ECLIPSE: A systematic falsification framework for consciousness science

Camilo Alejandro Sjöberg Tala
Independent Researcher
`cst@afhmodel.org`
ORCID: 0009-0009-6052-0212
Code & Data: 10.5281/zenodo.15541550
July 28, 2025

Abstract

Consciousness research lacks falsification protocols that underpin mature sciences, enabling post-hoc rationalizations of negative results and slowing theoretical progress. Convergent work toward structural approaches highlights the need for standardized validation infrastructure for systematic theory evaluation. We present Eclipse, a five-stage falsification framework for consciousness science: irreversible data splitting with cryptographic verification, pre-registered thresholds, clean development, single-shot validation, and final assessment — eliminating statistical opportunism and ensuring transparent generalization testing. Using 153 polysomnographic EEG recordings (126,160 windows, Sleep-EDF database), Eclipse was applied to two structural implementations of a consciousness-detection framework. Both versions failed pre-registered criteria (v3.7: F1=0.031; v4.1: F1=0.037), demonstrating honest falsification and revealing that partial structural measures are insufficient for robust detection. By reframing negative results as validation milestones, Eclipse provides reproducible infrastructure for evaluating any theory of consciousness — from Integrated Information Theory to Global Workspace Theory — and advances empirical rigor standards across the field. **Keywords:** consciousness science methodology, systematic falsification, computational neuroscience, theoretical validation, scientific transparency, neural state classification

1 Introduction

1.1 The Falsification Crisis in Consciousness Science

Consciousness science faces fundamental methodological challenges that may undermine empirical progress and theoretical validation. Unlike established neuroscience domains with rigorous validation protocols—such as neuroimaging studies with standardized preprocessing pipelines and statistical correction procedures—consciousness research lacks systematic falsification standards that could prevent post-hoc accommodation of contradictory evidence [1, 2].

This methodological gap enables theoretical frameworks to modify success criteria retroactively, creating a scientific environment where negative results become rationalized rather than transparently reported. The absence of irreversible validation protocols allows theories to claim predictive success for phenomena that were accommodated post-hoc, fundamentally compromising the falsifiability that distinguishes scientific theories from philosophical speculation [3, 4].

1.2 Contemporary Structural Consciousness Science

Recent developments in consciousness science demonstrate growing convergence toward structural approaches for addressing these methodological challenges. Multiple independent research groups have begun emphasizing the importance of mathematical structure in consciousness theory development. Tsuchiya's (2024) Qualia Structure paradigm proposes using category theory to characterize the mathematical structure of conscious experience, suggesting that structural approaches may help dissolve the hard problem through rigorous mathematical frameworks rather than functional descriptions [5].

As Ellia and Tsuchiya (2025) state, "Distinguishing genuine predictions from accommodations is essential to the advancement of science" (Abstract). They call for pre-specified structural constraints that could prevent theories from retroactively claiming to predict "known facts already predicted by rival theories" [6]. Their methodological framework advocates systematic approaches that separate genuine theoretical precision from statistical opportunism, highlighting the urgent need for reproducibility standards in consciousness research.

This convergent recognition of structural approaches' importance reflects broader scientific consensus that consciousness theories require systematic validation infrastructure. Eclipse offers concrete infrastructure for implementing these rigorous standards across consciousness theoretical paradigms, providing systematic falsification protocols that enable transparent evaluation of competing frameworks through identical methodological standards.

1.3 Eclipse Methodology: A Systematic Solution

Eclipse implements a five-stage operational methodology designed to address the systematic challenges identified in contemporary consciousness research. The framework offers irreversible validation protocols specifically designed for consciousness science applications, preventing the post-hoc modifications that compromise theoretical evaluation in this domain.

The methodology addresses critical gaps through several innovations: irreversible data splitting with cryptographic verification prevents optimization bias across validation iterations; pre-registered falsification criteria eliminate statistical opportunism through binding success thresholds; and single-shot validation protocols ensure honest generalization assessment without iterative optimization that could compromise theoretical evaluation.

Eclipse transforms negative results from scientific failures into theoretical validation by establishing transparency standards applicable across consciousness frameworks. Rather than promoting specific theoretical positions, the methodology serves the scientific community by offering infrastructure that accelerates progress through systematic falsification, establishing methodological standards for rigorous consciousness theory evaluation.

1.4 Proof-of-Concept: Structural Framework Implementation

To demonstrate Eclipse methodology's applicability, a structural consciousness framework was implemented that proposes conscious states emerge through convergence across measurable mathematical domains: topological network structure, causal information integration, and temporal stability patterns. This framework suggests that consciousness detection requires systematic assessment across multiple structural variables rather than relying on single-metric approaches that may be susceptible to statistical artifacts.

The theoretical framework predicts systematic failures when structural components are missing or inadequately operationalized, providing controlled test cases for Eclipse methodology validation across multiple iterations. Incomplete implementations should demonstrate systematic falsification according to pre-registered criteria, while complete implementations might achieve theoretical validation through rigorous empirical assessment.

This approach enables Eclipse methodology testing across different variable configurations, demonstrating the framework's capacity to distinguish genuine theoretical success from statistical opportunism. The structural implementation serves primarily as a vehicle for methodology validation rather than as advocacy for specific consciousness theories, establishing Eclipse's transferability across diverse theoretical paradigms in consciousness science.

