2018-2025), including work on axiomatic systems (da Costa, 2011; Giovannini & Schiemer, 2025), falsificationism (Hansson, 2006), concept clarity (Hua et al., 2024), and demarcation (Frigg, 2023; Holman & Wilholt, 2022), the paper demonstrates its alignment with the ongoing debates and audience. The framework not only responds to the challenges faced by modern scientific research (such as the research of MToE and complex systems) but also promotes the in-depth discussion of core issues in the philosophy of science such as the demarcation problem. It is believed that with the continuous development of science and the philosophy of science, the new scientific standards will be further improved and play a more important role in guiding scientific research and distinguishing science from pseudoscience.

Author Contributions

This is a single author paper.

Funding

This research was supported by the start-up funding from Westlake University under Grant Number 041030150118 and the scientific research project of Westlake University “Theoretical Research and Demonstration Application of Complex Systems and Deep-Sea Technology (Phase I)” under Grant Number WU2025A006.

Data Availability Statement

The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author CUI Weicheng.

Acknowledgments

During the preparation of this manuscript, the authors used Doubao (an AI developed by ByteDance) for providing support in searching information and language translation from Chinese to English throughout the research and paper drafting process. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Blachowicz, J. (2009). How science textbooks treat scientific method: A philosopher's perspective. *British Journal for the Philosophy of Science, 60*(2), 303-344.

Bohr, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review, 48*, 696-702.

Braude, S. E. (2018). Editorial. *Journal of Scientific Exploration, 32*(2), 255-264. Retrieved from https://doi.org/10.31275/2018/1330

Bunge, M. A. (1985). Cosmology: From myth to science. *The British Journal for the Philosophy of Science, 36*(2), 173-188. Retrieved from https://doi.org/10.1093/bjps/36.2.173

Case, R. (1985). *Intellectual development: Birth to adulthood.* London: Academic Press.

Cui, W. C. (2021). On an axiomatic foundation for a theory of everything. *Philosophy Study, 11*(4), 241-267. doi:10.17265/2159-5313/2021.04.001

Cui, W. C. (2024). Can science reveal the origin of the universe? *European Journal of Applied Sciences, 12*(1), 166-173. Retrieved from https://doi.org/10.14738/aivp.121.16256

Cui, W. C. (2026). Discussion on "Consequences of Undecidability in Physics on the Theory of Everything". Accepted and to be published in *Journal of Holography Applications in Physics, 6*(6), 1-4.

Cui, W. C., Li, R., & Pan, L. L. (2024). Toward a unified theory for complex systems. *European Journal of Applied Sciences, 12*(5), 69-103. doi:10.14738/aivp.125.17556

Cui, W. C., Li, R., & Pan, L. L. (2025). Unified Complex Systems Theory (UCST): Resolving materialist dilemmas through dualist ontology and active force. *European Journal of Applied Sciences, 13*(5), 517-545. doi:10.14738/aivp.1305.19511

Cui, W. C., Li, R., Pan, L. L., & Zeng, L. L. (2025). The problem of logical self-circulation in modern scientific theories and its resolution. *Philosophy Study, 15*(6), 291-314. doi:10.17265/2159-5313/2025.06.001

Da Costa, N. C. A. (2011). Gödel's incompleteness theorems and physics. *Principia, 15*(3), 453-459. Retrieved from https://doi.org/10.5007/1808-1711.2011v15n3p453

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology, 56*(6), 879-886. Retrieved from https://doi.org/10.1080/10635150701701083

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review, 47*, 777-780.

Ellis, G. F. R. (2014). On the philosophy of cosmology. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 46*(1), 5-23.

Faizal, M., Krauss, L. M., Shabir, A., & Marino, F. (2025). Consequences of undecidability in physics on the theory of everything. *Journal of Holographic Applied Physics, 5*, 10-21.

Frigg, R. (2023). *Models and theories a philosophical inquiry*. London and New York: Routledge.

