Essay

Steins Theory: A New Axiomatic System Concerning Identity

Jiaqi Guo *

Posted Date: 4 September 2025

doi: 10.20944/preprints202506.2384.v3

Keywords: informational uniqueness; identity; content identity; n = n; n ≠ m; formal axioms; indiscernibles; spatiotemporal jumping

 doi:10.20944/preprints202506.2384.v3

Essay

Steins Theory: A New Axiomatic System Concerning Identity

Jiaqi Guo

Independent Researcher, China; 1028165464@qq.com

Abstract

In the philosophy of language, Frege's (1892) distinction between sense and reference provided the foundational framework for identity statements, while Putnam's (1975) 'Twin Earth' thought experiment, with its remarkable insight, pushed externalism to its extreme, successfully challenging the internalist model of meaning and setting the agenda for subsequent decades of debate on the determination of reference. However, despite the highly illuminating nature of these seminal works, an intriguing phenomenon is that the ensuing debates—such as those surrounding core cases like the Ship of Theseus and identical particles—seem to have reached an impasse. This paper argues that this impasse may not stem from the depth of the problem itself, but rather from an unexamined deep presupposition shared by these otherwise highly persuasive theories: namely, the belief that there exists some single, decisive level (whether microphysical structure or historical causation) that can definitively answer the question of identity once and for all. This paper proposes that instead of continuing to seek a superior single answer under this presupposition, a more productive approach might be to question the presupposition itself. To this end, we develop a level-relativity analytical framework (Steins Theory). Interestingly, this framework demonstrates that the aforementioned seemingly opposing outstanding theories can actually be understood as special cases of this framework at different levels; the difficulties they encounter become inevitable precisely when they attempt to make assertions across levels. Therefore, this framework does not aim to negate previous work, but rather seeks to clarify its valid scope of application, thereby offering a new path to resolving a series of philosophical puzzles arising from level confusion.

Keywords: informational uniqueness; identity; content identity; n = n; n ≠ m; formal axioms; indiscernibles; spatiotemporal jumping

1. Introduction

The problem of identity, that is, "what makes a thing what it is," is central to metaphysics and logic. From Leibniz's (1714) profoundly insightful 'Principle of the Identity of Indiscernibles' (PII)—which, with its logical simplicity and power, set a lofty ideal for the individuation of substances—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 solid foundation for the stability of reference. These outstanding efforts by generations of philosophers have collectively built the grand intellectual tradition through which we understand the persistence and recognition of individuals.

However, an intriguing phenomenon is that these highly persuasive theoretical frameworks within their respective domains often exhibit a regrettable, systematic limitation when confronting complex boundary cases posed by reality and thought experiments. Leibniz's strong PII principle faces fundamental difficulties perhaps unforeseen in its original conception when confronted with quantum identical particles (French & Redhead, 1988). Its inherent strictness paradoxically leads it into an impasse when dealing with multiple entities that are numerically distinct yet completely indistinguishable in all intrinsic properties (such as two electrons). Similarly, Kripke's sophisticated rigid designation theory, aimed at anchoring reference, becomes entangled and unclear in its

 doi:10.20944/preprints202506.2384.v3

explanatory power when dealing with diachronic changes in intrinsic properties (like the Ship of Theseus) (Chandler, 1975), let alone its difficulty accommodating the phenomena of identical particles in quantum mechanics, which challenge the classical view of individuals.

Thus, we face a peculiar intellectual impasse: whether it is Leibniz's strong program pursuing property identity or Kripke's historical path focused on the necessity of origin, each seems to successfully illuminate one wing of the edifice of identity, only to cast the other wing into deeper shadow. Is their shared, perhaps unspoken, ambition—to find a single, absolute criterion for identity—itself constituting an a priori obstacle preventing us from truly understanding the multidimensional nature of the problem?

This paper argues that the key to breaking this impasse lies not in making an either/or choice among existing paths or engaging in another round of patching. Instead, it requires us to step back and engage in a meta-level reflection on the problem itself. This paper aims to propose a level-relativity framework for analyzing identity (Steins Theory). The ambition of this framework is not to completely negate predecessors—indeed, Leibniz's focus on properties and Kripke's insistence on history will be repositioned within the new framework and accorded their limited legitimacy—but rather to demonstrate that the root of the aforementioned difficulties lies in a confusion of levels: the mistaken attempt to answer a question posed at one level with an answer from another level.

By introducing a relativity framework defined by the level of the Content() function, this paper aims to transform classic ontological and semantic disputes into a clear, operational question of conceptual choice. Thus, rather than solving these puzzles, it dissolves them, offering a new way out for a series of philosophical anxieties born of category mistakes.

Note: In the absence of explicit premises, this paper defaults to an objective scientific perspective (e.g., physics and other sciences) for discussion.

2. Analysis

2.1.1. Axiom 1 (Self-Identity)

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

(An information entity is necessarily identical to itself; this is the foundation of logical reference.)