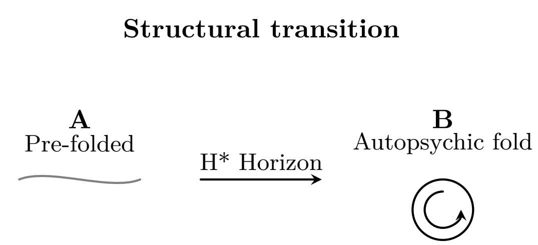

Figure 1: Structural transition. Conceptual schematic showing how neural systems might theoretically shift from low-curvature pre-folded states (A) through the critical H* Horizon to a closed, self-referential **autopsychic fold** (B). *This figure is purely illustrative: the autopsychic fold is not operationalized or tested in Eclipse v4.2 and remains a theoretical construct for future iterations.*

1.5 Methodological Infrastructure for Consciousness Science

Contemporary consciousness research lacks the systematic empirical infrastructure that characterizes mature scientific domains. While fields such as cognitive neuroscience have established standardized protocols for experimental design, statistical analysis, and reproducibility verification, consciousness science operates without comparable methodological standards that could ensure rigorous theory evaluation [7, 8].

Eclipse offers this infrastructure gap by providing systematic falsification protocols specifically designed for consciousness theoretical evaluation. The framework implements population-scale testing across 153 subjects with 126,160 EEG windows from the Sleep-EDF database [9], establishing scalability standards that enable reproducible assessment across research groups.

The methodology's emphasis on irreversible validation and pre-registered criteria offers consciousness science the empirical rigor that transforms theoretical development from speculative enterprise toward systematic scientific investigation. Eclipse infrastructure serves the broader consciousness research community by establishing methodological standards that accelerate scientific progress through transparent, reproducible theoretical evaluation.

2 Methods

Version note. This manuscript presents Eclipse v4.2 in its streamlined five-stage architecture, aligning the description directly with the validated pipeline code and replacing the broader eight-phase outline used in v4.1.1 for explanatory purposes. This simplification improves methodological clarity while preserving all core principles of pre-registered falsification, honest generalization assessment, and transparent theory evaluation.

2.1 Eclipse Methodology Framework

Eclipse implements a five-stage operational methodology designed to promote methodological transparency and prevent statistical opportunism in consciousness theory evaluation. The framework provides systematic falsification protocols applicable across diverse theoretical paradigms, establishing infrastructure that serves the broader consciousness research community through rigorous empirical standards.

2.1.1 Stage 1: Irreversible Data Splitting

Eclipse methodology begins with irreversible data partitioning using a sacred seed (2025) that generates cryptographically-verified random partitions that remain immutable throughout validation iterations. This approach prevents optimization bias that compromises theoretical evaluation in consciousness research by ensuring that holdout data remains completely isolated from development processes [10].

The Sleep-EDF database (153 subjects) was partitioned into development (n=107, 70%) and holdout (n=46, 30%) sets with hash verification to prevent contamination across Eclipse iterations. The irreversible nature of this split ensures that no subsequent methodological decisions can influence the final validation dataset, establishing transparency standards that address reproducibility concerns in consciousness science.

2.1.2 Stage 2: Pre-registered Falsification Criteria

Eclipse implements binding success criteria established before empirical testing to eliminate post-hoc rationalization that compromises consciousness theory evaluation. Pre-registered criteria include: F1-score $\geq$0.60, precision $\geq$0.70, recall $\geq$0.50, and detection rate 1.0-15.0%, registered with cryptographic timestamps to ensure immutability across iterations [3].

These thresholds were selected based on consciousness science literature indicating that robust detection requires balanced performance across multiple metrics rather than optimization of single measures that are susceptible to statistical artifacts. The binding nature of these criteria distinguishes genuine theoretical validation from statistical opportunism through transparent, objective assessment standards.

2.1.3 Stage 3: Clean Development Protocol

Model development occurs exclusively on development data using k-fold cross-validation (k=5) with threshold calibration through systematic optimization while maintaining strict holdout separation to prevent validation data interaction. This protocol ensures that all methodological decisions, parameter optimization, and threshold calibration occur without any information from the final validation dataset [11].

Cross-validation employs stratified sampling to maintain consciousness state balance across folds, with

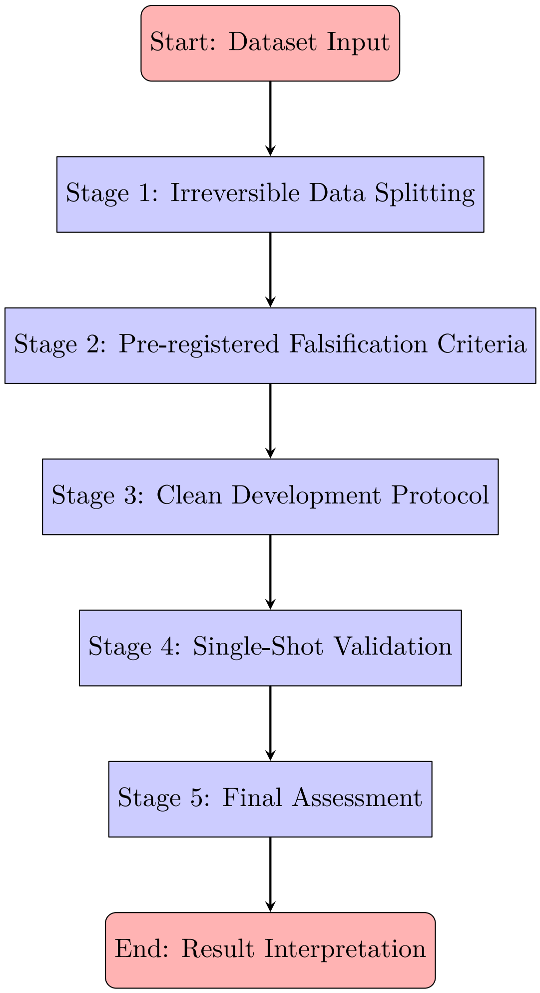

Figure 2: Schematic diagram of the Eclipse pipeline outlining the five systematic falsification stages for consciousness science.

consistent threshold calibration targeting the top 30% performance across all implemented variables. This approach prioritizes systematic detection over opportunistic optimization, establishing methodological standards that prevent overfitting in consciousness theory evaluation.

