Problem Identity as a Presupposition of Reproducibility Evaluation

A Formal Account of the DTC Conditions via Structural Isomorphism

K. Nabeya

Preprint. This manuscript has not yet undergone peer review.

March 22, 2026

Abstract

This paper argues that reproducibility evaluation presupposes problem identity and provides a formal framework for diagnosing failures of this presupposition. A problem space is defined as $M = (S, \tau, \Gamma)$, consisting of states, transitions, and admissible constraints. Problem identity is defined as structural isomorphism across these components, yielding three preservation conditions: D (distinguishability), T (transition), and C (constraint). Under Condition E—a methodological constraint on preserved representations, distinct from any ontological constraint on problems—the joint satisfaction of the DTC conditions suffices for reconstructability. This distinction has not been formally established in existing accounts. The DTC decomposition provides a systematic criterion for diagnosing when reported failures of reproducibility in fact arise from failures of problem identity—a misclassification that existing frameworks lack the structural resources to detect, as they lack an explicit structural representation of problem identity as a decomposable condition, thereby enabling their reclassification as structurally distinct kinds of failure.

Keywords: problem identity, reproducibility, structural isomorphism, philosophy of science, DTC conditions

1. The Hidden Presupposition of Reproducibility Discourse: What Does It Mean to Address "the Same Problem"?

This section shows that reproducibility discourse presupposes problem identity yet never makes this presupposition explicit; identifying it is the starting point of the entire argument.

1.1 The Ambiguity of Reproducibility

Leonelli (2018) clarifies the ambiguity of reproducibility by distinguishing three levels: direct, analytical, and conceptual. Direct reproducibility refers to the agreement of results under identical conditions; analytical reproducibility to agreement achieved by re-analysing existing data; and conceptual reproducibility to the confirmation of the same finding via different approaches. Each level carries its own standards and its own conditions for evaluation.

Despite this useful taxonomy, Leonelli's analysis does not make explicit that each level of assessment is grounded in a shared presupposition: that the studies under comparison address the same problem. Judging whether findings are "identical" in the sense required by conceptual reproducibility, for instance, already presupposes that both studies answer the same question. Yet the structural conditions implied by this presupposition remain unanalysed in Leonelli's account.

1.2 The Implicit Role of Problem Identity in the Replication Crisis

Ioannidis (2005) identifies a systematic failure in published research: results that cannot be reproduced across independent studies. His analysis focuses on statistical factors such as insufficient statistical power, selective reporting, and p-hacking. However, it leaves implicit a prior condition: that the studies being compared are testing the same hypothesis in the same problem space.

The question "are the same hypotheses being tested?" is logically independent of statistical failure, yet the two are treated without distinction in Ioannidis's framework. What Ioannidis targets is *failure of reproducibility*—obtaining different results for the same problem. But the question "are we dealing with the same problem?"—problem identity—functions as a *precondition* for reproducibility assessment, and as such is not itself an object of evaluation in his account.

1.3 Main Thesis of This Section

> **Thesis (this section):** *Reproducibility evaluation presupposes problem identity.*

Failure of reproducibility (obtaining different results for the same problem) and failure of problem identity (not addressing the same problem in the first place) are conceptually distinct. The former can be assessed only if the latter is presupposed to hold. This distinction has gone unnoticed because existing frameworks lack the tools to make problem identity explicit as a structural condition.

Failures of reproducibility and failures of problem identity have not been formally distinguished in existing accounts. This paper formally establishes this distinction and makes it possible to diagnose when a reported failure of reproducibility is in fact a failure of problem identity. This diagnostic capacity constitutes the substantive addition that the Structural Reconstructability Framework (SRF) makes to existing discourse: it enables the reclassification of failures that have been treated as a single phenomenon into two structurally distinct kinds.

The SRF acquires its theoretical profile by ruling out four positions. First, same procedure is not a sufficient condition for problem identity (same $\tau \neq$ same problem). Second, same outcome is not sufficient either (same outcome $\neq$ same problem). Third, approximate problem identity cannot be defined without a prior ideal criterion of exact isomorphism. Fourth, problem identity is not an intrinsic property of a problem but a structural-preservation relation between problem spaces (*relational identity*).

This distinction carries direct methodological implications. A portion of the failures that the replication crisis has treated as failures of reproducibility must be re-evaluated as failures of problem identity, requiring different diagnostic and corrective responses.

2. Why Does Natural-Language Description of Problem Identity Fail?

This section shows that "same data," "same method," and "same result" do not guarantee problem identity, establishing the need for a formal framework before any formalisation is attempted.

2.1 The Insufficiency of "Same Data"

Even when experimental data are nominally identical, the studies may not address the same problem if the interpretation of what states those data represent differs. The state set S in a problem space is always interpreted from a particular theoretical and methodological perspective, and that perspective can shift silently across studies.

This is not merely a theoretical concern. In machine learning, even when the same dataset is used, different pre-processing routines, treatments of missing values, or normalisation procedures may in effect target different state spaces. The description "same data" does not capture this difference, and therefore does not guarantee identity of S across studies.