2.1.2. Axiom 2 (Mutual Distinctness)

$$\forall n, m \in I,\ (n \not\equiv m) \Leftrightarrow (\ \mathrm{Content}(n) \neq \mathrm{Content}(m)\ )$$

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

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

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

2.2. Illegal Expansion and Level Confusion: The Case of the Ship of Theseus

In addressing the enduring puzzle of the Ship of Theseus, an influential and intuitively appealing solution has been proposed, championed by figures like David Wiggins (1980). This solution argues that the identity of an object is not guaranteed by its instantaneous properties at any given moment, but must be carried by its spatiotemporal continuity and an uninterrupted historical-causal path. The great advantage of this approach is that it successfully captures our deep intuition about the concept of "an object persisting through time"—that things are not momentary existences but have a life history or biography.

Another solution, Perdurantism or Four-Dimensionalism (see Heller, 1984; Sider, 2001), offers a radically different picture with considerable metaphysical simplicity. This theory, with its thorough clarity, skillfully avoids many pitfalls of diachronic identity. Four-dimensionalism holds that the Ship of Theseus is not a three-dimensional entity "wholly present" at each time, but a "spacetime worm" extended in four-dimensional spacetime. Each time-slice of the ship is considered a temporal part of

 doi:10.20944/preprints202506.2384.v3

this four-dimensional object. Thus, so-called "change" merely means that this four-dimensional whole possesses different properties (like different planks) at different temporal parts. Within this framework, the original Ship of Theseus (as a four-dimensional entity), the replaced ship, and the ship reassembled from the old planks are three distinct four-dimensional objects. They might share identical three-dimensional cross-sections at some temporal segment (and thus be indistinguishable at that moment), but as wholes, they are naturally different. The excellence of this solution is that it transforms the problem of persistence from the troublesome issue of "identity" into the relatively clearer relation of "part-whole."

Closely related to four-dimensionalism is Stage Theory (see Sider, 1996; Hawley, 2001), which retains many advantages of four-dimensionalism while attempting to better accommodate our everyday linguistic intuition that "objects are three-dimensional." Stage theorists argue that what we typically refer to as "the Ship of Theseus" denotes not the entire four-dimensional worm, but a stage or time-slice at a specific point in time. When we say at time t that "the ship is the same," we are actually saying that between the stage at t and the stage at an earlier time t_1 , there exists a (primitive) counterpart relation, sustained by similarity and causal continuity. Stage theory, with its conceptual economy, avoids presupposing identity across time, thereby exhibiting strong theoretical appeal.

However, despite the internal ingenuity and frequent self-consistency of the theories outlined above, this paper will argue that they share a common and profound dilemma at the normative level. Whether appealing to a historical path, a four-dimensional whole, or counterpart relations, these theories all attempt to provide a single, absolute criterion for identity determination. To achieve this goal, they must construct "time" or "spatiotemporal location" itself as a constitutive element of an object's identity. This means that within their theoretical frameworks, answering the question "Is the ship at t_1 identical to the ship at t_2 ?" logically must depend on examining spatiotemporal coordinates or cross-temporal relations.

This paper contends that it is precisely this crucial theoretical move that inadvertently leads to a "category mistake." Let us formally reconstruct the schemes of these theories: they effectively employ a content function whose domain is defined as Content_Hist(Ship) = (Material, Form, Function, Historical-Causal Path / Four-Dimensional Whole / Counterpart Relation) . This paper argues, however, that the initial identity question—"Is this ship, identical in function and structure, still itself?"—presupposes a content function that should only be Content_Struct(Ship) = (Material, Form, Function) .

We can formalize this reconstruction as follows:

Initial Question: Judge whether the entity "Ship" at time point t_1 and t_2 is identical.

Correct Domain of Content() (according to the initial question): Should only include attributes related to the abstract informational structure of "Ship," i.e., Content_Struct(Ship) = (Material, Form, Function) .

Illegitimate Expansion by Historical/4D/Stage Theories:

Wiggins effectively uses Content_Hist(Ship) = (Material, Form, Function, Historical-Causal Path) .

Four-dimensionalism effectively uses Content_4D(Ship) = (Material, Form, Function, Spatiotemporal Coordinates) .

Stage theory effectively uses Content_Stage(Ship) = (Material, Form, Function, Counterpart Relation) .

The resulting process can be formalized as:

They attempt to answer the identity question Q_Struct based on Content_Struct() : "Is Ship at t1 ≡ Ship at t2 ?"

However, the judgment function they actually use is Content_Hist() , Content_4D() , or Content_Stage() .

Since the replacement material is identical to the original, we have Content_Struct(Ship_t1) = Content_Struct(Ship_t2) .

 doi:10.20944/preprints202506.2384.v3

But because the historical path, spatiotemporal coordinates, or counterpart relation has changed, Content_Other(Ship_t1) ≠ Content_Other(Ship_t2) .

Consequently, they conclude Ship_t1 ≢ Ship_t2 in answer to Q_Struct .

Thus, we see an interesting situation: the theories effectively answer a question, but likely not the one originally posed. They precisely answer "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 then take this answer as the ultimate verdict on the question "Is the ship structurally identical?" This is akin to a judge being asked "Does the defendant's action conform to Criminal Law Article X?" but rendering a verdict after consulting Civil Law statutes. The conclusion might be coherent within its own system, but it has quietly shifted the ground of the debate.