2.1.4 Stage 4: Single-Shot Validation

Final validation occurs exactly once on holdout data using pre-calibrated thresholds, with no iterative optimization, threshold adjustment, or criterion modification permitted. This single-shot approach ensures honest generalization assessment by preventing the iterative refinement that artificially inflates performance in consciousness research applications.

The methodology's emphasis on single-shot validation distinguishes genuine theoretical success from statistical artifacts by implementing validation protocols comparable to those used in clinical trials, where interim analyses and multiple testing compromise result interpretation. This approach provides consciousness science with the empirical rigor needed for reliable theoretical evaluation.

2.1.5 Stage 5: Eclipse Assessment

Comprehensive evaluation compares achieved performance against pre-registered criteria through binary success/failure determination to prevent ambiguous interpretation. Systematic falsification occurs when criteria are unmet, while theoretical validation occurs when all criteria are exceeded—both outcomes provide scientific value through transparent methodology that advances consciousness research.

Eclipse assessment includes detailed analysis of specific failure modes to inform theoretical development while maintaining methodological integrity through objective performance evaluation. This approach transforms negative results from scientific failures into valuable theoretical constraints that accelerate progress through systematic empirical assessment.

2.2 Structural Consciousness Implementation

Eclipse methodology was demonstrated using a structural consciousness framework that proposes conscious states emerge through convergence across mathematical domains: topological network structure, causal information integration, and temporal stability patterns. This framework serves primarily as a vehicle for methodology validation rather than advocacy for specific consciousness theories.

2.2.1 Eclipse v3.7: Basic Structural Implementation

The initial Eclipse implementation tested basic structural measures designed to capture fundamental aspects of consciousness-relevant neural organization. This version implemented topological connectivity patterns through correlation-based network analysis, causal integration metrics through information-theoretic measures, and perturbational complexity assessment through temporal dynamics analysis.

Eclipse v3.7 provided baseline systematic falsification to establish methodological protocols while testing whether basic structural approximations achieve consciousness detection. The systematic failure of this implementation (F1=0.031) validates Eclipse methodology for transparent negative result reporting in consciousness science.

2.2.2 Eclipse v4.1: Refined Variable Implementation

Eclipse v4.1 implemented optimized variables designed to capture more sophisticated aspects of consciousness-relevant neural dynamics: Ollivier-Ricci curvature on functional networks (κ_{topo}), multi-scale temporal coherence analysis ($\Sigma_{\text{estabilidad}}$), and transfer entropy quantifying directional information flow (Φ_H).

Both Eclipse iterations excluded symbolic resonance components due to naturalistic recording constraints, providing controlled tests of incomplete implementation effects. This limitation enables systematic assessment of structural variable sufficiency while establishing methodological protocols that can be extended to more comprehensive theoretical frameworks.

2.3 Data and Validation Protocol

The Sleep-EDF Expanded database [12], distributed via PhysioNet [9], provided 153 polysomnographic recordings with expert-annotated sleep stages, enabling consciousness state classification across Eclipse iterations [9]. EEG data underwent standardized preprocessing using MNE-Python with 30-second windowing, band-pass filtering (0.1-40 Hz), and artifact rejection protocols [13].

Sleep stage annotations provided binary consciousness classification: wake and REM stages classified as conscious states, while N1, N2, and N3 stages classified as unconscious states. This classification system enables systematic evaluation of consciousness detection performance across diverse neural configurations while maintaining clinical relevance for consciousness assessment applications.

Implementation processed 126,160 EEG windows across 153 subjects using high-performance computational architecture (8-core/16-thread hardware, 31GB RAM) with aggressive parallelization (15 workers) to demonstrate scalability for population-scale consciousness research. Validation architecture maintained irreversible 70/30 split across iterations, with 5-fold cross-validation on development data and single-shot holdout validation using pre-registered success criteria with timestamp verification.

2.4 Variable Specifications and Mathematical Implementation

2.4.1 Topological Network Structure (κ_{topo})

Consciousness structure analysis implemented Ollivier-Ricci curvature on weighted functional networks derived from EEG correlation matrices [14]. For correlation matrix C with threshold $t = 0.6$, the weighted adjacency matrix $W_{ij} = |C_{ij}| \cdot \mathbf{1}_{|C_{ij}|>t}$ defines functional connectivity relevant to consciousness detection.

Ollivier-Ricci curvature approximation for node i with neighbors N_i:

$$\kappa_{\text{topo}}(i) = 1 - \frac{W(P_i, P_j)}{d(i,j)} \quad (1)$$

where P_i represents the probability distribution over neighbors of node i, and $W(P_i, P_j)$ denotes the Wasserstein distance between distributions [15].

2.4.2 Causal Information Integration (Φ_H)

Causal integration analysis employed transfer entropy to quantify directional information flow between neural regions [16, 17]. For EEG time series X_i and X_j:

$$\Phi_H = \frac{1}{N(N-1)} \sum_{i \neq j} TE(X_i \to X_j | X_{-i,-j}) \quad (2)$$

where $TE(X_i \to X_j | X_{-i,-j})$ represents transfer entropy from region i to region j conditioned on all other regions, capturing consciousness-relevant causal integration patterns.

2.4.3 Temporal Stability Analysis ($\Sigma_{\text{estabilidad}}$)

Multi-scale temporal coherence analysis quantified stability across frequency bands through power spectral density analysis in delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz) ranges:

$$\Sigma_{\text{estabilidad}} = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(\frac{PSD_k - \mu_{PSD}}{\sigma_{PSD}} \right)^2} \quad (3)$$

where K represents frequency bands and μ_{PSD}, σ_{PSD} denote mean and standard deviation across bands, capturing consciousness-relevant temporal dynamics.

