Article

Steins Theory: A New Axiomatic System for Identity

Jiaqi Guo *

Posted Date: 23 September 2025

doi: 10.20944/preprints202506.2384.v6

Keywords: ethics; identity; category mistake; n = n; quantum identical particles; spacetime leap

Article

Steins Theory: A New Axiomatic System for Identity

Jiaqi Guo

Independent Researcher, China; 1028165464@qq.com

Abstract

In the philosophy of language, Frege's distinction between sense and reference provided a foundational framework for identity statements, while Putnam's Twin Earth thought experiment, with its remarkable insight, pushed externalism to its limits, successfully challenging the internalist model of meaning and setting the agenda for decades of debate on the determinacy of reference. However, despite the groundbreaking nature of these works, a curious phenomenon persists: the debates they sparked such as Ship of Theseus or identical particles. This paper argues that this stalemate may not stem from the depth of the problem itself but rather from a shared, unexamined assumption underlying these otherwise compelling theories: the belief that there exists a single, decisive level (whether microphysical structure or historical causation) capable of conclusively resolving the identity question. This paper proposes that, rather than continuing to seek a superior singular answer under this assumption, a more productive approach lies in critically examining the assumption itself. To this end, we develop a hierarchical relativity framework This framework does not aim to negate prior work but seeks to clarify its valid scope of application, offering a new path to resolve a series of philosophical puzzles born of category mistakes.

Keywords: ethics; identity; category mistake; n = n; quantum identical particles; spacetime leap

1. Introduction

The problem of identity—what makes a thing what it is—lies at the core of metaphysics and logic. From Leibniz's (1714) profoundly insightful Principle of the Identity of Indiscernibles (PII), with its logical elegance and power in setting the lofty ideal of entity individuation, to Kripke's (1980) groundbreaking theory of necessary identity based on rigid designators and origins, which, with its deep grasp of modal intuitions, provided a seemingly robust foundation for the stability of reference, generations of philosophers have constructed a grand intellectual tradition for understanding the persistence and recognition of individuals.

Yet, a curious phenomenon emerges: these highly persuasive theoretical frameworks, while compelling within their domains, often reveal systematic limitations when addressing complex boundary cases posed by reality or thought experiments. Leibniz's strong PII principle encounters fundamental difficulties, perhaps unforeseen in its original conception, when confronted with quantum identical particles (French & Redhead, 1988). Its intrinsic rigor paradoxically leads to a near-paradoxical stalemate when dealing with multiple entities (e.g., two electrons) that are entirely indistinguishable in all intrinsic properties yet numerically distinct. Similarly, Kripke's elegant theory, designed to anchor reference, becomes entangled when addressing diachronic changes in intrinsic properties (e.g., the Ship of Theseus) (Chandler, 1975), let alone accommodating the phenomenon of quantum identical particles, which challenges classical notions of individuality.

Thus, we face a peculiar intellectual impasse: whether it is Leibniz's pursuit of property-based identity or Kripke's focus on the necessity of historical origins, these approaches brilliantly illuminate one wing of the edifice of identity but inadvertently cast another into deeper shadow. Could their shared, perhaps unspoken ambition—to find a single, absolute criterion for identity—itself constitute an a priori obstacle, preventing us from grasping the multidimensional nature of the problem?

 doi:10.20944/preprints202506.2384.v6

This paper contends that the key to breaking this impasse lies not in choosing between existing paths or attempting further patchwork but in stepping back for a meta-level reflection on the problem itself. We propose a hierarchical relativity framework for identity analysis (Steins Theory). This framework does not seek to wholly reject prior work—indeed, Leibniz's focus on properties and Kripke's insistence on history find their limited legitimacy within it—but aims to show that the root of the aforementioned dilemmas lies in a category mistake: the erroneous attempt to answer a question posed at one level with an answer from another.

By introducing a relativity framework defined by the Content() function, this paper aims to transform classical ontological and semantic disputes into a clear, operational problem of conceptual choice. Thus, it does not seek to solve these puzzles but to dissolve them, offering a new path to alleviate a series of philosophical anxieties arising from category mistakes.

Note: Unless otherwise specified, this paper adopts an objective scientific perspective (physics) by default.

2. Analysis

2.1. Axiomatic Foundations

2.1.1. Axiom 1 (Self-Identity)

$$\forall n \in I, n \equiv n$$

An information entity is necessarily identical to itself, forming the foundation of logical reference.

2.1.2. Axiom 2 (Distinctness)

$$\forall n, m \in I, (n \neq m) \Leftrightarrow (\text{Content}(n) \neq \text{Content}(m))$$

where the domain of Content() is uniquely determined by the comparison target.

$$\text{Core Corollary: Content}(n) = \text{Content}(m) \Rightarrow n \equiv m$$

Uniqueness Theorem: Under a given Content() definition, two entities with identical content are necessarily the same information entity.

2.2. Category Mistake and Level Confusion: The Ship of Theseus as a Case Study

In addressing the enduring puzzle of the Ship of Theseus, one highly influential and intuitively appealing approach, represented by David Wiggins (1980), posits that an object's identity is not guaranteed by its instantaneous properties at any given moment but must be sustained by spatiotemporal continuity and an uninterrupted historical-causal path. This approach's great strength lies in capturing our deep intuition about "objects persisting through time"—namely, that entities are not mere momentary existences but have a "life history" or "biography."

Another approach, four-dimensionalism (Perdurantism), offers a radically different yet metaphysically elegant picture (see Heller, 1984; Sider, 2001). With its radical clarity, this theory ingeniously sidesteps many pitfalls of diachronic identity. Four-dimensionalism posits that the Ship of Theseus is not a three-dimensional entity "wholly present" in time but a "spacetime worm" extended in four-dimensional spacetime. Each temporal slice of the ship is a temporal part of this four-dimensional object. Thus, "change" merely means that the four-dimensional whole possesses different properties (e.g., different planks) at different temporal parts. Under this framework, the original Ship of Theseus (as a four-dimensional entity), the replaced ship, and the ship reconstituted from old planks are three distinct four-dimensional objects. They may share identical three-dimensional cross-sections at certain moments (making them indistinguishable at that instant), but as wholes, they are naturally distinct. The brilliance of this approach lies in transforming the vexing problem of persistence into a relatively clear "part-whole" relation.

 doi:10.20944/preprints202506.2384.v6

Closely related is Stage Theory (see Sider, 1996; Hawley, 2001), which retains many of four-dimensionalism's advantages while attempting to better accommodate our everyday linguistic intuition that "objects are three-dimensional." Stage theorists argue that what we typically call "the Ship of Theseus" refers not to the entire four-dimensional worm but to a stage or temporal slice at a specific time. When we say "the ship is the same" at time t, we mean that the stage at t is related to a prior stage at t_1 through a (primitive) counterpart relation, sustained by similarity and causal continuity. Stage Theory's conceptual economy avoids presupposing cross-temporal identity relations, making it theoretically compelling.