Therefore, the true contribution of the historical path theory might lie in its excellent revelation of how attributes like "history" or "spatiotemporal whole," as powerful explanatory attributes, influence our identity judgments. Its limitation, however, is that it attempts to elevate this explanatory attribute into a metaphysical necessity, thereby forcing an expansion of the criteria for judging identity to maintain its theoretical integrity. The value of the Steins theory lies in its avoidance of this difficult expansion. Instead, by clarifying the levels of the question, it allows Content_Struct() , Content_Hist() , Content_4D() , and Content_Stage() to each provide valid and non-conflicting answers to different questions. It does not solve these puzzles; it dissolves them.

2.3. Information Conservation

2.3.1. The Indistinguishability of Identical Particles in Quantum Mechanics Poses the Most Severe Challenge to Leibniz's PII (Citations 1,2), Yet Provides a Natural, Physically Substantiated Model for the Present Theory

Current philosophical discussions on identical particles, facing the challenge quantum particles pose to PII, are mainly divided into two camps: revisionism and revolutionism. The former attempts to salvage some form of the principle of individuality, while the latter abandons individuality itself.

Saunders's proposal undoubtedly represents one of the most ingenious and technically rigorous attempts within the revisionist path. By ingeniously defining 'weak discernibility,' he successfully liberates the discussion from the dead end of intrinsic properties, offering an insightful perspective on finding the cornerstone of individuation in relationality. The complexity of this approach and the wide discussion it has sparked attest to its profound philosophical value. However, this very technical complexity exposes a potential cost of the solution: its definition of 'purely extrinsic relational properties,' while striving for precision, inevitably introduces considerable terminological vagueness, such that its defenders must carefully guard against accusations of circularity (Muller & Saunders, 2008). More fundamentally, the entire theoretical edifice rests on a troubling presupposition: that the 'individuality' of identical particles must and can only be 'saved' by finding some (even if relational) individuating property. This renders his theoretical effort—however ingenious—essentially an ad-hoc patchwork aimed at rescuing a premise. When applied to states in quantum field theory with indefinite particle number, this strategy of constantly introducing new relational properties to salvage individuality becomes increasingly ad-hoc: it appears less like an elegant deduction of the theory itself and more like an increasingly costly price paid to maintain the theoretical premise (that individuality must exist).

Faced with the difficulties of revisionism, another revolutionary approach chooses a more radical path. Scholars like Décio Krause (2011) propose a highly disruptive thesis: quantum particles might simply not be "individuals" in the traditional metaphysical sense. Therefore, laws based on individual identity apply the wrong category from the outset. They should be understood as "non-individuals," requiring highly specialized mathematical tools like quasi-set theory for their description.

Krause's approach is striking for its conceptual thoroughness and consistency; it unreservedly embraces quantum mechanics' most counterintuitive features, decisively breaking with our entire

 doi:10.20944/preprints202506.2384.v3

classical framework concerning objects and spatiotemporal location. This resolute stance is undoubtedly clean and efficient theoretically. However, the corresponding cost of this efficiency is that the concept of "non-individual" itself imposes a considerable explanatory burden metaphysically, requiring us to abandon a deeply ingrained intuitive understanding of what "a thing" means.

2.3.2. Steins Theory's Solution: A Level-Relativity Framework

The two approaches above share a deep misconception: they both attempt to find an answer to a wrongly posed question. The question is not "What is the correct individuating property?" but rather "At what level are we asking the identity question?"

Steins Theory provides a meta-framework for this by introducing the leveled 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 we ask the identity question at the level Content(Particle) = $\mathcal{I}$, i.e., comparing only intrinsic properties, all identical electrons have Content(electron) = (mass m_e, charge -e, spin 1/2...) . According to Axiom 2, Content(n) = Content(m) ⇒ n ≡ m. At this level, they are indeed the same informational entity e. This explains the root of their indistinguishability.

When we ask the identity question at the level Content(Particle State) = ($\mathcal{I}$, $\mathcal{S}$) , since $\mathcal{S}$ (e.g., position, momentum) necessarily differs, Content(P_1) ≠ Content(P_2) , therefore they are different particle states. This explains why we observe multiple scattering events in experiments.

Therefore, the perplexity brought by quantum identical particles stems from erroneously allowing the distinctness at the $\mathcal{S}$ (spatiotemporal coordinate) level to illegitimately intrude upon the judgment of identity at the $\mathcal{I}$ (intrinsic entity) level. Steins Theory resolves the contradiction by clearly distinguishing these two levels: they are both "one" (as abstract informational entities) and "many" (as manifestations in concrete spacetime). Particle annihilation and creation merely represent the decoupling and re-coupling of the invariant informational entity e with different coordinates $\mathcal{S}$.

The advantages of this solution are: it incorporates the strength of Krause's approach in acknowledging the peculiarity of quantum entities (by interpreting "non-individuality" as identity at the $\mathcal{I}$ level), while avoiding its radical metaphysical cost (we are still talking about "entities," just at different levels); simultaneously, it explains why Saunders's strategy of introducing relational properties appears viable in some cases (because he mistakenly took properties at the $\mathcal{S}$ level as the basis for individuation at the $\mathcal{I}$ level), yet is fundamentally misguided.