2.2 The Insufficiency of "Same Method"

Procedural agreement does not guarantee problem identity unless the procedure is identified with a specific transition structure τ within the problem space. The description of a method in natural language underdetermines which transitions it realises, and different realisations may correspond to structurally distinct problems.

Compression of data into summary statistics illustrates this failure. When experimental data are stored only as means and variances, the transition sequence of the data-generating process and the associated measurement constraints are irretrievably lost. "Applying the same statistical method" does not imply identity of transition structure; it leaves τ underdetermined.

2.3 The Insufficiency of "Same Result"

Agreement of results is not evidence of problem identity in itself. It functions as evidence of reproducibility only *if* problem identity is presupposed to hold—which is precisely what is at issue. Using result agreement to establish problem identity is therefore circular.

This circularity has practical consequences. In a black-boxed simulation, even if final outputs agree, if the history of parameter changes, execution-environment conditions, and grounds for initial-condition selection have been lost, the boundary conditions Γ of the problem space remain indeterminate. "Same result" does not imply preservation of problem boundary conditions.

2.4 The Distinction Between Procedural and Problem Reproduction

The three examples above converge on a single diagnostic point. *Procedural reproduction* (repeating the same steps) and *problem reproduction* (reconstructing the same problem space) are different achievements, and natural-language description systematically conflates them.

Natural-language description can capture procedural reproduction but has no systematic means of identifying the conditions for problem reproduction. It cannot classify which of D, T, or C has collapsed, and therefore cannot clearly distinguish failure of problem identity from failure of reproducibility. A formal framework is required to make this distinction tractable.

2.5 Conclusion of This Section

> **Conclusion:** *A formal framework is required to handle problem identity.*

The Structural Reconstructability Framework (SRF) developed in the following sections is proposed as such a framework. What SRF makes possible is the systematic identification of which component—S, τ, or Γ—suffers which collapse of D, T, or C, and under what conditions.

3. Formal Characterisation of Problem Spaces

This section introduces the three-component structure $M = (S, \tau, \Gamma)$, argues for the necessity and independence of each component, and positions SRF relative to prior formal frameworks.

3.1 Perspectival Fixing

Formalising problem identity as a structural notion requires delimiting the scope of analysis. *Perspectival fixing* refers to the situation in which it is settled in advance which states and transitions count as constituents of the problem space. Under this condition, the state set S is uniquely determined as the totality of states distinguishable from the perspective in question.

Cases in which the perspective shifts involve problems of re-description and translation (Kuhn 1962; Quine 1960), which lie outside the scope of this paper (Note 1). SRF does not require that elements of S be theory-neutral observations or that they be purely theory-laden constructs; S is defined relative to the investigator's perspective, and once that perspective is fixed, S can be treated as an external condition of the problem space.

The transition relation τ is a formal expression of the structural connection between states. SRF is neutral as to whether τ represents a causal, nomological, or model-theoretic relation—a stance we call *ontological neutrality*. This neutrality ensures that SRF applies across the full range of scientific disciplines without presupposing a particular metaphysics of science.

3.2 Definition of a Problem Space

The motivation for the three-component structure can be seen from a simple observation. Even when the same data (S) are subjected to the same operation (τ), different selection criteria or exclusion conditions yield different problems. The component that formalises this boundary condition is Γ, the admissibility family.

Definition 1 (Problem Space)**.**

$$M = (S,\ \tau,\ \Gamma)$$

S: state set

- $\tau \subseteq S \times S$: transition relation
- $\Gamma \subseteq \mathcal{P}(S) \times \mathcal{P}(\tau)$: admissibility family

Here $\mathcal{P}(X)$ denotes the power set of X.

Γ is no more arbitrary than S or τ; all three components are theory-laden relative to the fixed perspective, and the theory-ladenness of Γ is therefore no special objection to its inclusion as an independent component. The diagnostic value of SRF lies not in uniquely determining Γ but in identifying, as DTC collapse, the locus at which Γ has not been made explicit. The non-explicitness of Γ is not a defect in the framework; it is the very phenomenon that SRF is designed to illuminate.

This non-explicitness is not accidental. It derives from a structural feature of scientific practice: boundary conditions of problems are typically shared tacitly among practitioners. Γ is not written down precisely because it is shared; because it is not written down, its collapse goes undetected. What SRF illuminates is this **structural invisibility**.

While this paper does not provide a procedure for identifying Γ, such identification typically proceeds through the articulation of methodological conventions and theoretical commitments that structure scientific practice. However, the DTC diagnoses performed in each case study (Section 6) do partially render visible the constraints that Γ implicitly imposes, illustrating aspects of the constraints that such identification would have to capture. The identification of Γ lies outside this paper's scope, but lying outside the scope does not mean remaining unsolved: it is a task to be formalised in the second paper of the trilogy, as the change of Γ under perspectival shift within Problem Transformation Theory.

3.3 The Role and Independence of the Three Components

The three components (S, τ, Γ) are not an arbitrary decomposition. They correspond to three independently operative scientific activities: S to the specification of the domain of objects, τ to the description of structural relations or mechanisms between states, and Γ to the constraint conditions that regulate the admissible scope of explanation. The following three-stage argument shows that (S, τ, Γ) is the minimal sufficient structure for the purposes of this paper.