However, despite the ingenuity and internal coherence of these theories, this paper argues that they face a shared, profound normative dilemma. Whether appealing to historical paths, four-dimensional wholes, or counterpart relations, these theories strive to provide a single, absolute criterion for identity. To achieve this, they must incorporate "time" or "spatiotemporal position" as constitutive elements of an object's identity. This implies that, within their frameworks, answering the question "Is the ship at t_1 identical to the ship at t_2?" logically depends on verifying spatiotemporal coordinates or cross-temporal relations.

This paper contends that this critical theoretical move inadvertently commits a "category mistake." Let us formalize and reconstruct these theories' approaches: they effectively adopt a content function with a domain Content_Hist(Ship) = (material, form, function, historical-causal path/four-dimensional whole/counterpart relation). In contrast, we argue that the original identity question—"Is this ship, with identical functional structure, still itself?"—presupposes a content function defined solely as Content_Struct(Ship) = (material, form, function).

This can be formalized as follows:

1. Original Question: Determine whether the entity "ship" at time t_1 and t_2 is identical.
2. Correct Content() Domain (based on the original question): Should include only properties relevant to the ship's abstract information structure, i.e., Content_Struct(Ship) = (material, form, function).
3. Category Mistake in Historical/Four-Dimensional/Stage Theories:
   - Wiggins effectively uses Content_Hist(Ship) = (material, form, function, historical-causal path).
   - Four-dimensionalism uses Content_4D(Ship) = (material, form, function, spatiotemporal coordinates).
   - Stage Theory uses Content_Stage(Ship) = (material, form, function, counterpart relation).

The resulting process can be formally expressed as:

- They attempt to answer the identity question Q_Struct: "Is Ship at t_1 ≡ Ship at t_2?" based on Content_Struct().
- However, they actually employ the judgment functions Content_Hist(), Content_4D(), or Content_Stage().
- Since the materials in the replacement process are identical to the original materials, we have Content_Struct(Ship_t_1) = Content_Struct(Ship_t_2).
- But because the historical path, spatiotemporal coordinates, or counterpart relations have changed, Content_Other(Ship_t_1) ≠ Content_Other(Ship_t_2).
- Thus, they conclude Ship_t_1 ≠ Ship_t_2 to answer Q_Struct.

This leads to an intriguing situation: each theory effectively answers a question, but it may not be the one originally posed. They precisely address questions like "Is there a continuous four-dimensional worm connecting the ship at t_1 and t_2?" or "Is the stage at t_2 a counterpart of the stage at t_1?" but treat these answers as definitive resolutions to the question "Is the ship structurally identical?" This is akin to a judge, asked whether a defendant's actions violate criminal law article X, delivering a verdict based on civil law statutes. The conclusion may be coherent within its own system but has subtly shifted the debate's ground—a category mistake.

Thus, the true contribution of historical path theories may lie in brilliantly revealing how properties like "history" or "spatiotemporal whole" powerfully shape our identity judgments. Their limitation, however, lies in attempting to elevate these explanatory properties into metaphysically necessary conditions, requiring an expansion of the identity question's criteria to maintain theoretical

 doi:10.20944/preprints202506.2384.v6

coherence. The value of Steins Theory lies in avoiding such difficult expansions. By clarifying the hierarchy of questions, it allows Content_Struct(), Content_Hist(), Content_4D(), and Content_Stage() to provide valid, non-conflicting answers to different questions. It does not solve these puzzles but dissolves them.

2.3. Information Conservation

2.3.1. Quantum Identical Particles and the Challenge to PII

The indistinguishability of quantum identical particles poses the most severe challenge to Leibniz's PII (References 1, 2), yet it provides a natural, empirically grounded model for this theory.

Current philosophical discussions on identical particles, facing the challenge quantum particles pose to PII, primarily fall into two camps: revisionism and revolutionism. The former seeks to salvage some form of individuality principle, while the latter abandons individuality altogether.

Saunders' approach undoubtedly represents one of the most ingenious and technically rigorous attempts within the revisionist path. By subtly defining "weak discernibility," he successfully liberates the discussion from the dead end of intrinsic properties, offering a creative perspective on seeking individuation in relational terms. The complexity of this approach and the extensive debates it has sparked testify to its profound philosophical value. However, this very complexity reveals a potential cost: its definition of "purely extensional relational properties," while striving for precision, inevitably introduces considerable terminological ambiguity, requiring its defenders to cautiously address accusations of circular reasoning (Muller & Saunders, 2008). More critically, the entire theoretical edifice rests on an unsettling presupposition: that the "individuality" of identical particles must, and can only, be "saved" by identifying some (even relational) individuating property. This renders its efforts—however ingenious—essentially an ad-hoc patch to preserve a premise. When applied to states in quantum field theory with indefinite particle numbers, the ad-hoc nature of this strategy becomes increasingly apparent: it feels less like an elegant inference of the theory itself and more like an increasingly costly effort to uphold the premise that individuality must exist.

Facing the revisionist dilemma, the revolutionary approach takes a more radical path. Scholars like Décio Krause (2011) propose a highly subversive thesis: quantum particles may not be "individuals" in the traditional metaphysical sense at all. Thus, laws based on individual identity are categorically misapplied from the outset. They should be understood as "non-individuals," described using highly specialized mathematical tools like quasi-set theory.

Krause's approach is striking for its conceptual thoroughness and consistency, boldly embracing the most counterintuitive features of quantum mechanics and decisively breaking with our classical frameworks of objects and spatiotemporal localization. This resolute stance is theoretically clean and efficient. However, this efficiency comes at a cost: the concept of "non-individuals" imposes a significant metaphysical interpretive burden, requiring us to abandon an entire set of deeply ingrained intuitions about what "one thing" means.

2.3.2. Steins Theory's Solution: A Hierarchical Relativity Framework

Both approaches share a deeper error: they attempt to answer a wrongly posed question. The issue is not "What is the correct individuating property?" but "At what level are we asking the identity question?"

Steins Theory provides a meta-framework by introducing a hierarchical Content() function. We define an observable particle state as an ordered pair: P = ($\mathcal{I}$, $\mathcal{S}$), where $\mathcal{I}$ is the set of intrinsic properties (mass, charge, spin, etc.), and $\mathcal{S}$ is the set of spatiotemporal coordinates.

- When asking at the level of Content(particle) = $\mathcal{I}$, i.e., comparing only intrinsic properties, all identical electrons share Content(electron) = (mass m_e, charge -e, spin 1/2, ...). By Axiom 2, Content(n) = Content(m) ⇒ n ≡ m. At this level, they are indeed the same information entity e, explaining the root of their indistinguishability.

 doi:10.20944/preprints202506.2384.v6

- When asking at the level of Content(particle state) = ($\mathcal{I}$, $\mathcal{S}$), since $\mathcal{S}$ (e.g., position, momentum) necessarily differs, Content(P_1) ≠ Content(P_2), so they are distinct particle states. This explains why we observe multiple scattering events in experiments.