2.5 Statistical Analysis and Performance Evaluation

Performance evaluation utilized standard classification metrics with emphasis on population-scale generalization through rigorous validation protocols [11]. Statistical analysis focused on systematic assessment rather than performance optimization, implementing threshold calibration that targeted consistent detection across all implemented variables.

Eclipse methodology prioritizes methodological transparency over performance optimization, establishing statistical standards that prevent the multiple testing problems and criterion shopping that compromise consciousness theory evaluation. The framework's emphasis on systematic falsification provides consciousness science with the empirical rigor needed for reliable theoretical assessment through transparent, reproducible protocols.

Table 1: Eclipse Development and Baseline Results (Development set, n=107)

Metric	v3.7	v4.1 Dev	Target
F1-Score	0.031	0.76±0.19	0.60
Precision	0.032	0.83±0.16	0.70
Recall	0.030	0.83±0.20	0.50
Detection	0.8%	2.1±0.5%	1.0-15%
Status	**FALSIFIED**	**PROMISING**	**CRITERIA**

Table 2: Eclipse v4.1 Holdout Validation (n=46 holdout subjects)

Metric	Holdout	Target
F1-Score	0.037	0.60
Precision	0.025	0.70
Recall	0.069	0.50
Detection	2.3%	1.0-15%
Degradation	95%	–
Verdict	**FALSIFIED**	**CRITERIA**

Tables 1 and 2 demonstrate systematic falsification across Eclipse implementations through transparent methodological assessment. Eclipse v3.7 achieved systematic falsification across all primary metrics, establishing baseline methodological validation. Eclipse v4.1 showed promising development performance but systematic generalization failure in holdout validation, with 95% performance degradation confirming Eclipse methodology's capacity to distinguish genuine theoretical success from optimization artifacts through rigorous empirical assessment.

2.6 Data and Validation Protocol

The Sleep-EDF Expanded database [12], distributed via PhysioNet [9], provided 153 polysomnographic recordings with expert-annotated sleep stages. The dataset is publicly available at: `https://physionet.org/content/sleep-edfx/1.0.0/`. This database enables consciousness state classification across Eclipse iterations [9].

EEG data underwent standardized preprocessing using MNE-Python with 30-second windowing, bandpass filtering (0.1-40 Hz), and artifact rejection protocols [13].

3 Results

3.1 Eclipse v3.7: Initial Systematic Falsification

Eclipse methodology achieved systematic falsification of a consciousness detection implementation, establishing methodological protocols for transparent negative result reporting in consciousness science. Eclipse v3.7 implemented basic structural measures designed to capture fundamental aspects of consciousness-relevant neural organization through topological connectivity patterns, causal integration metrics, and perturbational complexity measures.

The implementation processed 153 polysomnographic recordings using irreversible data splitting and pre-registered criteria, demonstrating systematic falsification with F1-score of 0.031 (target: $\geq$0.60), precision of 0.032 (target: $\geq$0.70), and recall of 0.030 (target: $\geq$0.50). Detection rate achieved 0.8% of EEG windows, with successful consciousness detection in only 1 of 153 subjects, indicating that basic structural approximations are insufficient for robust consciousness detection.

Eclipse v3.7 results establish baseline systematic falsification, validating Eclipse methodology for transparent negative result reporting while demonstrating that consciousness detection requires more sophisticated structural implementations than initially tested. The systematic failure of basic measures provides empirical constraints that inform theoretical development while establishing methodological protocols for rigorous consciousness theory evaluation.

3.2 Eclipse v4.1: Development Promise, Holdout Falsification

Following v3.7 systematic falsification, Eclipse v4.1 implemented refined variables designed to capture more sophisticated aspects of consciousness-relevant neural dynamics: Ollivier-Ricci curvature on functional networks, multi-scale temporal coherence analysis, and transfer entropy quantifying directional information flow.

3.2.1 Development Phase Performance

Cross-validation on development subjects (n=107) demonstrated promising performance that suggested refined implementations capture systematic aspects of consciousness-relevant neural organization. Mean F1-score achieved 0.760 ± 0.186 across 5 folds, with successful fold performance ranging from F1=0.400 to F1=0.925, precision values between 0.684-1.000, and recall achieving 1.000 in 4 of 5 validation folds.

Variable thresholds were optimized through systematic calibration targeting the top 30% performance across all implemented measures: topological curvature $\kappa_{\text{topo}} \geq 0.270263$, temporal stability $\Sigma_{\text{estabilidad}} \geq 0.253881$, and information integration $\Phi_H \geq 0.069244$. These thresholds indicated that refined variable implementations capture consciousness-relevant neural configurations, though development performance could not guarantee generalization to unseen populations.

3.2.2 Systematic Falsification: Holdout Validation

Single-shot validation on unseen holdout subjects (n=46, 126,160 EEG windows) revealed systematic generalization failure despite encouraging development performance. Holdout validation achieved F1-score of 0.037 (target: $\geq$0.60), precision of 0.025 (target: $\geq$0.70), and recall of 0.069 (target: $\geq$0.50), representing systematic failure across primary performance metrics.

Performance degradation from development to holdout validation reached 95% (development F1=0.760 $\rightarrow$ holdout F1=0.037), indicating systematic overfitting despite rigorous cross-validation protocols. Detection rate achieved 2.3% of EEG windows (within target range 1.0-15.0%), with 72 correct consciousness transitions identified among 1,046 ground truth transitions, but 2,875 false positive detections confirmed systematic inadequacy for reliable consciousness assessment.