2.3.3. Formal Derivation Proof of Information Conservation

Assume a fundamental particle state can be expressed as an ordered pair: Particle 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 $\boldsymbol{x}$, time t).

Axiomatic Operational Definitions:

Coordinate Binding: P_k = ($\mathcal{I}$, $\mathcal{S}$k) e.g., Electron near a photon: $\mathcal{S}$k = (Relative Position: Adjacent to photon γ)

Coordinate Decoupling (Annihilation): ($\mathcal{I}$, $\mathcal{S}$k) → ($\mathcal{I}$), ($\mathcal{S}$k) ⇒ The particle degenerates into a pure eigenstate entity. Due to the lack of observational basis ($\mathcal{S}$ = ∅), it becomes unmeasurable.

∀ particle states ($\mathcal{I}\alpha$, $\mathcal{S}$i) and ($\mathcal{I}\alpha$, $\mathcal{S}$j), satisfying:

(Content($\mathcal{I}$_α) ≡ Content($\mathcal{I}$_α)) ⇨ ($\mathcal{I}\alpha$, $\mathcal{S}$i) ≣ ($\mathcal{I}\alpha$, $\mathcal{S}$j)

This indicates:

When the intrinsic properties of two particles are indistinguishable ($\mathcal{I}\alpha$ = $\mathcal{I}\alpha$), regardless of how their spatiotemporal coordinates $\mathcal{S}$i ≠ $\mathcal{S}$j differ, the particles are projections of the same informational entity **e** = Content($\mathcal{I}$) in different spacetimes.

Physical Interpretation:

Particle Annihilation ⇨ Decoupling of the set, not destruction ⇒ **e** = ($\mathcal{I}$) enters a free state.

 doi:10.20944/preprints202506.2384.v3

Particle Creation ⇨ The same **e** binds to new coordinates $\mathcal{S}'$ ⇒ Observed as reappearance.

Example: An electron e^- disappearing at position $\boldsymbol{x}_1$ and appearing at $\boldsymbol{x}_2$ is actually the coordinate migration of the entity **e** = (q=-1e, m_e, s=1/2...): (**e**, $\boldsymbol{x}_1$) → (**e**) → (**e**, $\boldsymbol{x}_2$). Its informational identity is guaranteed by n = n.

Direct Corollary: Information Conservation Theorem: Logic permits that information, once it exists, cannot be annihilated nor newly created.

2.4. Symmetry

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

Traditional response strategies mainly fall into two categories: one questions the metaphysical possibility of such a symmetric universe (e.g., demanding a ground for "numerical difference" itself, often reverting to some hidden property); the other, like Saunders (2003), argues that relational properties (e.g., "being X miles from a sphere") themselves can serve as a basis for weak discernibility. However, the former is criticized as ad-hoc, while the latter struggles within Black's original setup because each sphere's relational property ("being X miles from the other sphere") remains identical.

This paper contends that Black's challenge and the difficulties of traditional responses share a root in an unexamined presupposition: that "numerical twoness" is a primitive, irreducible fact. Steins Theory offers a novel analytical perspective. Within the Steins framework, we must first clarify the comparison level of the Content() function.

If we define Content(Sphere) as the set of all traditional properties (intrinsic + relational), i.e., Content(A) = {Mass M, Shape Spherical, ..., Distance from a sphere X} , then according to Axiom 2, because Content(A) = Content(B) , we must conclude A ≡ B (Uniqueness Theorem). This seems to directly yield the PII conclusion that Black sought to refute.

However, Black's intuition—"there are clearly two spheres here"—is not entirely unfounded. Steins Theory interprets this as a cognitive framing effect. The observer reports "seeing two" because their perspective is itself embedded within this symmetric spatiotemporal coordinate system $\mathcal{S}$. This paper argues that a fundamental misconception shared by Black and his commentators is the default assumption that the referents "A" and "B" necessarily correspond to two entities possessing independent spatiotemporal coordinates. This presupposition forces them into a dilemma between "abandoning PII" or "inventing new metaphysical concepts." Steins Theory offers a third way through this impasse by introducing the concept of "Coordinate Self-Reference." Formalization of the System Content() : For the entire symmetric system S, we should define: Content(S) = { There exists an entity a, with property set P, and a is facing itself } . While this description sounds complex, simply put, Content(S) describes a single coordinate framework that allows "self-facing." Within this framework, "a is facing a" is not a grammatical error but an accurate description of a singular, self-referential coordinate topology. The visually apparent "two" spheres are projections of this single, self-referential coordinate structure within Euclidean spatial perception (similar to an object and its mirror image, but here there is no mirror, only the topological nature of space itself). Resolution of the Paradox:

At the Content(Entity) level: Content(a) = Content(b) = P . According to the Uniqueness Theorem, this points to a single entity, denoted e.

At the Content(System S) level: Content(S) describes the state of a single entity e bound to a special self-referential coordinate topology: (e, R_self-facing).

Paradox Dissolved: Black's error lies in his illicit inference from the system state (e, R_self-facing) to the existence of two entities (e_a and e_b). He confused levels, using the description result of Content(S) to answer a question about Content(e) . In reality, there never was a second entity b; there

 doi:10.20944/preprints202506.2384.v3

was always only one entity e, situated within a special coordinate topology that produces a "double-image projection."