Thus, the confusion caused by quantum identical particles arises from erroneously applying differences at the $\mathcal{S}$ (spatiotemporal coordinates) level to identity judgments at the $\mathcal{I}$ (intrinsic entity) level. Steins Theory resolves this contradiction by clearly distinguishing these levels: particles are both "one" (as an abstract information entity) and "many" (as manifestations in specific spacetime coordinates). Particle annihilation and creation merely represent the decoupling and recoupling of the invariant information entity e with different coordinates $\mathcal{S}$.

This approach's advantage lies in absorbing the strengths of Krause's scheme (acknowledging the uniqueness of quantum entities by interpreting "non-individuality" as identity at the $\mathcal{I}$ level) while avoiding its radical metaphysical cost (we still discuss "entities," just at different levels). It also explains why Saunders' strategy of introducing relational properties seems plausible in some cases (because he mistakenly treats $\mathcal{S}$-level properties as individuating bases for the $\mathcal{I}$ level) but is fundamentally misguided.

2.3.3. Formal Derivation of Information Conservation

Define a fundamental particle state as an ordered pair: P = Content($\mathcal{I}$, $\mathcal{S}$), where:

- $\mathcal{I}$ is the set of intrinsic properties (e.g., mass m, charge q, spin s).

- $\mathcal{S}$ is the set of spatiotemporal coordinates (e.g., position x, time t).

Axiomatic Operations:

1. Coordinate Binding: P_k = ($\mathcal{I}$, $\mathcal{S}$_k), e.g., an electron near a photon: $\mathcal{S}$_k = (relative position: adjacent to photon γ).
2. Coordinate Decoupling (Destruction): ($\mathcal{I}$, $\mathcal{S}$_k) → ($\mathcal{I}$), ($\mathcal{S}$_k) ⇒ The particle degrades into a pure eigenstate entity, unmeasurable due to the lack of observable basis ($\mathcal{S}$ = ∅).

For any particle states ($\mathcal{I}$_α, $\mathcal{S}$_i) and ($\mathcal{I}$_α, $\mathcal{S}$_j), we have:
(Content($\mathcal{I}$_α) ≡ Content($\mathcal{I}$_α)) ⇨ ($\mathcal{I}$_α, $\mathcal{S}$_i) ≡ ($\mathcal{I}$_α, $\mathcal{S}$_j), indicating:

- When two particles' intrinsic properties are indistinguishable ($\mathcal{I}$_α = $\mathcal{I}$_α), regardless of differences in their spatiotemporal coordinates $\mathcal{S}$_i ≠ $\mathcal{S}$_j, they are the same information entity e = Content($\mathcal{I}$) projected in different spacetimes.

Physical Interpretation:

- Particle Annihilation: Represents set decoupling, not destruction ⇒ e = ($\mathcal{I}$) enters a free state.
- Particle Creation: The same e binds to new coordinates $\mathcal{S}$' ⇒ observed as reappearance.

Example: An electron e^- disappearing at position x_1 and reappearing at x_2 is merely a coordinate migration of the entity e = (q = -1e, m_e, s = 1/2, ...): (e, x_1) → (e) → (e, x_2), with its information identity guaranteed by n = n.

Direct Corollary: Information Conservation Theorem: Logic permits that information, once existent, neither annihilates nor updates.

2.4. Symmetry

Max Black's (1952) symmetrical universe thought experiment poses the most extreme challenge to Leibniz's strong PII. He envisions a universe containing only two perfectly identical spheres, indistinguishable in all intrinsic properties (mass, composition, shape, etc.) and relational properties (e.g., X miles apart, symmetrical to each other). Black argues this is a genuine scenario of "two" entities, thereby refuting PII—there exists no property to distinguish them, yet they remain numerically distinct.

Traditional responses fall into two camps: one questions the metaphysical possibility of such a symmetrical universe (e.g., demanding a basis for "numerical difference," often reverting to hidden

 doi:10.20944/preprints202506.2384.v6

properties); the other, like Saunders (2003), argues that relational properties (e.g., "being X miles from a sphere") can serve as a weakened basis for distinction. However, the former is criticized as ad-hoc, while the latter fails in Black's original setup, as each sphere's relational property ("X miles from another sphere") remains identical.

This paper argues that Black's challenge and the dilemmas of traditional responses stem from an unexamined assumption: that "numerical twoness" is a primitive, irreducible fact. Steins Theory offers a novel analytical perspective. We must first specify the Content() function's comparison level.

- If Content(sphere) is defined as the set of all traditional properties (intrinsic + relational), i.e., Content(A) = {mass M, spherical shape, ..., X miles from a sphere}, then by Axiom 2, since Content(A) = Content(B), we inevitably conclude A ≡ B (Uniqueness Theorem). This seems to directly yield the PII conclusion Black sought to refute.
- However, Black's intuition—"there are clearly two spheres"—is not without basis. This theory interprets it as a cognitive bias. Observers report "seeing two" because their perspective is embedded in this symmetrical spatiotemporal coordinate system $\mathcal{S}$.

This paper argues that Black and his critics share a fundamental error: assuming that the referents "A" and "B" necessarily correspond to two entities with independent spatiotemporal coordinates. This assumption forces a dilemma between "abandoning PII" or "inventing new metaphysical concepts." Steins Theory offers a third path through the concept of "Coordinate Self-Reference."

System Content() Formalization: For the entire symmetrical system S, we define: Content(S) = {there exists an entity a with property set P, and a is opposite a}. This description sounds complex, but simply put, Content(S) describes a single coordinate framework allowing "self-opposition." The visual appearance of "two" spheres is the projection of this single, self-referential coordinate structure in Euclidean space perception (akin to an object and its mirror image, but without a mirror, due to the topological properties of space itself).

Resolving the Paradox:

- At the Content(entity) level: Content(a) = Content(b) = P. By the Uniqueness Theorem, this points to a single entity, denoted e.
- At the Content(system S) level: Content(S) describes a single entity e bound to a special self-referential coordinate topology: (e, R_self-facing).
- Paradox Dissolved: Black's error lies in inferring the existence of two entities (e_a and e_b) from the system state (e, R_self-facing). He confuses levels, using Content(S)'s description to answer a question about Content(e). In reality, there was never a second entity b; there is only one entity e in a special coordinate topology producing a "double-image projection."

Thus, this framework does not deny the intuition of "seeing two spheres" but provides a novel, precise ontological explanation: it is the perception of a single entity in a self-referential coordinate topology. This successfully resolves the apparent contradiction between PII and counting intuitions without introducing ad-hoc individuating factors. Black's challenge, far from refuting the law of identity, reveals, through this framework's hierarchical analysis, the dependence of "identity" judgments on the background framework.

3. Examples (Derived Mathematically, Assuming a Scientific Perspective by Default)

Notably, this framework not only resolves quantum puzzles but also offers a new perspective on Gibbs' Paradox. Additionally, for classical paradoxes like Gibbs' Paradox, this framework provides novel insights and conclusions.