1. **Stage 1: S alone is insufficient.** Identifying the set of relevant states leaves undetermined what relations hold among them, and therefore what counts as a *solution* to the problem.
2. **Stage 2: (S, τ) alone is insufficient.** Even with states and transitions specified,

the boundary condition “which substructures are admissible for this problem” remains undetermined, so distinct problems cannot be individuated.

3. **Stage 3:** (S, τ, Γ) **suffices.** Adding Γ settles the boundary condition and makes it possible to define problem identity as structural isomorphism.

D, T, and C are therefore the *minimal* independent decomposition: each can fail independently of the others (as shown in § 4.3–4.4), and no strictly coarser partition of the isomorphism conditions in Definition 2 preserves this mutual independence.

3.4 Philosophical Status of Γ

Γ is not merely a boundary condition on the problem space. It is the *individuation condition* (in the sense of Fine 1994) that constitutes S and τ as a particular scientific problem. Without Γ, identical S and τ cannot distinguish different problems.

That C-collapse is observed in all four case studies of Section 6 reflects the central role that Γ plays in problem individuation. Note, however, that Γ alone does not constitute a scientific problem: Γ specifies the *type* of a problem, but the concrete *token* is given by the pair (S, τ). The individuation is always a joint product of all three components.

3.5 Formal Positioning Relative to Prior Work

Relation to Nickles (1981). Nickles’s theory of problem reduction treats a problem as a “question plus constraints” and explicitly handles boundary conditions (corresponding to Γ in this paper). However, it does not provide the three-component decomposition or define re-identification via structural isomorphism, leaving problem identity as a conceptual rather than formal notion.

Comparison with Suppes (2002). Suppes’s representation theorem asks about *representation*: how theoretical structure is expressed in an empirical model. Homomorphism suffices for this purpose, since one needs only one-way structure preservation. SRF, by contrast, asks about *reconstruction*: whether the original problem space can be uniquely recovered from a preserved representation. Reconstruction requires that distinct problem spaces not be mapped to the same representation—a requirement not guaranteed by mere homomorphism. The demand for injectivity follows necessarily from the purpose of reconstruction and is not an additional constraint on Suppes’s framework but an independently motivated condition arising from a different question.

Comparison with Klir (1985). Klir's Reconstructability Analysis treats reconstruction as a probabilistic question: what can be approximately recovered? The reconstructability R in SRF is an isomorphism-theoretic concept that admits no approximation. For positions that treat problem identity as a graded notion, Klir's framework may be more appropriate; this is explicitly acknowledged as a scope limitation of SRF and as a direction for future extension (SRF-g; see Future Directions in Section 8).

Positioning relative to erotetic logic. The semantics of questions developed by Hintikka (1976) analyses the *content* of questions—what counts as an admissible answer to a given question. SRF analyses *identity conditions* for problems: under what structural conditions are two problems the same? The two frameworks address different questions and are complementary, not competing.

4. Problem Identity as Structural Isomorphism

This section defines problem identity as structural isomorphism $M_1 \cong M_2$, names the three DTC conditions, and proves their mutual logical independence.

4.1 Definition of Re-identification

Definition 2 (Structural Isomorphism; Re-identification (following Hodges 1993))**.** Given two problem spaces $M_1 = (S_1, \tau_1, \Gamma_1)$ and $M_2 = (S_2, \tau_2, \Gamma_2)$, **re-identification** holds if and only if there exists a map $\varphi : S_1 \to S_2$ satisfying:

1. φ is a bijection.
2. $(s, t) \in \tau_1 \iff (\varphi(s), \varphi(t)) \in \tau_2$
3. $(S', \tau') \in \Gamma_1 \iff (\varphi[S'], \varphi[\tau']) \in \Gamma_2$

We write $M_1 \cong M_2$ in this case. Here $\varphi[A]$ denotes the pointwise image of a set A under φ. We use bracket notation $\varphi[A]$ for set images and parentheses $\varphi(s)$ for pointwise mappings, to distinguish the two types throughout.

This definition follows the standard model-theoretic formulation of isomorphism (Hodges 1993), and no mathematical novelty is claimed. The philosophical contribution consists in applying this standard framework as a formal criterion for problem identity in philosophy of science.

The choice of structural isomorphism ($\cong$) as the criterion of problem identity has a methodological rationale. The purpose of SRF is to make explicit the ideal criterion—the minimal structural conditions under which problem identity holds completely. Only once exact isomorphism has been specified as an ideal criterion does it become possible to diagnose partial failures of the form "which condition has broken down and to what degree?" The DTC collapse analysis of Section 6 (using the three-valued notation ✓ / △ / ×) is a practical application of this diagnostic purpose, and specifying the ideal criterion is required in order to describe deviations from it.

4.2 The Three DTC Conditions: Definitional Remarks

Remark 1 (Distinguishability Preservation, D)**.** The bijectivity of φ (condition (1)) means that the distinguishability of states is preserved from S_1 to S_2. We call this **D-preservation (D)**.