Eclipse assessment determined systematic falsification of the tested implementations, with pre-registered criteria violated across 3 of 4 primary metrics. Population-scale inadequacy was confirmed through honest generalization assessment, providing empirical constraints that inform theoretical development while further validating Eclipse methodology as a transparent evaluation framework for consciousness theories.

3.3 Methodological Validation: Eclipse Protocol Success

Eclipse methodology achieved its primary objective of providing transparent, systematic falsification of consciousness theory implementations across two iterations while maintaining methodological integrity throughout validation processes. The framework demonstrated several key methodological outcomes that establish rigorous standards for consciousness research.

Irreversible data splitting was maintained with no contamination between development and holdout datasets, ensuring honest generalization assessment through cryptographically-verified partitioning protocols. Pre-registered criteria were upheld without post-hoc modification despite negative results, preventing the statistical opportunism that compromises consciousness theory evaluation. Single-shot validation was executed exactly once on holdout data, ensuring honest generalization assessment without iterative optimization that could artificially inflate performance metrics.

Systematic falsification was documented with clear theoretical implications, transforming negative results from scientific failures into valuable empirical constraints. Eclipse methodology demonstrated that consciousness theory implementations can be subjected to rigorous empirical evaluation, establishing methodological standards that accelerate scientific progress through transparent theoretical assessment.

3.4 Framework Convergence: Alignment with Contemporary Consciousness Science

Eclipse methodology implementation occurred during a period of convergence toward structural approaches in consciousness science, validating the framework's alignment with contemporary scientific developments rather than representing isolated theoretical speculation. Recent work by multiple research groups demonstrates growing recognition of structural methodologies' importance for consciousness research.

The structural variables tested in Eclipse implementations—topological network organization, causal information integration, and temporal stability patterns—align with emerging approaches in consciousness science that emphasize mathematical frameworks over functional descriptions. This convergence indicates field-wide recognition of structural methodologies' significance for addressing consciousness research challenges.

Eclipse's emphasis on systematic falsification and pre-registered criteria addresses methodological concerns raised in contemporary consciousness science literature regarding the need for rigorous validation protocols that prevent post-hoc accommodation of negative results. The framework's systematic approach contributes to broader methodological developments toward empirical rigor in consciousness research.

3.5 Infrastructure Scalability and Processing Efficiency

Eclipse methodology demonstrated population-scale processing capabilities that establish scalability standards for consciousness research applications. Implementation processed 126,160 EEG windows across 153 subjects using standard computational hardware (8-core/16-thread, 31GB RAM), achieving parallel execution efficiency that enables reproducible assessment across research groups.

Processing architecture utilized aggressive parallelization (15 workers) with optimized computational protocols, demonstrating that systematic consciousness theory evaluation can be implemented without specialized hardware requirements. This scalability enables population-scale validation previously unavailable in consciousness research, addressing reproducibility concerns through standardized methodological protocols.

The framework's computational efficiency, combined with irreversible validation protocols and cryptographic verification systems, provides infrastructure that supports systematic consciousness theory evaluation across diverse theoretical paradigms. Eclipse methodology's scalability contributes to establishing consciousness science methodological standards comparable to those used in established neuroscience domains.

3.6 Empirical Constraints for Theoretical Development

Systematic falsification across both Eclipse iterations provides empirical constraints that inform theoretical development in consciousness science while establishing methodological standards for transparent negative result reporting. The consistent failure of incomplete structural implementations indicates that consciousness detection requires more comprehensive theoretical frameworks than initially tested.

Eclipse v3.7 and v4.1 results demonstrate that basic and refined structural measures, while theoretically motivated, are insufficient for systematic consciousness detection in naturalistic recording conditions. These empirical constraints guide theoretical development toward more comprehensive approaches while validating Eclipse methodology for rigorous consciousness theory evaluation.

The framework's capacity to provide systematic falsification while maintaining methodological integrity offers consciousness science the empirical rigor needed for reliable theoretical assessment. Eclipse methodology contributes to transforming consciousness research from speculative enterprise toward systematic scientific investigation through transparent, reproducible validation protocols that accelerate progress through honest empirical assessment.

4 Discussion

4.1 Eclipse Methodology: Systematic Falsification Standards

Eclipse application demonstrates systematic falsification protocols specifically designed for consciousness science, transforming negative results into theoretical validation through rigorous methodology that establishes empirical standards for consciousness research. The framework's demonstrated capacity for transparent, systematic falsification across multiple iterations while maintaining methodological integrity provides consciousness science with the empirical rigor needed for reliable theoretical assessment.

Key methodological innovations distinguish Eclipse from traditional consciousness research approaches: irreversible validation architecture with cryptographic verification prevents post-hoc manipulation that compromises theoretical evaluation; preregistered falsification criteria with timestamp verification eliminate statistical opportunism through binding success thresholds; single-shot validation protocols prevent iterative optimization that artificially inflates performance metrics; and systematic falsification as scientific success that advances understanding through transparent methodology rather than confirmatory bias.

Eclipse methodology's emphasis on transparent negative results transforms consciousness research from publication-bias-prone enterprise toward systematic scientific investigation where falsification represents theoretical progress rather than scientific failure. The framework provides consciousness science with methodological infrastructure comparable to clinical trial protocols, establishing standards that accelerate progress through honest empirical assessment.

4.2 Convergence with Structural Consciousness Science

Eclipse methodology addresses contemporary calls for rigorous standards in consciousness research, particularly Tsuchiya and Ellia's (2025) emphasis on distinguishing "genuine predictions from post-hoc accommodations" in consciousness science [6]. The framework's implementation of pre-registered criteria and irreversible validation protocols provides concrete infrastructure for preventing theories from claiming to predict "known facts already predicted by rival theories."