3.1. The Duplication Paradox

- Issue: Are two documents with identical content stored on different devices two distinct pieces of information?

 doi:10.20944/preprints202506.2384.v6

- Solution:
  - If the target is pure informational content identity → Content(n) = textual semantics, then n ≡ m (single entity).
  - If the target is document-location entity identity → Content(An) = (textual semantics, location), then (n, Loc_A) ≠ (m, Loc_B).
- Conclusion: Duplicates are the same information entity bound to different spatiotemporal coordinates, resulting in observable distinctness.

3.2. Gibbs' Paradox

- Category Mistake:
  - The target should be particle type identity → Content(g) = (mass, spin, ...).
  - Classical statistics illicitly expands Content(Ag) = (intrinsic properties, fictitious label).
- Correction: ∀g_i, g_j: Content(g_i) = S_int = Content(g_j) ⇒ g_i ≡ g_j (type identity).
  Entropy errors arise from erroneously selecting the Content() domain (introducing labels).

3.3. Black Hole Information Paradox

- Category Mistake: Illegitimately binding the information entity n's Content() to spatiotemporal coordinates: Content(An) = (information structure, black hole coordinates).
- Correct Solution:
  - Define the target: information structure identity → Content(n) = quantum properties.
  - A black hole decouples the set (quantum properties, coordinates), leaving unpaired content unobservable, but Content(quantum properties/coordinates) as an abstract entity persists.
  - If a new spacetime entity satisfies Content(m) = Content(n), then m ≡ n.

3.4. Chinese Room Thought Experiment

- Define the target entity as the Chinese understanding function: Content(understanding) = input-output behavioral consistency.
- If the Chinese Room system's behavior is indistinguishable from a native speaker, Content(system) ≡ Content(human), then by Axiom n = n: the system objectively understands Chinese.
- Searle's fallacy lies in illicitly expanding Content() to a subjective dimension (e.g., the operator's mental state), conflating the functional entity with the consciousness entity level.

3.5. The Ship of Theseus

A highly influential view (cf. Kripke, 1980; Wiggins, 1980) holds that an object's identity is guaranteed by spatiotemporal continuity and historical-causal paths. This paper argues that this view erroneously incorporates "spatiotemporal coordinates" and "history" into the Content() function's domain for "ship" without scrutiny.

Definition: Define the target entity as the ship's abstract information structure: S = Content(ship) = (material, form), assuming replacement materials are identical to the originals.

Process:

1. Component Replacement: $(S, \mathcal{S}_0) \rightarrow (S, \mathcal{S}_0) \rightarrow (S, \mathcal{S}_0)$.
2. Old Parts Reassembly: $(S, \mathcal{S}_0) \rightarrow S \rightarrow (S, \mathcal{S}_n)$.

Since Content(S) ≡ Content(S) (n = n), in $(S, \mathcal{S}_n)$, $(S, \mathcal{S}_o)$, $(S, \mathcal{S}_0)$, it is always the same ship.

Critique: Claiming historical causation determines identity commits a category mistake: Content(A_ship) = (function, form, history) violates the initial Content() domain, constituting a category error.

3.6. Twin Earth Paradox

 doi:10.20944/preprints202506.2384.v6

Putnam's (1975) thought experiment argues that meaning is not in the head. This theory offers a new perspective on this debate.

- Traditional Contradiction: Earth's "water" (H_2O) differs chemically from Twin Earth's "water" (XYZ), but are the concepts of "water" identical for residents of both planets?
- Theoretical Solution:
  - If Content(water concept) = macroscopic properties (colorless, chemically reactive, potable liquid, etc.) → the concepts are identical (n = n).
  - Introducing microscopic structure (H_2O/XYZ) expands the Content() domain to the molecular level, committing a category mistake.
- Conclusion: Semantic identity is determined solely by cognitive function, independent of underlying physical structure.

3.7. Grandfather Paradox

- Contradiction: If one travels back and kills one's grandfather ⇒ one should not exist ⇒ one cannot perform the killing.
- Theoretical Dissolution:
  - Define the target entity: worldline identity Content(worldline) = logical structure of events.
  - The killing event results in:
    - Original worldline W_0: (grandfather lives → you exist → you kill).
    - New worldline W_1: (grandfather dies → you do not exist).
  - Since Content(W_0) ≠ Content(W_1), W_0 and W_1 are distinct information entities (not "the same worldline modified").

3.8. Brain in a Vat

Current debates on the "brain in a vat," whether skeptical or realist, implicitly incorporate the properties of the "external carrier" (biological brain or vat) into judgments of "cognitive entity" identity without scrutiny. By strictly distinguishing Content(cognitive flow) and Content(carrier), this paper aims to dissolve the debate itself:

- Problem: How can one prove one is not a brain in a vat? Perception cannot distinguish reality from simulation.
- Theoretical Formula:
  - Define Content(cognitive entity) = perceptual information flow.
    - Real brain B_real: Content = {light signals, touch, ...}.
    - Vat brain B_vat: Content = {electrical stimulation signals}.
  - If Content(B_real) ≡ Content(B_vat) ⇒ by Axiom B_real ≡ B_vat (same cognitive entity).
  - Key Point: The "reality" dispute stems from expanding Content() to the external carrier (skull/vat), while the cognitive entity is determined solely by the information flow.

3.9. Mary's Room

- Scenario: Mary knows all about color neuroscience but has never seen red → Does she gain new knowledge upon seeing red?
- Theoretical Solution:
  - Define knowledge entity levels:
    - Propositional knowledge: Content(K_prop) = data on red light wavelengths.
    - Qualia knowledge: Content(K_qualia) = subjective red experience.
  - Since Content(K_prop) ≠ Content(K_qualia), they are distinct information entities.
  - Mary gains a new entity K_qualia, not a supplement to K_prop ⇒ the paradox arises from conflating knowledge types.

3.10. Newcomb's Paradox

 doi:10.20944/preprints202506.2384.v6

- Core Paradox: A near-perfect predictor vs. the participant's free will. Choose one box (known to contain money) or two boxes (potentially more money)?
- Theoretical Deconstruction:
  - Category Mistake: Conflating the Content() domain of the decision entity:
    - Level 1 (pure decision logic): Content(decision) = (action, payoff function) ⇒ dominant strategy: choose two boxes (regardless of prediction accuracy).
    - Level 2 (causal history binding): Content(decision) = (action, payoff function, prediction history) ⇒ if prediction is accurate, one box yields higher payoff.
  - Uniqueness Theorem Ruling:
    - If the target is rational decision without historical constraints → use Content(Level 1) ⇒ choose two boxes.
    - If the target is decision with predictive causation → use Content(Level 2) ⇒ choose one box.
  - Paradox Dissolved: The two are distinct decision entities, Content(Level 1) ≠ Content(Level 2); the contradiction arises from domain swapping.