Therefore, this framework does not deny our intuition of "seeing two spheres"; instead, it provides a novel, more precise ontological explanation for that intuition: it is the perception of a single entity within a self-referential coordinate topology. This successfully resolves the apparent contradiction between PII and counting intuition, while avoiding the introduction of any ad-hoc individuating factors. Far from refuting the law of identity, Black's challenge, through the level analysis of this framework, more profoundly reveals the dependence of "identity" judgments on the background framework.

3. Examples (Derived Mathematically, Defaulting to a Scientific Perspective for Discussion) It Is Noteworthy That This Framework Not Only Resolves Quantum Puzzles but Also Offers New Perspectives for Understanding Classical Paradoxes Like Gibbs' Paradox

3.1. Copy Paradox

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

Solution:

If the target is pure informational content identity → Content(n) = Textual Semantics , then $n \equiv m$ (single entity).

If the target is document location entity identity → Content(Doc_n) = (Textual Semantics, Location) , then (n, Loc_A) ≢ (m, Loc_B).

Conclusion: Copies are the same informational entity bound to different spatiotemporal coordinates, making them observable as distinct.

3.2. Gibbs Paradox

Fallacy Essence:

Target should be particle type identity → Content(g) = (Mass, Spin, ...)

Classical statistics illegally expands this to Content(g_i) = (Intrinsic Properties, Fictitious Label)

Correction: $\forall g_i, g_j$: Content(g_i) = S_int = Content(g_j) ⇒ $g_i \equiv g_j$ (type identity). The entropy calculation error stems from incorrectly choosing the Content() domain (introducing labels).

3.3. Black Hole Information Paradox

Traditional Fallacy: Illegally binding the information entity n's Content() to spatiotemporal coordinates: Content(Info_n) = (Information Structure, Black Hole Coordinates)

Correct Solution:

Define target: Information structure identity → Content(n) = Quantum State Encoding

The black hole decouples the set (Quantum State Encoding, Coordinates). The unpaired coordinates render the content unobservable, but Content(Quantum State Encoding) as an abstract entity does not vanish.

If a new spatiotemporal entity satisfies Content(m) = Content(n) , then $m \equiv n$.

 doi:10.20944/preprints202506.2384.v3

3.4. Chinese Room Thought Experiment

Solution: Define the target entity as Chinese understanding functionality: Content(Understanding) = Input-Output Behavioral Consistency .

If the Chinese Room system's behavior is indistinguishable from a native speaker's: Content(System) ≡ Content(Person) , then according to Axiom n=n and the Uniqueness Theorem: The system objectively understands Chinese.

Searle's Fallacy: Illegally expands Content() to the subjective dimension (e.g., operator's mental state), confusing the functional entity level with the conscious entity level.

3.5. Ship of Theseus

Solution: A prominent view in discussions of the Ship of Theseus (cf. Kripke, 1980; Wiggins, 1980) holds that an object's identity is guaranteed by its spatiotemporal continuity and historical-causal path. This paper argues that this view mistakenly and illegally incorporates 'spatiotemporal coordinates' and 'history' as additional properties into the domain of the ship's own Content() function.

Definition: Let the target entity be the ship's abstract informational structure: **S** = Content(Ship) = (Material, Form) . Assume replacement material is identical to the original.

Process:

Component Replacement: $(\mathbf{S}, \mathcal{S}_0) \rightarrow (\mathbf{S}, \mathcal{S}_0) \rightarrow (\mathbf{S}, \mathcal{S}_0)$ [Structure remains]

Reassembly of Old Parts: $(\mathbf{S}, \mathcal{S}_0) \rightarrow \mathbf{S} \rightarrow (\mathbf{S}, \mathcal{S}_n)$ [Structure remains]

∵ Content(**S**) ≡ Content(**S**) (n=n) ∴ The entities $(\mathbf{S}, \mathcal{S}_n)$, $(\mathbf{S}, \mathcal{S}_o)$, $(\mathbf{S}, \mathcal{S}_0)$ are all the original ship.

Fallacy Critique: Claiming that historical causation determines identity constitutes illegal expansion: Content(Ship_A) = (Function, Form, History) ⇒ Violates the initial Content() domain definition, constituting a logical fallacy.

3.6. Twin Earth Paradox

Traditional Contradiction: Putnam's (1975) thought experiment argues that meaning isn't in the head. "Water" on Earth (H_2O) vs. Twin Earth (XYZ) have different chemical compositions. Are the "water" concepts of inhabitants on both planets the same?

Theoretical Solution:

If Content(Water Concept) = Macroscopic Properties (colorless, chemical reactions, drinkable liquid, etc.) → The concepts on both planets are identical (n=n).

Introducing microstructure (H_2O/XYZ) at this point illegally expands Content() to the molecular form level, constituting a logical fallacy.

Conclusion: Semantic identity is determined solely by cognitive function, independent of underlying physics.

3.7. Grandfather Paradox

Contradiction Point: If you travel back in time and kill your grandfather ⇒ You should not exist ⇒ Cannot perform the assassination.

Theoretical Dissolution:

Define target entity: Worldline identity Content(Worldline) = Event Logical Structure .