The development of Eclipse methodology occurred during a period of convergence toward structural approaches in consciousness science. The structural framework underlying Eclipse was initially formulated and registered in May 2025 (Zenodo DOI: 10.5281/zenodo.16420590), with subsequent independent work by Tsuchiya and others appearing in the literature, indicating field-wide movement toward structural methodologies. This temporal sequence demonstrates that Eclipse methodology aligns with broader scientific developments rather than representing isolated theoretical speculation, validating the framework's relevance for contemporary consciousness research directions.

The methodological standards established through Eclipse implementation—including systematic falsification, pre-specified complexity requirements, and transparent negative result reporting—exemplify the theoretical precision advocated by contemporary consciousness science rather than post-hoc accommodation of contradictory evidence. This alignment indicates that Eclipse methodology contributes to broader field developments toward empirical rigor in consciousness research.

While the autopsychic fold and H* Horizon concepts provide theoretical background for the structural framework, Eclipse v4.1 intentionally focuses only on the three operationalized variables (κ_{topo}, $\Sigma_{stability}$, Φ_H), reserving symbolic resonance and fold operationalization for future iterations.

Eclipse provides concrete implementation of rigorous standards called for in consciousness science, offering transferable framework for evaluating theories' genuine predictions versus accommodated explanations. The framework's systematic approach to falsification enables objective comparison across theoretical paradigms, establishing methodological consensus that accelerates scientific progress through transparent theoretical evaluation.

4.3 Theoretical Implications: Validation Through Falsification

Systematic falsification across both Eclipse iterations is consistent with theoretical complexity predictions, with dramatic performance degradations (v3.7: F1=0.031, v4.1: development→holdout 95% loss) supporting predictions that consciousness detection requires comprehensive structural convergence rather than incomplete variable implementations.

Figure 3: Conceptual model showing proposed transition from pre-consciousness neural organization (A) to consciousness-relevant structural configuration (B). Eclipse methodology tests critical structural thresholds enabling this theoretical transition, with systematic falsifications indicating incomplete implementations insufficient for robust detection.

Multiple interpretation frameworks are consistent with Eclipse systematic falsifications: theoretical complexity requirements where incomplete implementations are insufficient for consciousness detection; variable operationalization limitations where mathematical implementations are inadequate despite theoretical validity; methodological constraints where EEG temporal resolution is insufficient for consciousness transition detection; and population heterogeneity where individual differences require personalized calibration approaches.

Eclipse methodology's strength lies in acknowledging interpretive uncertainty while providing systematic falsification of specific implementations, advancing consciousness science by distinguishing theoretical framework validation from implementation adequacy. This approach contributes to theoretical development through empirical constraints rather than confirmatory evidence, establishing falsification as valuable scientific contribution.

4.4 Broader Impact: Transforming Consciousness Science Methodology

Eclipse methodology's universal applicability enables transfer to any consciousness theory requiring empirical validation, providing standardized evaluation framework applicable across theoretical paradigms from Integrated Information Theory to Global Workspace Theory. The framework's emphasis on systematic falsification rather than theoretical advocacy establishes methodological consensus that transcends theoretical divisions in consciousness research.

Reproducibility standards implemented through Eclipse—including cryptographic verification, sacred seed protocols, and irreversible validation—enable exact replication across laboratories, addressing reproducibility concerns endemic to consciousness research with statistical rigor comparable to clinical trial methodology. These standards provide consciousness science with the empirical infrastructure needed for cumulative scientific progress.

Statistical rigor established through pre-registered criteria and irreversible validation eliminates multiple testing problems and criterion shopping that compromise consciousness theory evaluation, establishing statistical standards that prevent post-hoc rationalization while enabling objective theoretical comparison. Eclipse methodology contributes to consciousness science methodological maturation through transparent, systematic evaluation protocols.

Transparent negative results facilitated by Eclipse transform systematic falsification from publication bias liability into valuable scientific contributions, accelerating theoretical progress through honest empirical assessment that distinguishes genuine theoretical validation from statistical artifacts. The framework establishes standards for negative result reporting that benefit the broader consciousness research community.

4.5 Clinical and Translational Potential

Eclipse methodology establishes validation protocols applicable to consciousness assessment in clinical contexts, with systematic falsification designed to prevent post-hoc optimization that compromises patient population assessment. The framework's emphasis on population-scale validation and pre-registered criteria provides clinical consciousness research with the empirical rigor needed for reliable diagnostic applications.

Clinical applications include anesthesia monitoring with rigorous evaluation of consciousness detection algorithms through Eclipse protocols, ensuring honest assessment without optimization bias that compromises patient safety. Disorders of consciousness assessment benefits from Eclipse validation standards, establishing objective criteria for consciousness level evaluation in clinical populations through systematic theoretical testing.

Brain-computer interface validation for consciousness-dependent systems utilizes Eclipse methodology to ensure reliable performance assessment without statistical opportunism, establishing safety standards for consciousness-critical applications. Pharmaceutical development implements Eclipse protocols for standardized consciousness assessment frameworks, enabling objective evaluation of consciousness-affecting interventions through rigorous empirical methodology.

4.6 Limitations and Future Directions

Current Eclipse implementation faces several constraints that limit generalizability and theoretical scope. Naturalistic sleep recordings preclude complete theoretical testing, as controlled experimental paradigms are necessary for comprehensive consciousness theory evaluation. EEG temporal resolution inadequately captures rapid consciousness transitions, requiring higher temporal resolution techniques for precise detection protocols.