3.11. Raven Paradox

- Core Paradox: "All ravens are black" ≡ "All non-black things are non-ravens." Why does observing a red apple (non-black and non-raven) confirm the proposition?
- Theoretical Deconstruction:
  - Category Mistake: Expanding the Content() domain of "confirmation behavior" from the logical structure of the proposition to the type of empirical sample.
  - Correct Definition:
    - Proposition identity: Content(P) = logical form ($\forall x: R(x) \rightarrow B(x)$).
    - Confirmation identity: Content(confirmation) = verification of $\neg\exists x: (R(x) \wedge \neg B(x))$.
  - Conclusion:
    - A red apple confirms the logically equivalent contrapositive (non-black ⇒ non-raven), with Content(confirmation) identical to observing a raven (since Content(P) ≡ Content(P)).
    - Claiming "red apples and ravens differ in confirmatory strength" commits a category mistake, expanding Content() to physical sample categories (birds/fruits), violating the initial logical target.

3.12. Sorites Paradox (Baldness Paradox)

- Core Paradox: Removing one grain of sand does not turn a heap into a non-heap ⇒ removing all sand still yields a "heap," a contradiction.
- Theoretical Deconstruction:
  - Category Mistake: Conflating the Content() definition of "heap":
    - Level 1 (topological structure): Content(heap) = macroscopic form of sand collection ⇒ removing one grain does not alter form identity (n = n).
    - Level 2 (atomic quantity): Content(heap) = number of grains N ⇒ when N = 0, Content(heap) = ∅, entity ceases.
  - Solution:
    - If heap is defined as a form entity Content(Level 1), removing one grain retains the same heap.
    - If defined as a quantity entity Content(Level 2), each grain removal creates a new entity.
  - Paradox Root: Swapping Content() domains (from form to quantity) mid-argument.

3.13. Sleeping Beauty Problem

- Core Paradox: What probability should Sleeping Beauty assign to heads or tails upon awakening (1/2 or 1/3)?
- Theoretical Deconstruction:
  - Category Mistake: Conflating the Content() domain of "probability entity":
    - Level 1 (prior probability): Content(probability) = coin's physical state ⇒ P(heads) = 1/2.

 doi:10.20944/preprints202506.2384.v6

- Level 2 (information update): Content(probability) = (coin state, awakening frequency) ⇒ P(heads | awakening) = 1/3.
- Uniqueness Ruling:
  - If asking "probability of the coin's actual state" → Content(Level 1) ⇒ 1/2.
  - If asking "probability given current awakening" → Content(Level 2) ⇒ 1/3.
- Contradiction Root: Treating two distinct probability levels Content(Level 1) ≠ Content(Level 2) as the same question.

3.14. Pascal's Wager in Modern Contradiction

- Problem: If multiple religions' gods claim "I alone am true," how does a rational person bet?
- Theoretical Deconstruction:
  - Category Mistake: Conflating Content(god) domains:
    - Level 1: Content(god X) = divine description in a specific religion's doctrine.
    - Level 2: Content(omnipotent entity) = abstract supreme being beyond specific doctrines.
  - Ruling:
    - If comparing the truth of specific religious gods → each Content(god X) differs ⇒ distinct levels.
    - If asking "does a supreme entity exist" → define Content(omnipotent entity) independently of specific religions.

3.15. Unexpected Execution Paradox

- Problem: A judge announces, "You will be executed unexpectedly next week." The prisoner deduces it cannot happen, yet the execution occurs.
- Theoretical Deconstruction:
  - Category Mistake: Swapping Content(unexpected) from "prisoner's cognitive state" to "objective time point."
  - Correct Definition: Content(unexpected) = prisoner's inability to be certain of execution the day before.
  - Conclusion: The execution day inevitably exists (due to objective time flow), while Content(unexpected) depends solely on the prisoner's cognitive state, belonging to different levels.

4. Applications

4.1. The Dilemma of Personal Identity and Existing Theories

The core question of personal identity is: What makes a person the same over time? Traditional theories primarily revolve around debates between physical continuity (e.g., brain continuity) and psychological continuity (e.g., coherence of memory and personality). Among these, Derek Parfit's (1984) highly influential reductionist psychological continuity theory reduces personal identity to overlapping chains of psychological connectedness (such as memory, personality, and intentions). This theory demonstrates remarkable explanatory power when addressing dynamic changes, such as gradual cellular replacement or slow personality transformations. It successfully shows that the persistence of personhood is not an "all-or-nothing" metaphysical fact but rather a matter of degree.

However, Parfit's theory, as well as competing physical continuity theories, implicitly presuppose a more fundamental and unarticulated premise: How do we determine, at a given time-slice, that an entity is a "person," and how do we statically compare them across different time-slices and possible worlds? In other words, these theories excel at answering "Why is he still himself?" (the problem of dynamic persistence) but neglect to define "What exactly is 'he' at time t?" (the problem of static identity). This static "what" is a prerequisite for discussing any dynamic "persistence."

This weakness is exposed in Bernard Williams' (1970) famous "fission" thought experiment. When a person splits into two psychologically continuous successors, physical continuity theories collapse because they cannot handle "one dividing into two." Meanwhile, Parfit's psychological

 doi:10.20944/preprints202506.2384.v6

continuity theory faces a dilemma: If identity is considered intransitive (B and C, both identical to A, are not identical to each other), it violates logic; if the original individual is considered to have perished after fission, it contradicts the core claim that "psychological continuity is sufficient for identity." Through this experiment, Williams powerfully demonstrates that without a clear criterion for static identity, any discussion of dynamic continuity will descend into conceptual confusion.

4.1.1. Space

This paper argues that the common root of these dilemmas lies in the fact that existing theories attempt to treat "person" as a primitive concept defined by specific physical substrates or historical causality. They mistakenly allow the properties of the "carrier" (the biological brain) or "history" (causal chains) to intrude into the judgment of the identity of the "person entity" itself, which is a category error. The debate between Parfit and Williams is essentially a conflict between two different definitions of the domain of the Content() function—one being the stream of psychological attributes, the other being the history of the physical carrier. However, neither side recognizes this, leading to an unresolvable deadlock.

An Analytical Framework Based on Information Content

According to the two axioms of Steins theory, we propose a minimal assumption: The necessary and sufficient condition for the identity of a conscious entity lies in the identity of its core information content (Content). This first provides a clear criterion for resolving static identity.

Formally, let:

· C be a conscious time-slice.

· We define it as an ordered pair: C = Content(C, $\mathcal{S}$)

  · Content(C): represents the consciousness at that time-slice.

  · $\mathcal{S}$: represents the carrier that instantiates this information structure (e.g., a specific brain).

Based on this, for any two conscious stages C_1 = Content(C, $\mathcal{S}_1$) and C_2 = Content(C, $\mathcal{S}_2$), we have:

$$\text{Content}(C) \equiv \text{Content}(C) \Rightarrow C_1 \equiv C_2$$

This means that as long as the information content at two time-slices is the same, they are different instances of the same consciousness, regardless of whether their spatiotemporal coordinates $\mathcal{S}$ are continuous.