Remark 2 (Transition Preservation, T)**.** Condition (2) means that the transition relation is preserved in both directions. We call this **T-preservation (T)**.

Remark 3 (Constraint Preservation, C)**.** Condition (3) means that the admissibility family is preserved in both directions. We call this **C-preservation (C)**.

4.3 Independence of Γ

Proposition 1 (Independence of the Admissibility Family)**.** Let $M_1 = (S, \tau, \Gamma_1)$ and $M_2 = (S, \tau, \Gamma_2)$ share the same S and τ. If $\Gamma_1 \neq \Gamma_2$ then $M_1 \ncong M_2$.

Proof. Suppose $M_1 \cong M_2$. Since S and τ are common to M_1 and M_2, the identity map $\mathrm{id} : S \to S$ satisfies conditions (1) and (2). But $\Gamma_1 \neq \Gamma_2$ implies that there exists (S', τ') with $(S', \tau') \in \Gamma_1$ and $(S', \tau') \notin \Gamma_2$. (Here we use the identity as φ, so $\varphi[S'] = S'$, $\varphi[\tau'] = \tau'$.) Condition (3) then fails. Hence $M_1 \ncong M_2$. □

Proposition 1 guarantees that C cannot be logically derived from D and T. This independence provides the formal ground for the intuition established in Section 2 that "same method and same result are insufficient" to establish problem identity.

4.4 Mutual Independence of D and T

Independence of D from T. When $\tau = \emptyset$, T holds vacuously, but D fails if φ is not injective. As a counterexample, let $S = \{s_1, s_2\}$, $\tau = \emptyset$, and let $U = \{u_1\}$ be any singleton set with $\varphi(s_1) = \varphi(s_2) = u_1$. T-preservation holds vacuously but D fails, showing that $T \wedge C \Rightarrow D$ does not hold in general.

Independence of T from D. To show this independence, we consider mappings into an auxiliary representation space E of the kind formalised in Section 5. Let $S = \{s_1, s_2\}$, $\tau = \{(s_1, s_2)\}$, $E_S = \{e_1, e_2\}$, $E_\tau = \emptyset$. Setting $\iota_S(s_1) = e_1$ and $\iota_S(s_2) = e_2$ establishes D. However, the coherence condition for ι_τ requires $\iota_\tau(s_1, s_2) = (e_1, e_2) \in E_\tau$, which fails since $E_\tau = \emptyset$. Hence D holds but T fails, showing that $D \wedge C \Rightarrow T$ does not hold in general.

4.5 Conclusion of This Section

> **Conclusion:** Problem identity can be understood as structural preservation. The holding of $M_1 \cong M_2$ is equivalent to the simultaneous holding of D, T, and C; the collapse of any single condition entails failure of problem identity.

5. Preserved Representations and Reconstructability of Problem Spaces

This section introduces Condition E for preserved representations, defines the reconstruction operation $\Pi(E)$*, and proves that DTC together with Condition E suffices for reconstructability (DTC Sufficiency) while each condition is individually necessary (DTC Minimality).*

5.1 Motivation for Condition E

The DTC conditions are conditions on the structural preservation of the problem space M itself. They say nothing about whether the preserved representation E has the formal properties needed for reconstruction to succeed. Even when DTC holds, the preserved representation E need not uniquely determine the reconstruction operation; Condition E is required as an independent condition to close this gap.

> **Philosophical status of Condition E.** Condition E is a *methodological* condition, not an ontological one. The ontological individuation conditions for the problem space $M = (S, \tau, \Gamma)$ are given by Definition 2 (structural isomorphism). Condition E is a technical condition that ensures the *soundness* of the reconstruction operation Π—that $\Pi(E)$ faithfully recovers the structure of M—and is entirely independent of the ontology of problems.

5.2 Definition of Preserved Representations (with Condition E)

Definition 3 (Preserved Representation; Condition E). $E = (E_S, E_\tau, E_\Gamma)$ satisfies:

- E_S: set; $E_\tau \subseteq E_S \times E_S$; $E_\Gamma \subseteq \mathcal{P}(E_S) \times \mathcal{P}(E_\tau)$

[Condition E]

- **(E1) Finite generation:** $|E_S| < \infty$
- **(E2) Structural closure:** For all $(e, e') \in E_\tau$: $e, e' \in E_S$; for all $(A, B) \in E_\Gamma$: $A \subseteq E_S$ and $B \subseteq E_\tau$.