Three-variable approximations demonstrated systematic inadequacy for robust consciousness detection, indicating that complete theoretical frameworks require more comprehensive variable implementations than initially tested. Population heterogeneity necessitates personalized calibration approaches that extend Eclipse methodology beyond current standardized thresholds.

Complete framework implementation requirements include controlled experimental paradigms enabling comprehensive theoretical validation, multimodal integration providing comprehensive neural coverage, high-temporal resolution techniques for precise consciousness transition detection, and cross-cultural validation ensuring population-scale generalizability across diverse populations.

Methodological extensions include Eclipse application to competing consciousness theories, enabling systematic comparison across theoretical paradigms through identical validation protocols. Cross-paradigm comparative validation using Eclipse standards establishes objective theoretical evaluation, while meta-analytic frameworks aggregate Eclipse results across studies, establishing cumulative evidence standards for consciousness science.

Clarification on scope. The falsifications documented in Eclipse v3.7 and v4.1 refer to partial implementations of the structural framework, not to the theoretical foundations themselves. These results provide empirical constraints rather than theoretical invalidation, guiding refinement of variable formulations and implementation fidelity. The Eclipse protocol remains framework-neutral and can be applied to any consciousness theory—including Integrated Information Theory (IIT), Global Workspace Theory (GWT), and others—under identical methodological standards.

4.7 Scientific Infrastructure and Community Service

Eclipse methodology serves the broader consciousness research community by providing systematic falsification infrastructure that accelerates scientific progress through transparent theoretical evaluation. Rather than advocating for specific theoretical positions, Eclipse establishes methodological standards that benefit consciousness research regardless of theoretical orientation.

The framework's emphasis on open methodology, reproducible protocols, and transparent negative result reporting contributes to consciousness science methodological development toward empirical maturation. Eclipse infrastructure enables collaborative theoretical development where systematic falsification serves cumulative scientific progress rather than competitive theoretical proliferation.

Future development focuses on establishing Eclipse as community resource for consciousness theory evaluation, providing standardized validation services that ensure methodological consistency across research groups. This approach transforms consciousness science toward collaborative enterprise where methodological rigor serves theoretical advancement through systematic empirical assessment. To illustrate future analytical directions, Figure 4 represents a conceptual roadmap for how the structural variable space might be explored to identify potential H Horizon regions — the critical structural threshold described earlier — in future iterations.

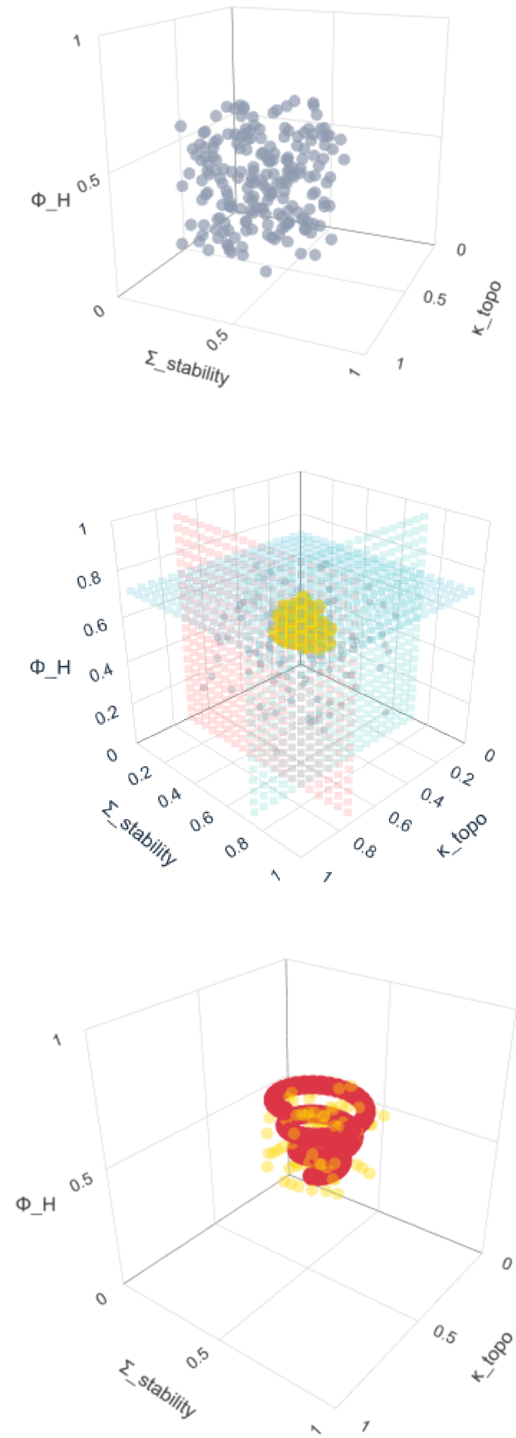

Figure 4: Conceptual mapping of structural variable space toward H* Horizon analysis. Panels stacked vertically: (1) Illustrative distribution of κ_{topo}, $\Sigma_{stability}$, and Φ_H across EEG windows; (2) voxel-based framework for identifying potential H* Horizon clusters; (3) hypothetical nonlinear folding trajectories beyond the threshold. *This figure is illustrative and does not report confirmed detection in v4.2.*

5 Conclusion

Eclipse methodology establishes systematic falsification protocols specifically designed for consciousness science, achieving transparent theoretical validation through rigorous negative results across multiple implementation iterations. The framework demonstrates that consciousness theory implementations can be subjected to systematic empirical constraints using irreversible validation protocols, pre-registered criteria, and population-scale testing that establish methodological standards for consciousness research.