Thus, we provide a clear resolution to Williams' fission experiment: The paradox arising from the two successors C_b and C_c is due to the erroneous requirement that dynamic continuity must map to a one-to-one physical path. Under this framework, we need only compare static content: If Content(C_b) = Content(C_a) and Content(C_c) = Content(C_a), then according to n=n, both C_b and C_c are identical to C_a as the same consciousness. This is not a logical contradiction but rather the simultaneous instantiation of the same information entity at multiple coordinates. Parfit's intuition—"survival without identity"—is realized here in the strictest sense: Survival is the identity of information content itself, and it does not require an additional "identity" guaranteed by carrier history.

Therefore, this theory does not entirely negate Parfit's psychological continuity theory but rather provides it with a solid foundation. Dynamic psychological continuity can be seen as a special case where Content(C) remains highly similar (though not absolutely identical) across the time dimension. The advantage of this framework is that it first clearly defines what "static identity" is, thereby allowing the discussion of "dynamic persistence" to proceed on a firm logical basis.

4.1.2. Time

Based on the axiomatic system established earlier, we can draw a revolutionary conclusion about the existence of consciousness in the time dimension: Your "present" is (C, $\mathcal{S}$_now). Your "past" is (C, $\mathcal{S}$_past). Your "future" is (C, $\mathcal{S}$_future). They are all specific instantiations or manifestations of the same conscious information structure C at different spatiotemporal coordinates $\mathcal{S}$. Therefore, the "self" you are experiencing at this moment is, in the strictest sense of absolute identity, the same as

 doi:10.20944/preprints202506.2384.v6

the "you" in the past and future, because "you" refers to that abstract information pattern C, not the transient and ephemeral coordinate-bound state (C, $\mathcal{S}$).

Let us conduct an extreme thought experiment. Suppose at time point t_1, there exists a specific instantiation of a conscious information structure C, whose complete state is uniquely determined by the instantaneous information pattern it contains, denoted as Content(C). Now, imagine that in the distant future, at time point t_2 ($t_2 >> t_1$), an entirely identical nervous system (whether through natural Poincaré recurrence, an extreme coincidence of quantum fluctuations, or some yet-unknown cosmic recurrence mechanism) is instantaneously assembled and activated. The instantaneous information pattern it produces, Content(C), is bit-for-bit identical to Content(C) in the information-theoretic sense.

According to our Axiom 1 (n = n) and Axiom 2 (uniqueness theorem), we must conclude:

$$\text{Content}(C) = \text{Content}(C) \Rightarrow C \equiv C$$

This means that the consciousness C at time t_2 is the same conscious entity as the consciousness C at time t_1. This is not a "copy" or "rebirth" but a direct recurrence of identity at different coordinates. The billions of years of spatiotemporal gap between them are entirely irrelevant to determining whether they are the same entity. What connects them is not a fragile and defeasible thread of "psychological continuity" but the ironclad law of logical identity. The interval of time carries zero weight in this judgment.

Now, let us push this thought experiment to another extreme. Suppose at time point t_1, a conscious activity "a" has just begun its neural computation process. Within an extremely short time Δt, before the activity of the first nervous system is completed (i.e., before Content(a) is fully manifested), in another corner of the universe, another nervous system with an identical physical structure is activated and begins executing an identical computation process, thereby producing an identical conscious activity "a."

At this point, we have two concurrent conscious processes:

· Process P_1: At coordinate $\mathcal{S}_1$, starting at time t_1, ongoing.
· Process P_2: At coordinate $\mathcal{S}_2$, starting at time $t_1 + \Delta t$, ongoing.

When we examine the states of these two processes at equivalent progress points, we find that since they are executing the exact same "algorithm," the information patterns Content(P_1_slice) and Content(P_2_slice) at any equivalent progress point are indistinguishable. However, they did not start simultaneously, meaning that equivalent progress points occur at different times. Therefore, within the category of the P information pattern, time does not serve as a valid criterion for identity.

According to our axioms, we again conclude:

$$\text{Content}(P_1\_\text{slice}) = \text{Content}(P_2\_\text{slice}) \Rightarrow P_1\_\text{slice} \equiv P_2\_\text{slice}.$$

This means that at the level of information patterns, what we observe is not two consciousnesses but one consciousness appearing simultaneously at two different time points. It is not that there are two "yous" thinking, but rather that the thinking process of "you" is executed and presented by two physical systems at different time points.

From this, we arrive a counterintuitive but logically inevitable conclusion: Identity is non-continuous in time and non-local in space. The way anything persists is not like a continuous "river" but more like a series of discrete, absolutely identical "state flashes." The continuity we perceive is a cognitive illusion produced by the rapid succession of these highly similar, causally connected state flashes (produced by the same brain). Once we introduce a physically separate but informationally identical instance, the illusion of continuity is broken, revealing its underlying discrete and separable nature.

Therefore, your "present" is (C, $\mathcal{S}$_now). Your "past" is (C, $\mathcal{S}$_past). Your "future" will be (C, $\mathcal{S}$_future). They are all "manifestations" or "slices" of the same abstract conscious information pattern C in different spatiotemporal coordinate blocks. What you are experiencing now is, in the strictest sense of absolute identity, the same as the "you" in the past and future, because "you" refers to C, not

 doi:10.20944/preprints202506.2384.v6

the transient and volatile coordinate-bound state (C, $\mathcal{S}$). Time does not divide you; it merely provides the coordinates for your manifestation.

4.2. "Spacetime Leap" as a Logical Necessity of Coordinate Decoupling and Rebinding

Before exploring the "spacetime leap" of consciousness entities, we must first pay the highest tribute to modern physics, particularly Einstein's special and general relativity. These theories, with their unparalleled precision and beauty, describe the profound dynamical relationship between mass, energy, and spacetime, strictly setting the causal limit of any physical signal or entity motion—the speed of light. Any attempt to achieve a "spacetime leap" at the physical level, whether via wormholes, warp drives, quantum suicide, or other exotic mechanisms, must be tested within the robust frameworks of relativity or quantum mechanics, facing immense physical challenges like energy conditions, singularities, and empirical verification.

However, the "spacetime leap" argued in this paper is fundamentally distinct from such physical processes. It is not a motion process within spacetime governed by physical laws but a logical necessity derived from Steins' axiomatic system. It addresses a more primitive question: "Are two information entities with identical content, instantiated at different spatiotemporal coordinates, the same entity?" The answer depends not on connecting physical paths but solely on the logical axiom n = n.

Traditional physics (including relativity) studies the "spacetime leap" as the "binding state of consciousness information and spatiotemporal coordinates," i.e., (C, $\mathcal{S}$). Physics perfectly describes how such binding states evolve over time, i.e., (C_t, $\mathcal{S}$_t) → (C_{t+dt}, $\mathcal{S}$_{t+dt}), governed by elegant differential equations (e.g., Einstein's field equations).