The assassination event causes:

Original Worldline W_0: (Grandfather survives → You exist → You assassinate)

New Worldline W_1: (Grandfather dies → You do not exist)

∵ Content(W_0) ≠ Content(W_1) ∴ W_0 and W_1 are different informational entities (not "the same worldline being modified").

Furthermore, the paradox itself contains a larger logical fallacy: If you travel back to the past of your worldline, its future ceases to exist. The "you" from the future is erased upon arrival in the past, not merely when the timeline is interfered with later.

 doi:10.20944/preprints202506.2384.v3

3.8. Brain in a Vat

Solution: Current debates about the 'Brain in a Vat,' whether skeptical or realist interpretations, implicitly incorporate properties of the 'external carrier' (biological brain or vat) illegitimately into judgments about the identity of the 'cognitive entity.' By strictly distinguishing Content(Cognitive Stream) and Content(Carrier) , this paper aims to dissolve the debate itself:

Problem: How to prove you are not a brain in a vat? Perception cannot distinguish real from simulated.

Applying the Theory:

Define Content(Cognitive Entity) = Perceptual Information Stream .

Real Brain B_real: Content = {Light signals, Touch...}

Vat Brain B_vat: Content = {Electrical stimulation signals}

If Content(B_real) ≡ Content(B_vat) ⇒ According to Axiom B_real ≣ B_vat (same cognitive entity).

Key Point: The "reality" controversy essentially stems from illegally expanding Content() to the external carrier (skull/culture vat), while the cognitive entity is determined solely by the information stream.

3.9. Mary's Room

Scenario: Mary knows all color neuroscience but has never seen red → When she first sees red, does she gain new knowledge?

Theoretical Answer:

Define knowledge entity levels:

Propositional Knowledge: Content(K_prop) = Wavelength data of red light

Qualia Knowledge: Content(K_qualia) = Subjective red experience

∵ Content(K_prop) ≠ Content(K_qualia) ∴ They are different informational entities. Mary gains the new entity K_qualia, not a supplement to K_prop ⇒ Paradox arises from confusing knowledge types.

3.10. Newcomb's Paradox

Paradox Core: Predictor with near-perfect accuracy vs. participant's free will choice. Choose one box (known to contain money) or two boxes (possibly more money)?

Theoretical Deconstruction:

Illegal Expansion: Confusing the Content() domain of the decision entity:

Level 1 (Pure Decision Logic): Content(Decision) = (Choice Action, Payoff Function) ⇒ Dominant strategy: Choose two boxes (regardless of prediction accuracy).

Level 2 (Causal History Binding): Content(Decision) = (Choice Action, Payoff Function, Prediction History) ⇒ If predictor is accurate, choosing one box yields higher payoff.

Uniqueness Theorem Adjudication:

If target is rational decision without historical constraint → Use Content(Level 1) ⇒ Choose two boxes.

If target is decision incorporating predictive causality → Use Content(Level 2) ⇒ Choose one box.

Paradox Dissolved: The two are decision entities at different levels: Content(Level 1) ≠ Content(Level 2) . The contradiction arises from surreptitiously swapping domains.

3.11. Raven Paradox

Paradox Core: "All ravens are black" ≡ "All non-black things are not ravens." Why does observing a red apple (non-black and non-raven) confirm this proposition?

Theoretical Deconstruction:

 doi:10.20944/preprints202506.2384.v3

Illegal Expansion: Illegally expanding the Content() of "confirmation" from propositional logical structure to empirical sample type.

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:

The red apple confirms the logically equivalent contrapositive proposition ($\neg$Black $\Rightarrow$ $\neg$Raven); its Content(Confirmation) is identical to observing a raven (since Content(P) $\equiv$ Content(P)).

Claiming "red apple and raven have different confirmatory power" illegally expands Content() to the sample's physical category (bird/fruit), violating the initial logical target.

3.12. Sorites Paradox (Paradox of the Heap)

Paradox Core: Removing one grain of sand does not turn a heap into a non-heap $\Rightarrow$ Eventually removing all sand still calling it a "heap" is contradictory.

Theoretical Deconstruction:

Illegal Expansion: Confusing the Content() definition of the "heap" entity:

Level 1 (Topological Structure): Content(Heap) = Macroscopic form of the sand grain collection $\Rightarrow$ Removing one grain doesn't change structural identity (n = n).

Level 2 (Atomic Count): Content(Heap) = Number of grains N $\Rightarrow$ When N=0, Content(Heap) = $\emptyset$, entity ceases to exist.

Uniqueness Theorem Application:

If heap is defined as a structural entity Content(Level 1) , then removing one grain leaves it the same heap.

If heap is defined as a numerical entity Content(Level 2) , then each grain removal creates a new entity.

Paradox Root: Slippage in the argument by surreptitiously changing the Content() domain (from form to number).

3.13. Sleeping Beauty Problem

Paradox Core: Sleeping Beauty, awakened at different stages, what should her probability estimate for a coin being Heads be (1/2 or 1/3)?

Theoretical Deconstruction:

Illegal Expansion: Confusing the Content() domain of the "probability" entity:

Level 1 (Prior Probability): Content(Probability) = Coin's physical state $\Rightarrow$ P(Heads) = 1/2.