Giovannini, E. N., & Schiemer, G. (2025). Hilbert's early metatheory revisited. *Erkenntnis*. Retrieved from https://doi.org/10.1007/s10670-025-00959-z

Gödel, K. (1931). About formally undecidable sentences of the principia mathematica and related systems I. *Monthly Magazines for Mathematics and Physics, 38*, 173-198.

Hansson, S. O. (2006). Falsificationism falsified. *Foundations of Science, 11*(3), 275-286. doi:10.1007/s10699-004-5922-1

Hilbert, D. (1902). Mathematical problems. *Bulletin of the American Mathematical Society, 8*(10), 437-479.

Hines, T. (2003). *Pseudoscience and the paranormal* (2nd ed.). Amherst, N.Y.: Prometheus Books.

Holman, B., & Wilholt, T. (2022). The new demarcation problem. *Studies in History and Philosophy of Science Part A, 91*(C), 211-220. doi:10.1016/j.shpsa.2021.11.011

Hua, D., Petrina, N., Young, N., Cho, J.-G., & Poon, S. K. (2024). Understanding the factors influencing acceptability of AI in medical imaging domains among healthcare professionals: A scoping review. *Artificial Intelligence in Medicine, 147*, 102698. Retrieved from https://doi.org/10.1016/j.artmed.2023.102698

Johansson, L.-G. (2016). *Philosophy of science for scientists*. NY: Springer International Publishing Switzerland.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago: University of Chicago Press.

Maudlin, T. (2018). Ontological clarity via canonical presentation: Electromagnetism and the Aharonov-Bohm effect. *Entropy, 20*(6), 465-485. Retrieved from https://doi.org/10.3390/e20060465

Mukerji, N., & Ernst, E. (2022). Why homoeopathy is pseudoscience. *Synthese*, 200, Article 394. Retrieved from https://doi.org/10.1007/s11229-022-03882-w

Newton, I. (1846). *Newton's principia: The mathematical principles of natural philosophy*. (A. Motte, Trans.). New York: Daniel Adee. (Original work published 1687)

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560-1563. Retrieved from https://doi.org/10.1126/science.1133755

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics, 133*(2), 553-609. Retrieved from https://doi.org/10.1086/696699

Planck Collaboration. (2020). Planck 2018 results. VI. Cosmological parameters. *Astronomy & Astrophysics, 641*, A6. Retrieved from https://doi.org/10.1051/0004-6361/201833910

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge* (2002 ed.). London: Routledge.

Robertson, D. S. (2000). Goedel's theorem, the theory of everything, and the future of science and mathematics. *Complexity, 5*(5), 22-27. doi:10.1002/1099-0526(200005/06)5:53.0.CO;2-0

Rovelli, C. (2018). *Reality is not what it seems: The journey to quantum gravity*. New York: Riverhead Books.

Walach, H. (2019). Beyond a materialist worldview-towards an expanded science. *Scientific and Medical Network*. Retrieved from www.explore.scimednet.org

Weber, M. (2002). Economy and society: An outline of interpretive sociology. In N. W. Biggart (Ed.), *Readings in economic sociology* (Chap. 3). Malden, Massachusetts: Wiley. Retrieved from https://doi.org/10.1002/9780470755679

Weinberg, S. (2011). *Dreams of a final theory: The scientist's search for the ultimate laws of nature*. New York: Knopf Doubleday Publishing Group. ISBN:978-0-307-78786-6

Whitaker, A. (2006). *Einstein, Bohr and the quantum dilemma—From quantum theory to quantum information* (2nd ed.). Cambridge: Cambridge University Press.

Wu, Y. L., Zhang, H., & Li, J. Q. (2024). Quantum gravitational corrections to Schwinger pair production in strongly curved spacetimes: Implications for black hole accretion disks. *Physical Review D, 109*(8), 084032. Retrieved from https://doi.org/10.1103/PhysRevD.109.084032