5.3 Formal Definitions of the Three DTC Conditions

Definition 4 (D-preservation). $\iota_S : S \hookrightarrow E_S$ (injective)

Definition 5 (T-preservation). $\iota_\tau : \tau \hookrightarrow E_\tau$ (injective) and $\iota_\tau(s, t) = (\iota_S(s), \iota_S(t))$ (coherence condition)

Definition 6 (C-preservation). $\iota_\Gamma : \Gamma \hookrightarrow E_\Gamma$ (injective) and $\iota_\Gamma(S', \tau') = (\iota_S[S'], \iota_\tau[\tau'])$ (coherence condition)

5.4 Reconstruction Operation and Reconstructability

Definition 7 (Reconstruction Operation).

$$\Pi(E) := (S^*, \tau^*, \Gamma^*),$$

$$\text{where} \quad S^* := \mathrm{Im}(\iota_S),\ \tau^* := \mathrm{Im}(\iota_\tau),\ \Gamma^* := \mathrm{Im}(\iota_\Gamma).$$

Definition 8 (Reconstructability).

$$R(S, \tau, \Gamma) := \exists E = (E_S, E_\tau, E_\Gamma) \text{ satisfying Condition E such that } \Pi(E) \cong (S, \tau, \Gamma)$$

5.5 Pathological Case When Condition E Is Violated

To establish the non-triviality of Proposition 2, we construct a case in which all three DTC conditions hold yet reconstructability fails because Condition E is violated.

Let $M = (S, \tau, \Gamma)$ with $S = \{s_1, s_2\}$, $\tau = \{(s_1, s_2)\}$, and $\Gamma = \{(\{s_1, s_2\}, \{(s_1, s_2)\})\}$. Set $E_S = \{e_1, e_2, e_3\}$, $E_\tau = \{(e_1, e_2), (e_1, e_3)\}$, and $E_\Gamma = \{(\{e_1, e_2\}, \{(e_1, e_2)\})\}$. All three DTC conditions hold under the mappings $\iota_S(s_i) = e_i$ (for $i = 1, 2$), $\iota_\tau(s_1, s_2) = (e_1, e_2)$, and the corresponding ι_Γ.

However, E violates (E2): the spurious transition $(e_1, e_3) \in E_\tau$ is not the image of any transition in τ. Because the same E could arise from a different problem space M' having an extra transition, E does not uniquely encode M. Condition E is therefore violated, and this E falls outside the scope of the existential quantifier in Definition 8. The construction shows that DTC alone does not suffice for reconstructability without Condition E, establishing the non-triviality of the next proposition.

5.6 DTC Sufficiency

Proposition 2 (DTC Sufficiency)**.** If D, T, and C hold and E satisfies Condition E, then $R(S, \tau, \Gamma)$ holds:

$$(D \wedge T \wedge C \wedge \text{CondE}) \Rightarrow R(S, \tau, \Gamma)$$

Proof. Define $\varphi : S \to S^*$ by $\varphi(s) := \iota_S(s)$. Injectivity of ι_S gives injectivity of φ; since $\text{Im}(\iota_S) = S^*$, it is surjective, hence bijective. For transition preservation, the coherence condition of T gives $\iota_\tau(s, t) = (\iota_S(s), \iota_S(t)) \in \tau^*$; Condition E (E2) prevents spurious transitions from entering τ^*, since every element of E_τ is closed over E_S. For constraint preservation, the coherence condition of C gives $\varphi[S'] = \iota_S[S']$ and $\varphi[\tau'] = \iota_\tau[\tau']$. Hence φ witnesses $(S, \tau, \Gamma) \cong (S^*, \tau^*, \Gamma^*)$, giving $R(S, \tau, \Gamma)$. □

5.7 DTC Minimality

Proposition 3 (DTC Minimality)**.** $\neg D \Rightarrow \neg R, \quad \neg T \Rightarrow \neg R, \quad \neg C \Rightarrow \neg R$

Proof. By counterexample. $\neg D$: if φ is not injective, $|S^*| < |S|$, so no bijection exists. $\neg T$: Let $\tau \neq \emptyset$. If $E_\tau = \emptyset$, then $\tau^* = \emptyset \neq \text{Im}(\iota_\tau)$, so $\Pi(E) \ncong M$. $\neg C$: by Proposition 1, $\Gamma_1 \neq \Gamma_2$ implies $M_1 \ncong M_2$. □

5.8 Bridging Proposition

Bridging Proposition (from problem identity to reconstructability)

Given a problem space $M = (S, \tau, \Gamma)$, if re-identification ($M_1 \cong M_2$) is defined, then the DTC conditions readable from that definition hold; and if a preserved representation E satisfies Condition E, reconstructability R is guaranteed. The structural conditions

> for problem identity are, under Condition E, sufficient for reconstructability.

This paper establishes only the sufficient condition. The general proof of the necessary condition ($R \Rightarrow$ DTC) lies outside the current scope; analysis of the equivalence when Condition E is relaxed is a natural extension of SRF.

5.9 Ontological Implications of SRF

This subsection draws out the ontological interpretation of SRF as Structural Problem Ontology (SPO) and delimits its scope relative to the paper's primary purpose.

SRF is a formal criterion that individuates problems as structures of the form (S, τ, Γ). This criterion carries ontological implications: the conditions for problem identity are not merely methodological rules but constitute individuation conditions for problems as objects. Problem identity is not determined by the individual judgment of an investigator but by the preservation of the structure $M = (S, \tau, \Gamma)$.

Within the scope where the perspective is fixed, two problems are treated as identical as long as this structure is preserved. This position can be located as an application of Worrall's (1989) structural realism to the domain of problems, and may be called **Structural Problem Ontology (SPO)**. SPO holds that problems are individuated by their structure rather than by any intrinsic, non-relational essence.