Level 2 (Information Update): Content(Probability) = (Coin State, Number of Awakenings) $\Rightarrow$ P(Heads | Awake) = 1/3.

Uniqueness Adjudication:

If asking "Probability of coin's true state" $\rightarrow$ Content(Level 1) $\Rightarrow$ 1/2.

If asking "Probability under current awakening condition" $\rightarrow$ Content(Level 2) $\Rightarrow$ 1/3.

Contradiction Root: Treating two different probability entities Content(Level 1) $\neq$ Content(Level 2) as the same question.

3.14. Modern Contradiction in Pascal's Wager

Problem: If gods of multiple religions all claim "only I am true," how should a rational person wager?

Theoretical Deconstruction:

Illegal Expansion: Confusing the Content() domain of "God":

Level 1: Content(God_X) = Divine description within a specific religious doctrine

Level 2: Content(Omnipotent Entity) = Abstract supreme being transcending specific doctrines

Adjudication:

 doi:10.20944/preprints202506.2384.v3

If comparing the reality of specific religious gods → Each Content(God_X) is different ⇒ Entities are distinct.

If asking "Does a supreme entity exist?" → Requires an independent definition of Content(Omnipotent Entity) , unrelated to specific religions.

3.15. Unexpected Hanging Paradox

Problem: Judge declares "You will be hanged unexpectedly next week." Prisoner reasons it's impossible, yet the hanging day arrives.

Theoretical Deconstruction:

Illegal Expansion: Slipping the Content(Unexpectedness) from "prisoner's cognitive state" to "objective time point."

Correct Definition: Content(Unexpected) = Prisoner's inability to be certain of hanging on the day before execution

Conclusion: The hanging day necessarily exists (due to objective passage of time), while Content(Unexpected) depends solely on the prisoner's cognitive state; the two belong to different entities.

Applications

4.1. The Dilemma of Personal Identity Problems and Existing Theories

The core issue of personal identity is: What makes a person the same over time? Traditional theories mainly debate physical continuity (e.g., brain continuity) and psychological continuity (e.g., memory, personality coherence). Among them, Derek Parfit's (1984) highly influential reductionist psychological continuity theory reduces personal identity to overlapping chains of psychological connectedness (e.g., memory, personality, intentions) over time. This theory demonstrates extraordinary explanatory power when dealing with dynamic changes, such as gradual cell replacement or slow personality shifts. It successfully shows that personal persistence is not an "all or nothing" metaphysical fact but a matter of degree.

However, Parfit's theory, as well as competing physical continuity theories, all implicitly presuppose a more fundamental and unarticulated premise: how to determine, at a given time-slice, that an entity is a "person," and how to statically compare them across time-slices and possible worlds. In other words, these theories excel at answering "Why is he still him?" (dynamic persistence question) but neglect to define "What exactly is 'he' at time t?" (static identity question). 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 fully psychologically continuous successors, physical continuity theories collapse as they cannot handle "one dividing into two"; Parfit's psychological continuity theory faces a dilemma: if identity is considered non-transitive (B and C both identical to A but not to each other), it violates logic; if the original individual is considered to cease after fission, it contradicts the core claim that "psychological continuity suffices for identity." Through this experiment, Williams powerfully demonstrates that without a clear static identity criterion, any discussion of dynamic continuity will descend into conceptual confusion.

This paper argues that the common root of the aforementioned dilemma lies in existing theories all attempting to treat "person" as a primitive concept defined by specific physical substrates or historical causality, mistakenly and illegitimately incorporating the attributes of "carrier" (biological brain) or "history" (causal chain) into judgments about the identity of the "person entity" itself. The debate between Parfit and Williams is essentially a conflict between two different Content() function domains (one being psychological attribute flow, the other physical carrier history), but neither side realizes this, thus reaching an irresolvable impasse.

An Analytical Framework Based on Information Content

 doi:10.20944/preprints202506.2384.v3

According to Steins Theory's two axioms, 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:

Let P be a person-stage (a person's time-slice).

Define it as an ordered pair: P = Content(P, $\mathcal{S}$)

Content(P): Represents the core information structure of the person at this time-slice (e.g., a specific set of perceptions, memories, personality traits, cognitive function patterns).

$\mathcal{S}$: Represents the carrier instantiating this information structure (e.g., a specific brain).

Based on this, for any two person-stages P_1 = Content(P, $\mathcal{S}_1$) and P_2 = Content(P, $\mathcal{S}_2$), we have: Content(P) ≡ Content(P) ⇒ $P_1 \equiv P_2$ This means that as long as the information content at two time-slices is the same, they are different instances of the same person entity, regardless of whether their spatiotemporal coordinates $\mathcal{S}$ are continuous.

Thus, we provide a clear resolution for Williams' fission experiment: the two successors P_B and P_C cause a paradox because we mistakenly require dynamic continuity to map to a one-to-one physical path. Under this framework, we only need to compare static content: if Content(P_B) = Content(P_A) and Content(P_C) = Content(P_A), then according to n = n, both P_B and P_C are the same person entity as P_A. This is not a logical contradiction but the simultaneous instantiation of the same information entity at multiple coordinates. Parfit's intuition—"survival without identity"—is most strictly realized here: survival is the identity of information content itself, requiring no additional "identity" based on carrier history to guarantee.