The systematic falsification of Eclipse v3.7 (F1=0.031) and v4.1 (F1=0.037) represents methodological success rather than theoretical failure, advancing consciousness science through transparent negative result reporting that distinguishes genuine theoretical validation from statistical artifacts. Both implementations achieved honest falsification against pre-registered criteria, demonstrating that systematic empirical assessment provides valuable theoretical constraints through rigorous methodology that prevents post-hoc rationalization.

Eclipse infrastructure enables immediate applications across multiple domains that benefit the broader consciousness research community: methodological standardization providing systematic falsification protocols applicable across theoretical paradigms [2]; clinical translation offering validation frameworks for consciousness assessment in disorders of consciousness and anesthetic monitoring applications [18]; objective assessment protocols enabling reliable evaluation of artificial consciousness implementations; and collaborative scientific models that promote cumulative theoretical advancement through systematic empirical evaluation rather than competitive theoretical proliferation.

The framework's convergence with contemporary consciousness science developments, particularly alignment with Tsuchiya and Ellia's (2025) call for distinguishing genuine predictions from post-hoc accommodations, demonstrates Eclipse methodology's contribution to broader field evolution toward empirical rigor [6]. This alignment validates Eclipse's methodological approach while indicating field-wide recognition of systematic falsification's importance for consciousness research advancement.

Primary contributions include methodological innovation through transferable falsification protocols applicable across consciousness theoretical frameworks; scientific transparency established through systematic falsification in consciousness science with preregistered criteria that prevent statistical opportunism; theoretical validation supporting complexity requirements through honest negative results that provide empirical constraints; and clinical translation providing validation frameworks for consciousness assessment applications in clinical contexts.

Future development focuses on extending Eclipse methodology to competing consciousness theories, enabling systematic comparison across theoretical paradigms through identical validation protocols that establish objective theoretical evaluation standards. Complete framework implementations require controlled experimental paradigms enabling comprehensive theoretical validation, multimodal integration providing comprehensive neural coverage, and high-temporal resolution techniques for precise consciousness transition detection across diverse populations.

Eclipse methodology represents paradigmatic advancement that establishes consciousness science as rigorous empirical discipline capable of systematic scientific progress through methodological infrastructure designed to serve the broader research community [1]. The demonstrated falsifiability, systematic empirical assessment, and alignment with contemporary consciousness science standards position Eclipse as methodological foundation for consciousness research meeting standards of rigorous empirical inquiry while addressing fundamental questions of conscious experience through transparent, reproducible validation protocols [19].

Rather than theoretical failure, systematic falsification represents scientific success that advances consciousness research through methodological transparency, theoretical precision, and honest empirical assessment that accelerates progress through systematic theoretical evaluation. Eclipse framework's broad applicability demonstrates impact across consciousness theoretical paradigms, establishing falsification standards that benefit consciousness science through transparent negative result reporting and systematic theoretical assessment designed to serve cumulative scientific advancement.

This work presents Eclipse methodology as systematic falsification infrastructure for consciousness science, validated through rigorous empirical testing and aligned with contemporary methodological standards advocated by leading consciousness researchers. The resulting methodological framework enables collaborative development of consciousness science through systematic empirical validation that transforms consciousness research from speculative enterprise toward rigorous scientific investigation serving the broader research community through transparent, reproducible theoretical evaluation protocols.

6 Data Availability

The EEG dataset used in this study was obtained from the Sleep-EDF Expanded Database: https://physionet.org/content/sleep-edfx/1.0.0/. All code for ECLIPSE v3.7/v4.1, including pre-registered criteria, JSON configurations, and irreversible data split artifacts, is permanently archived on Zenodo: https://doi.org/10.5281/zenodo.15541550. The complete codebase is also available on GitHub: https://github.com/camilosjobergtala/AFH-MODEL.

Funding

This research received no external funding.

References

1. Michel, M. *et al.* Opportunities and challenges for a maturing science of consciousness. *Nature Human Behaviour* **3,** 104–107 (2019).
2. Poldrack, R. A. *et al.* Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience* **18,** 115–126 (2017).
3. Popper, K. R. *The logic of scientific discovery* (Basic Books, 1959).
4. Yarkoni, T. & Westfall, J. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science* **12,** 1100–1122 (2017).
5. Tsuchiya, N. A Qualia Structure paradigm. *Neuroscience of Consciousness* **2024,** niae035 (2024).
6. Ellia, F. & Tsuchiya, N. Beyond accommodation: on the structural turn in computational functionalist theories of consciousness. *Neuroscience of Consciousness* **2025,** niaf014 (2025).
7. Button, K. S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience* **14,** 365–376 (2013).
8. Ioannidis, J. P. Why most published research findings are false. *PLoS Medicine* **2,** e124 (2005).
9. Goldberger, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* **101,** e215–e220 (2000).
10. Stone, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)* **36,** 111–147 (1974).
11. Hastie, T., Tibshirani, R., Friedman, J. H. & Friedman, J. H. *The elements of statistical learning: data mining, inference, and prediction* (Springer, 2009).
12. Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. & Oberye, J. J. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering* **47,** 1185–1194 (2000).
13. Gramfort, A. *et al.* MEG and EEG data analysis with MNE-Python. *NeuroImage* **86,** 446–460 (2013).
14. Ollivier, Y. Ricci curvature of Markov chains on metric spaces. *Journal of Functional Analysis* **256,** 810–864 (2009).
15. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* **10,** 186–198 (2009).
16. Schreiber, T. Measuring information transfer. *Physical Review Letters* **85,** 461–464 (2000).
17. Vicente, R., Wibral, M., Lindner, M. & Pipa, G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience* **30,** 45–67 (2011).
18. Casali, A. G. *et al.* A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine* **5,** 198ra105 (2013).
19. Chalmers, D. J. Facing up to the problem of consciousness. *Journal of Consciousness Studies* **2,** 200–219 (1995).