The purpose of this paper, however, is not to develop the ontological details of SPO but to present formal conditions for problem identity. The mode of existence of problems (Platonism, constructivism, or structuralism), the ontology of problem granularity, and the persistence and disappearance of problem identity through scientific change are all deferred to future work on SPO as a primary subject. In this paper SPO is positioned as an ontological interpretation suggested by the formal consequences of SRF, not as a doctrine that the paper undertakes to defend.

6. SRF Analysis of Reproducibility Failures

This section applies SRF to four cases drawn from contemporary science, showing in each that natural-language analysis cannot identify the structural source of failure whereas DTC analysis can; the cases are chosen for structural diversity, not representativeness, and make no frequency claim.

To indicate the degree to which DTC conditions hold, the notation $\triangle$ (partial collapse) is used alongside $\times$ (full collapse) and ✓ (preserved), distinguished from the strict isomorphism judgement. These are diagnostic indicators of structural collapse, not rigorous isomorphism determinations.

6.1 Case 1: The Psychology Replication Crisis (OSC 2015)

The Open Science Collaboration (2015) conducted replication attempts for 100 psychology papers and reported that only approximately 36% reproduced a statistically significant result. This case is typically discussed in terms of statistical underpowering and publication bias; SRF analysis reveals an additional structural layer.

D (distinguishability): The original study and the replication differ in their subject populations. D is formally preserved in many cases, but the interpretation of S changes tacitly due to individual differences, cultural background, and temporal context.
T (transition): Even when experimental procedures are described, experimenter effects, context effects, and subtle procedural variations affect τ. T-collapse is the most frequently occurring collapse among the three components, since procedural descriptions underdetermine τ.
C (constraint): The subject selection criteria, exclusion criteria, and admissible range of measurement (Γ) that the original study implicitly set are not made explicit; the replication cannot recover them. C-collapse is therefore structurally inevitable.

> **SRF diagnosis:** D preserved, T collapsed, C collapsed (D✓ T× C×)

Natural-language analysis can only say "the procedure was the same but the result differed." SRF analysis identifies the failure of problem identity as T- and C-collapse, showing that part of what Ioannidis called "statistical failure" is in fact a failure of problem identity. Context dependency maps onto C-collapse, and experimenter effects map onto T-collapse; SRF unifies these descriptive observations as structural failures of the same kind.

6.2 Case 2: Training Reproducibility of Large Language Models (GPT-family)

Training runs of large language models are in principle held to be reproducible, but it has been reported that training results are not reproduced even with the same architecture, dataset, and hyperparameters (Bouthillier et al. 2021). This case is notable because the apparent conditions for reproducibility are satisfied, making the failure structurally puzzling.

D: The final parameter set (weight matrices) is preserved as the correlate of S. D formally preserved.

T: The gradient-update history, the order of mini-batch selection depending on random seeds, and the non-deterministic execution order of data-parallel processing constitute the transition sequence corresponding to τ; these are not normally recorded or stored. T-collapse.
C: The grounds for hyperparameter scheduling, the admissible range of architectural design choices, and the convergence criterion (Γ) are only partially made explicit. C-collapse.

> **SRF diagnosis:** D preserved, T collapsed, C collapsed (D✓ T× C×)

The "different results" reported in this literature should, at least in part, be described as failure of problem identity rather than failure of reproducibility. The two diagnoses call for different responses: improving reproducibility calls for fixing the random seed and recording execution order (T), while addressing problem identity requires making Γ explicit (C).

6.3 Case 3: Statistical Re-analysis and Compressed Storage (Begley & Ellis 2012)

Begley & Ellis (2012) reported that only 6 of 53 "landmark" cancer research papers could be reproduced. A characteristic feature of many studies was that data had been stored in compressed form as summary statistics (means, variances, p-values) rather than as raw experimental records.

D: Compression into summary statistics destroys the distinguishability of individual observations. D partially collapsed (D△).
T: The data-generating process (the sequence of experimental operations) corresponds to τ; with raw data lost, recovery of τ is impossible. T-collapse.
C: The grounds for experimental condition selection, admissible range, and exclusion criteria (Γ) are not reported. C-collapse.

> **SRF diagnosis:** D partial collapse, T collapsed, C collapsed (D△ T× C×)

SRF analysis makes explicit that the failure of re-analysis originates in D-collapse at the data-storage stage, and provides a structural answer to the question "what should be stored?" The answer is: sufficient raw data to preserve distinguishability (S), the full data-generating sequence (τ), and the boundary conditions of the study (Γ).

6.4 Case 4: Reproducibility of Climate Simulation Models (CMIP)

In climate-change prediction models (CMIP), significant differences arise in long-run simulation results even when multiple research groups use models implementing "the same physical equations." This case is philosophically important because the apparent agreement on physical theory coexists with systematic divergence in simulation outcomes.

D: States set as initial conditions (temperature, pressure, etc.) are preserved as S. D preserved.

T: The choice of numerical integration scheme, spatial discretisation method, and parameterisation-scheme implementation induces variation in τ. "Same physical equations" does not imply identity of the transition structure that those equations induce in a specific numerical implementation. T partially collapsed (T△).