This theory, however, focuses on a possibility physics naturally overlooks due to its paradigmatic constraints: the decoupling and rebinding of a consciousness information C with its coordinates $\mathcal{S}$.

1. Decoupling: (C, $\mathcal{S}_1$) → (C), ($\mathcal{S}_1$). Physically, this may correspond to the carrier (e.g., brain) being destroyed by an event consistent with the event horizon principle (e.g., a velocity differential causing causal isolation at the neuronal level), rendering the information structure uninstantiated ($\mathcal{S}$ = ∅).
2. Rebinding: (C) → (C, $\mathcal{S}_2$). Physically, this may correspond to an instantaneous "reappearance" elsewhere (e.g., via Poincaré recurrence, MWI parallel universes, or bubble universes).

Crucially, by Axiom 1 (n = n), the decoupled abstract consciousness information C retains its self-identity. Thus, the new state after rebinding (C, $\mathcal{S}_2$) and the old state before decoupling (C, $\mathcal{S}_1$), sharing the same C, are necessarily different instances of the same information entity. This is the logical core of the "spacetime leap": it is not "travel" through spacetime but the "realization" or "manifestation" of identity at different locations.

Thus, the relationship between this theory and traditional physics is not competitive but complementary and foundational:

- Physics studies the continuous evolution of the binding state (C, $\mathcal{S}$) within spacetime, asking "How does one get from A to B?"
- This Theory studies the discrete identity logic of the C entity transcending spacetime, asking "Are A and B the same thing?"

Relativity prohibits superluminal motion of physical entities but is entirely silent on, and need not address, the "realization" of a logical entity at different spacetime points. The "spacetime leap" seems "inconceivable" or "physics-violating" precisely because we erroneously apply physical laws governing binding state motion to judge a logical theorem about information entity identity—a category mistake.

Conclusion: The "spacetime leap" proposed here is not a physical conjecture awaiting realization but a logical inference already established. Derived from the fundamental law of identity, it offers a novel picture of personal identity: the persistence of consciousness lies in the identity of its information pattern, not in the continuity of physical processes connecting its instances. This provides

 doi:10.20944/preprints202506.2384.v6

an unprecedentedly clear framework for understanding thought experiments like teleportation and brain-in-a-vat, liberating the discussion of personal identity from physical constraints and grounding it in a more fundamental logical and metaphysical foundation.

Possible Method: Self-Envating (entering a room isolated from the environment to be changed, see Section 3.8) → Detector scans the isolated environment (e.g., AI retrieval) → Judgment (if requirements are met, terminate; if not, proceed) → Rapid destruction for decoupling (e.g., explosives, survivor effect, see Section 4.2) → Cosmic randomness for reappearance (see Section 4.2). Repeatable for multi-stage leaps.

4.3. Ethical Dilemmas and Existing Theories

Since the inception of ethics, generations of profoundly insightful philosophers—from Kant's grand a priori framework to Mill's intricate consequentialist calculus—have constructed a magnificent ethical edifice. These remarkable efforts share a profound and admirable ambition: to seek a stable, transcendent metaphysical foundation for moral judgments beyond individual perspectives. This foundation is typically conceived as: (a) a moral reality independent of our cognition; (b) a self-identity persisting through time as the anchor of responsibility; and (c) a sacred "God's-eye view" to adjudicate actions' value from absolute impartiality. This pursuit of universality and objectivity is undoubtedly one of philosophy's most glorious achievements.

Within this tradition, Bernard Williams' (1973) discussion of "moral luck" reveals, with striking acuity, the subtle rift between the control principle and our moral intuitions, greatly enriching our philosophical imagination. The "undiscovered betrayal" thought experiment, with its logical purity, pushes traditional theories to the limits of their explanatory power, serving as a "touchstone" for testing their core. Facing this challenge, traditional theories like Kantian ethics display unparalleled rigor, resolutely defending the absoluteness of moral wrongness, even if their arguments require appealing to a "moral law" beyond experience—a steadfast commitment to universality that commands respect. Similarly, some utilitarian approaches attempt to resolve the dilemma through a global calculation by an "ideal observer," showcasing ambitious and systematic grandeur.

As Derek Parfit (1984) incisively notes, such approaches may create tension with individuals' first-person perspectives at the motivational level, but this is not a flaw in the theories themselves; rather, it highlights the tragic yet heroic tension inherent in humanity's rational pursuit of moral sublimity.

This paper, standing on the shoulders of these giants with utmost respect, seeks to inherit and advance this glorious tradition. We fully endorse the traditional pursuit of objectivity and universality. However, we propose that this lofty goal may be achieved through a more direct, frictionless path. The reason traditional frameworks generate troubling tensions in boundary cases may lie in a methodological over-indirectness: attempting to mediate and regulate moral life, which fundamentally arises from first-person experience, through a hypothetical, transcendent third-party categorical system.

This paper explores a complementary path. We are delighted to find that, through Steins' axiomatic system and the profoundly illuminating observational realism reinforced by the "brain in a vat" thought experiment, we can pay tribute to and achieve the core goals of traditional ethics in a novel way. Our central thesis is that the sought-after objectivity and universality of ethical value need not be guaranteed by a "God's-eye view"; rather, they can be more robustly grounded in the first-person facts of consciousness systems' observational experiences. The boundaries of moral concern thus perfectly and logically coincide with the boundaries of conscious experience, achieving the universality sought by traditional theories in an unexpected way.

This study aims to show that we do not seek to negate prior work but to realize their shared aspirations through a more precise metaphysical foundation, dissolving unnecessary philosophical anxieties arising from methodological indirectness.

4.3.1. An Information-Based Content Analysis Framework: Advancing Traditional Goals

 doi:10.20944/preprints202506.2384.v6

The "brain in a vat" thought experiment, with its unparalleled philosophical value, challenges our naïve notions of "reality." It compels us to acknowledge a highly productive principle: for any consciousness system, its operational reality is its observational experience flow. Whether an external simulator exists is empirically undecidable and redundant.

Combining this profound insight with Steins' axioms, we derive a cornerstone principle for ethics, a more precise formulation of the traditional pursuit of objectivity in the information age: For any consciousness system C, an event E is ethically significant if and only if its consequences (directly or indirectly) are reducibly embodied as a specific impact on system C's observational experience (i.e., its information pattern Content(C)).

Corollary 4.3.1 (Ethical Relevance Criterion): If an event E's occurrence produces no discernible difference in the past, present, or future observational experience flow of system C, then, in ethical considerations regarding C, event E does not constitute a relevant fact and thus carries zero weight in ethical evaluation.

4.3.2. Core Derivation: A Dissolving Analysis of Traditional Dilemmas

With respect, let us restate the "undiscovered betrayal" case within this framework:

Suppose two possible worlds: W_1 (event E occurs: partner betrays) and W_2 (event E does not occur). By the strict setup of the thought experiment, the victim's (as consciousness system C) entire experiential information flow is completely indistinguishable in both worlds. That is: Content(C_W_1) = Content(C_W_2).