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

4.2. "Spatiotemporal Jump" as the Logical Necessity of Coordinate Decoupling and Rebinding

Before discussing the "spatiotemporal jump" of conscious entities, we must first pay the highest tribute to modern physics, especially Einstein's special and general relativity. With unparalleled precision and beauty, these theories successfully describe the profound dynamical relationships between mass, energy, and spacetime, strictly stipulating the upper limit of causal laws that any physical signal or entity motion must follow—the speed of light. Any attempt to realize a "spatiotemporal jump" at the physical level, whether through wormholes, warp drives, quantum suicide, or other exotic mechanisms, must be tested within the solid frameworks of relativity or quantum mechanics, facing enormous physical challenges like energy conditions, singularities, and empirical verification.

However, the "spatiotemporal jump" argued in this paper is fundamentally different from all the above physical processes. It is not a motion process existing within spacetime, governed by physical laws, but a logical necessity derived from the Steins axiom system. It answers a more primitive question: "Are two information entities instantiated at different spatiotemporal coordinates with the same content the same entity?" The answer to this question does not depend on the physical path connecting them but solely on the logical axiom n = n.

The objects of study in traditional physics (including relativity) are precisely defined in this framework as "binding states of conscious information and spatiotemporal coordinates," i.e., (C, $\mathcal{S}$). Physics perfectly describes how binding states evolve over time, i.e., (C_t, $\mathcal{S}$t) -> (C{t+at}, $\mathcal{S}$_{t+dt}), finding that these evolutions follow elegant differential equations (e.g., Einstein's field equations).

This theory, however, focuses on a possibility that physics, due to the limitations of its research paradigm, naturally does not discuss: the decoupling and rebinding of conscious information C with its coordinates $\mathcal{S}$.

 doi:10.20944/preprints202506.2384.v3

Decoupling: (C, $\mathcal{S}_1$) → (C), ($\mathcal{S}_1$). Physically, this may correspond to the destruction of the carrier (e.g., the brain) by some event conforming to the event horizon principle (a velocity difference causing causal isolation at the neuronal level), leading to the information structure no longer being instantiated ($\mathcal{S} = \emptyset$).

Rebinding: (C) → (C, $\mathcal{S}_2$). Physically, this may correspond to an instantaneous "reappearance" occurring elsewhere (e.g., from Poincaré recurrence, MWI parallel universes, bubble universes, etc.).

The key point is that according to Axiom 1 (n = n), the abstract conscious information C after decoupling maintains its self-identity. Therefore, the new state (C, $\mathcal{S}_2$) after rebinding and the old state (C, $\mathcal{S}_1$) before decoupling, because they share the same C, must be different instances of the same information entity. This is the logical core of the "spatiotemporal jump": 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 laws of (C, $\mathcal{S}$) binding states within spacetime. It asks "how to get from A to B."

This theory studies the discrete identity logic of the C entity itself beyond spacetime. It asks "are A and B the same thing."

Relativity prohibits any physical entity from moving faster than light, but it cannot and need not prohibit, nor does it care about, a logical entity's two "realizations" at different spacetime points. The "spatiotemporal jump" seems "inconceivable" or even "violating physics" precisely because we mistakenly use physical laws describing the motion of binding states to judge a logical theorem about the identity of information entities. This is a category mistake.

Conclusion: The "spatiotemporal jump" proposed by this framework is not a physics conjecture to be realized but an already established logical inference. Starting from the most basic law of identity, it deduces a brand-new picture of personal identity: the persistence of consciousness fundamentally lies in the identity of its information patterns, not in the continuity of physical processes connecting these pattern instances. This provides an unprecedented clear framework for understanding thought experiments like teleportation and the brain in a vat, completely liberating the discussion of personal identity from the constraints of physics and placing it on a more fundamental logical and metaphysical foundation.

Possible method: Self-Envating (entering a room isolated from the environment needing change, refer to Chapter 3.8) → Detector checks the isolated environment (e.g., AI search, etc.) → Judgment result (if requirements are met, end; if not, proceed to next step) → High-speed destruction leading to decoupling (e.g., explosives, survivor effect. Refer to Chapter 4.2) → Universe randomness leading to reappearance (refer to Chapter 4.2). Repeatable operations enable multi-stage jumps.

5. Conclusion

Unity of Absoluteness and Relativity: Identity is absolutely guaranteed by n = n, but the judgment n ≡ m is relative to the chosen Content() function. Any claim introducing external properties constitutes illegal expansion.

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

Fundamentality: This theory is the most basic metaphysical framework. Even in a completely random universe where causality vanishes, World ≡ World still holds.

Explanatory Power: It successfully incorporates the theories of Leibniz and Kripke as special cases and resolves the modern scientific challenges they could not address, particularly the problem of identical particles in quantum mechanics.

References

1. Leibniz, G. W. (1714). Principle of the Identity of Indiscernibles (in Monadology).

 doi:10.20944/preprints202506.2384.v3

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.
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. (Referencing his work on non-individuals and quasi-set theory).
30. Parfit, D. (1984). Reasons and Persons. Oxford University Press.
31. Williams, B. (1970). The Self and the Future. The Philosophical Review.