C: The admissible range of substructures treated as "the real climate system" (Γ)—e.g., the admissible range of cloud parameterisation, coupling conditions with ocean circulation models— differs across research groups. C-collapse.

> **SRF diagnosis:** D preserved, T partial collapse, C collapsed (D✓ T△ C×)

The claim "we are modelling the same physical system" does not guarantee problem identity unless Γ agrees. Disagreement about which physical processes fall within the admissible scope of a climate model is a disagreement about Γ, not about the physical equations per se.

6.5 Comparative Analysis and Methodological Implications

C-collapse is observed in all four case studies. This observation does not assert an inductive generalisation from four cases; rather, it is understood as a consequence of the definitional non-explicitness of Γ identified in Section 3. Γ is the component of scientific practice most resistant to explicit articulation: what counts as admissible for a problem typically depends tacitly on methodological conventions and theoretical commitments in ways that S (recorded data) and τ (described procedure) do not.

The highest priority for improving reproducibility is therefore making Γ explicit. Data preservation (D) and procedural recording (T) have been discussed in existing reproducibility improvement proposals; but making Γ explicit—as the sharing of boundary conditions that specify what problem is being addressed—has not been treated systematically in existing discourse.

This diagnostic function constitutes the ground for positioning SRF not as a merely descriptive framework but as one that imposes structural explanatory constraints on failures of problem identity. The case analyses serve as structural illustrations of DTC collapse patterns, not as empirical evidence for their distribution; the structural implications drawn in Section 8 rest on the formal properties of the framework, not on the frequency of the cases examined here.

7. The Distinction Between Failure of Reproducibility and Failure of Problem Identity

This section states the central Distinction Proposition, traces the formal chain that supports it, contrasts it with prior accounts, and draws the overall conclusion.

7.1 The Distinction Proposition

> **Distinction Proposition**
>
> Failure of reproducibility (obtaining different results for the same problem) and failure of problem identity (not addressing the same problem in the first place) are formally distinguishable; the latter is a precondition of the former. Failure of problem identity is identified as DTC collapse; reproducibility assessment is meaningful only on the presupposition that problem identity holds.

This proposition formally establishes the "tacit presupposition of problem identity" identified in Section 1 by means of Definition 2, Proposition 1, and Proposition 3, and demonstrates its applicability via the case analyses of Section 6. In existing literature this distinction is at best suggested conceptually; no example of its formalisation as a structural condition exists, including in the erotetic-logic frameworks of Hintikka (1976) and Sintonen (1984). The contribution of this paper lies in formalising this intuition as explicit structural conditions.

7.2 The Theoretical Chain Underlying the Distinction Proposition

The proposition rests on a tight chain of formal results. Definition 2 (structural isomorphism) defines "being the same problem" as $M_1 \cong M_2$. Remarks 1–3 establish that the holding of this isomorphism is definitionally equivalent to the simultaneous holding of D, T, and C. Proposition 1 (independence of Γ) formally establishes that even when D and T hold, failure of C alone constitutes failure of problem identity. Proposition 3 (DTC minimality) establishes that D, T, and C are each independently indispensable: the collapse of any single one is sufficient to cause failure of problem identity.

7.3 Contrast with Prior Accounts

Leonelli (2018) distinguishes three levels of reproducibility but does not formalise the structural conditions implied by the identity of problems at each level. SRF fills this gap with the three

DTC conditions, supplying the missing formal layer beneath Leonelli's taxonomy.

In Ioannidis (2005) the identity of the problem space functions as a tacit presupposition. SRF formalises this circular presupposition as DTC conditions and positions problem identity as the antecedent condition for reproducibility assessment. The effect is to transform an implicit and unexamined assumption into a diagnosable structural requirement.

What existing reproducibility discourse treated as a single phenomenon under "failure of reproducibility" in fact contains two structurally distinct failures. The first is genuine failure of reproducibility: different results for an identical problem. The second is failure of problem identity: not addressing the same problem in the first place. The four case studies of Section 6 showed that the second type actually occurs in the reproducibility problems of contemporary science and is indistinguishable by natural-language analysis alone.

7.4 Conclusion of This Section

> **Central Thesis. Reproducibility evaluation presupposes problem identity.**
>
> Problem identity is defined as the simultaneous holding of the three DTC conditions; the collapse of any single one entails failure of problem identity. Failure of reproducibility and failure of problem identity are systematically distinguishable by the formal criteria of SRF; SPO provides the ontological foundation for this distinction.

8. Conclusion

8.1 Structure of the Argument

This paper is structured as a linear argument in eight steps.