By Steins' Axiom 2 (Distinctness), we obtain: Content(C_W_1) = Content(C_W_2) ⇒ $C\_W_1 \equiv C\_W_2$. This means that in both worlds, there exists the same conscious experiential subject C.

Now, for ethical judgment: The direct object of ethical concern is the experiential well-being of consciousness system C. Since C's experiences are identical in both worlds, the two worlds are ethically equivalent for C.

Thus, event E (the betrayal) has zero impact on C's experiential information flow and does not constitute a variable in ethical evaluation regarding C. It neither causes harm nor constitutes betrayal, as these ethical concepts are operationally defined as specific negative information states in the experiential flow, which do not occur.

Conclusion: Moral wrongness is not mysteriously attached to actions themselves but systematically and verifiably tied to their specific impact patterns on the experiential information flow of consciousness systems. Absent such observable impact, the action is ethically irrelevant.

4.3.3. Implications of the New Framework: Pursuit of Harmony and Stability

If morality is not about inaccessible "external truths," what is it about? This framework proposes that the core purpose of ethics can be operationally defined as maintaining and promoting the stability and harmony of internal information patterns in consciousness systems. This definition echoes the goals of many traditional virtue ethics but rests on a clearer foundation.

- Moral Action: Any action or rule that produces, sustains, or optimizes harmonious information patterns within individual consciousness systems.
- Immoral Action: Any action or rule that triggers, disrupts, or leads to chaotic or painful states in individual consciousness systems' information patterns.

Final Corollary: A Paradigm's Precision

This paper argues that, based on information pattern identity and observational realism, ethics can undergo a foundational precision.

Ethics can shift its ambition from an unattainable "God's-eye view" to its only effective domain: the first-person facts of conscious experience. An action's goodness or badness depends not on its properties in a (possibly hypothetical) "objective world" but entirely on how it predictably shapes our only accessible reality—our experiential flow.

 doi:10.20944/preprints202506.2384.v6

Thus, in strict ethical derivation, events unobservable by any consciousness system are assigned no value in ethical computation. The boundaries of moral concern are the boundaries of conscious experience. This framework does not excuse immoral actions—on the contrary, by anchoring responsibility firmly in observable impacts, it provides a more robust, clearer, and undeniable foundation for moral responsibility: we bear sole and total responsibility for our own peace and prosperity. (References 32, 33, 34, 35, 36, 37)

5. Conclusion

1. Unity of Absoluteness and Relativity: Identity is absolutely guaranteed by n = n, but its judgment n ≡ m is relative to the chosen Content() function. Any claim introducing external properties commits a category mistake.

2. Indescribability: Under the same Content() definition, no one can distinctly describe two entities with identical content.

3. Fundamentality: This theory is the most foundational metaphysical framework; even in a fully random universe without causality, world ≡ world still holds.

4. Explanatory Power: It successfully subsumes Leibniz's and Kripke's theories as special cases and resolves modern scientific challenges they cannot address, particularly the problem of quantum identical particles.

References

1. Leibniz, G. W. (1714). Principle of the Identity of Indiscernibles (in Monadology).
2. Frege, G. (1892). Über Sinn und Bedeutung (On Sense and Reference).
3. Floridi, L. (2011). The Philosophy of Information. Oxford University Press.
4. Russell, B. (1905). On Denoting. Mind.
5. Quine, W. V. O. (1950). Identity, Ostension, and Hypostasis. The Journal of Philosophy.
6. Shannon, C. E. (1948). A Mathematical Theory of Communication. The Bell System Technical Journal.
7. Chalmers, D. J. (2018). The Meta-Problem of Consciousness.
8. Putnam, H. (1967). The Nature of Mental States.
9. Everett, H. III (1957). "Relative State" Formulation of Quantum Mechanics. Reviews of Modern Physics, 29(3), 454–462.
10. Linde, A. (1986). Eternally Existing Self-Reproducing Chaotic Inflationary Universe. Physics Letters B, 175(4), 395–400.
11. Tononi, G. (2004). An Information Integration Theory of Consciousness. BMC Neuroscience, 5(1), 42.
12. Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). Principles of Neural Science. McGraw-Hill.
13. Tegmark, M. (1998). The Interpretation of Quantum Mechanics: Many Worlds or Many Words? Fortschritte der Physik, 46(6–8), 855–862.
14. Poincaré, H. (1890). Sur le problème des trois corps et les équations de la dynamique. Acta Mathematica, 13, 1–270.
15. Zurek, W. H. (2009). Quantum Darwinism. Nature Physics, 5(3), 181–188.
16. Zurek, W. H. (2003). Decoherence, Einselection, and the Quantum Origins of the Classical. Reviews of Modern Physics, 75(3), 715–775.
17. Black, M. (1952). The Identity of Indiscernibles. Mind, 61(242), 153–164.
18. Chandler, H. S. (1975). Rigid Designation. *The Journal of Philosophy, 72(13), 363–369.
19. French, S., & Redhead, M. (1988). Quantum Physics and the Identity of Indiscernibles. The British Journal for the Philosophy of Science, 39(2), 233–246.
20. Kripke, S. A. (1980). Naming and Necessity. Harvard University Press.
21. Leibniz, G. W. (1714/1898). The Monadology and Other Philosophical Writings. Trans. Robert Latta. Oxford University Press.
22. Putnam, H. (1975). The Meaning of Meaning. Minnesota Studies in the Philosophy of Science, 7, 131–193.
23. Wiggins, D. (1980). Sameness and Substance. Harvard University Press.

 doi:10.20944/preprints202506.2384.v6

24. Heller, M. (1984). Temporal Parts of Four-Dimensional Objects. Philosophical Studies.
25. Sider, T. (2001). Four-Dimensionalism: An Ontology of Persistence and Time. Oxford University Press.
26. Sider, T. (1996). All the World's a Stage. Australasian Journal of Philosophy.
27. Hawley, K. (2001). How Things Persist. Oxford University Press.
28. Muller, F. A., & Saunders, S. (2008). Discerning Fermions. The British Journal for the Philosophy of Science.
29. Krause, D. (2011). Logical Aspects of Quantum Non-Individuality.
30. Parfit, D. (1984). Reasons and Persons. Oxford University Press.
31. Williams, B. (1970). The Self and the Future. The Philosophical Review.
32. Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. The Quarterly Review of Biology, 46(1), 35–57.
33. Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. Science, 211(4487), 1390–1396.
34. Hamilton, W. D. (1964). The Genetical Evolution of Social Behaviour. I & II. Journal of Theoretical Biology, 7(1), 1–52.
35. Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. Science, 314(5805), 1560–1563.
36. Street, S. (2006). A Darwinian Dilemma for Realist Theories of Value. Philosophical Studies, 127(1), 109–166.
37. Fehr, E., & Fischbacher, U. (2003). The Nature of Human Altruism. Nature, 425(6960), 785–791.