1. Reproducibility discourse (Leonelli 2018; Ioannidis 2005) presupposes problem identity, but this presupposition is not made explicit. (§ 1)

2. Natural-language descriptions (same data, same method, same result) cannot guarantee the conditions for problem identity. (§ 2)

3. A problem can be formally characterised as $M = (S, \tau, \Gamma)$. (§ 3)

4. Problem identity is defined as $M_1 \cong M_2$ and is equivalent to the simultaneous holding of the three DTC conditions. (§ 4)

5. Under Condition E, the three DTC conditions guarantee reconstructability R. (§ 5)

6. SRF carries ontological implications interpretable as SPO. (§ 5, final subsection)

7. In four case studies, SRF analysis achieved structural identification of failure of problem identity. (§ 6)

8. Failure of reproducibility and failure of problem identity are formally distinguishable; the latter is the precondition of the former. (§ 7)

The role of SRF is not to present an ideal of problem identity but to diagnose the structural factors in which problem identity fails. D-, T-, and C-collapse are the typology of this structural failure; SRF is positioned as a framework that imposes structural explanatory constraints on failures of problem identity.

8.2 Scholarly Positioning

The mathematical content is an application of the standard model-theoretic framework of structural isomorphism (Hodges 1993); no mathematical novelty is claimed. The philosophical contributions are fourfold.

1. **Proposal of Structural Problem Ontology.** Positioning problems as entities individuated by structure provides a formal ontological foundation for the theory of problems that Nickles and Bromberger discussed conceptually. This is the first framework to formally describe the individuation conditions, granularity, and the persistence and disappearance of problem identity through scientific change.

2. **Establishment of the Bridging Proposition.** The formal chain according to which the structural conditions for problem identity suffice for reconstructability under Condition E has not been made explicit in existing accounts.

3. **Derivation and structural illustration of the Distinction Proposition.** The formal distinction between "failure of reproducibility" and "failure of problem identity" is established within reproducibility discourse, and its applicability to contemporary science is demonstrated through SRF analysis of four case studies.

4. **Establishment of SRF's explanatory constraints and structural implications.** Any attempt to explain reproducibility failure is insufficient if it does not take into account the possibility of Γ-branching (explanatory constraint). Reproducibility failure is classifiable into the finite patterns of DTC collapse (structural implication).

8.3 Future Directions

1. **Proof of DTC necessity.** The general proof of $R \Rightarrow$ DTC is a natural extension of SRF.

2. **Extension to graded identity (SRF-g).** Relaxing the binary definition of exact isomorphism to admit structural approximation is pursuable as a connection to the Klir approach.

3. **Extension to the temporal axis (SRF-t).** An extension that equips τ with a partial-order structure and treats temporal separation of preservation and reconstruction as a formal object is possible.

4. **Connection between SPO and Fine's essentialism.** Formalising Γ as Fine's essential truths would further refine the ontological foundations of SPO.

5. **Relaxing perspectival fixing.** Analysis of re-identification conditions under perspectival change constitutes a separate research task (Note 1).

Notes

Note 1: Perspectival fixing is not the claim that the perspective is always fixed in practice. It is a methodological measure for isolating the formal conditions for problem identity; the scope of this paper is limited to these conditions. Problem identity under perspectival shift is an extension task for the next study in this research programme, to be addressed through the introduction of weak structural similarity; the second paper of the trilogy is planned to treat this as its primary subject.

Note 2: This paper does not claim that $M = (S, \tau, \Gamma)$ is the unique possible decomposition. For example, $M' = (S, \tau, p, \Gamma)$ incorporating stochastic transitions could be effective for other purposes. The position of this paper is limited to the claim that the three components provide a sufficient and mutually independent framework for the purposes of this paper.

Note 3: Remarks 1–3 are presented as "nominal assignments of the preservation conditions readable from Definition 2." They are not new independent theorems; they are definitional remarks that serve as bridges to the Bridging Proposition of Section 5.

References

[1] Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533.

[2] Bouthillier, X., Delaunay, P., Bronzi, M., Trofimov, A., Nichyporuk, B., Szeto, J., Mohammadi, S., Beckers, N., Kahou, S. E., & Vincent, P. (2021). Accounting for variance in machine learning benchmarks. *Proceedings of Machine Learning and Systems*, *3*, 747–769.

[3] Bromberger, S. (1966). Why-questions. In R. Colodny (Ed.), *Mind and Cosmos* (pp. 86–111). Pittsburgh: University of Pittsburgh Press.

[4] Fine, K. (1994). Essence and modality. *Philosophical Perspectives*, *8*, 1–16.

[5] Hintikka, J. (1976). *The Semantics of Questions and the Questions of Semantics*. Amsterdam: North-Holland.

[6] Hodges, W. (1993). *Model Theory*. Cambridge: Cambridge University Press.

[7] Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

[8] Klir, G. J. (1985). *Architecture of Systems Problem Solving*. New York: Plenum Press.

[9] Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

[10] Leonelli, S. (2018). Rethinking reproducibility as a criterion for research quality. *Philosophy of Science*, *85*(5), 1295–1307.

[11] Nickles, T. (1981). What is a problem that we may solve it? *Synthese*, *47*(1), 85–118.

[12] Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

[13] Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: MIT Press.

[14] Sintonen, M. (1984). *The Pragmatics of Scientific Explanation*. Helsinki: Finnish Society for Philosophy.

[15] Suppes, P. (2002). *Representation and Invariance of Scientific Structures*. Stanford: CSLI Publications.

[16] Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica*, *43*(1–2), 99–124.