Universal Collapse Theory: Two Worked Demonstrations

The Research Stack and the Applied Stack, Worked — and the Line Between Them

Flagship Demonstration · read alongside the *Program Map*

Jeremy C. Jones
HoldingLight LLC — ORCID 0009-0007-2515-3774 — universalcollapse.com

Version: v1.0 (2026-06) | License: CC BY 4.0
DOI: 10.17605/OSF.IO/Z7HY2

What this document is

A program map states a program's architecture; a worked demonstration shows the architecture doing something. This document is the second. It walks two cases end to end — one on the research side of Universal Collapse Theory (UCT), one on the applied side — so a reader can see, concretely, how the same coherence-audit logic operates in each, and why the two are kept on opposite sides of a deliberate line.

Part A is the **research-stack** flagship. It follows one structural signature — S_2 — down its full vertical: a formal bound proved under stated assumptions, an audit protocol that turns the bound into a field test with a sharp falsifier, and a live empirical run on real data. It is chosen deliberately because that run recorded a *failure* alongside its support, which is what a research flagship should show.

Part B is the **applied-stack** flagship. It walks one reverse-audit from the AI Integrity Protocol (AIP) — the same coherence-audit logic translated into a diagnostic for deployed AI and algorithmic socio-technical systems — applied to a public, recognizable case. It shows what the method surfaces that a narrower review is not designed to adjudicate, and the claim discipline it runs under.

Between them sits the line. The two demonstrations use related machinery, but they do *not* have the same evidentiary role: Part A counts toward the program's research ledger; Part B does not, unless independently audited under the same public standards. Showing both, with the line stated between them, is the point of pairing them — it demonstrates the firewall the program asserts elsewhere, rather than merely claiming it.

> **What this document does not claim**
>
> This document does not prove the UCT kernel or validate the Law of Coherence. Part A reports one empirical run — including a failed prediction — not a confirmed theory. Part B reports a structural diagnostic, not a legal, regulatory, or fairness finding, and its commercial-side results are not offered as evidence for the research program. The pairing demonstrates how the program separates evidence from application; it does not collapse the two.

The two demonstrations at a glance

Before the worked detail, the whole document in one table. The rightmost columns are the point: the two demonstrations have different evidentiary roles, and this is where the firewall is stated plainly.

Demonstration	Stack	Claim type	Counts toward research ledger?	What it shows	What it does not show
Part A S_2 / COGITATE	Research	Level-3 empirical run	**Yes**	Proof → protocol → data, with a recorded failure inside it.	Does not validate UCT; it is one run.
Part B AIP / Apple Card	Applied	Structural diagnostic	**No** unless converted under the stated rule	Field utility and claim discipline on a real case.	Not a legal, fairness, or regulatory verdict.

Part A — The research stack, worked: S_2 from bound to data

The program's empirical core is a repeated three-layer structure: a formal bound, an audit protocol that operationalizes it, and an empirical run that can come out against it. Part A follows the S_2 signature — the claim that, in systems resolving competing outcomes by evidence accumulation under constraint, weaker effective bias yields longer expected resolution — down all three layers, on one dataset.

A.1 The formal bound

The S_2 bound is proved in the Technical Note *Neutrality Delays Resolution*. The simplest case is modeled as a one-dimensional drift-diffusion process with fixed symmetric absorbing boundaries and start point $z = 0$. As the drift magnitude $|\mu|$ decreases toward zero, the expected first-passage time $E[\tau]$ increases strictly monotonically and attains the diffusion-only ceiling a^2/σ^2 at zero drift. The proof is self-contained from the Kolmogorov backward equation. This is a Level-3 (domain-discriminator) result, and it can fail in two distinct ways that the program keeps separate: the formal bound fails mathematically if the first-passage analysis is incorrect or the monotonicity does not follow under its stated assumptions; the domain claim fails empirically if the assumptions hold but the observed field behavior does not match. A.2 and A.3 expose it to the second kind of failure.

A.2 The audit protocol

The Methods paper *Auditing Constraint Asymmetry* turns that bound into a field test. The standard mistake it guards against is circular: fitting the drift parameter to the same response-time distribution it is then used to explain. Instead, constraint asymmetry ΔK is specified independently, and two identifiable objects are estimated from non-circular channels — the latency ceiling $C = a^2/\sigma^2$ and the drift-link parameter λ. Four features of the predicted latency curve $E[\tau](\Delta K) = C \cdot \tanh(\lambda\Delta K)/(\lambda\Delta K)$ (with the $\Delta K = 0$ value defined by the continuous limit, $E = C$) are then compared against observation: the ceiling at $\Delta K = 0$, the near-ceiling curvature, symmetry around zero, and full-curve shape — with a fifth axis handling a known confound. The audit's sharp falsifier is the failure of the observed latency curve to peak near neutrality, or to match the predicted shape under the stated model and an independently audited asymmetry. The protocol is valid only when ΔK, the response coordinate, the formalism, and the predicted curve are specified without circularity.

A.2a Mapping note: why this is an S_2 test

The bridge from the formal claim to the empirical predictions is specific, and the demonstration should show enough of it that a reader need not leave the document. On the constraint-architecture model, finer discrimination under higher constraint load resolves later — the perceptual analog of S_2, where a signal near a constraint boundary collapses more slowly. Task relevance supplies the constraint: a task-relevant non-target must be discriminated against the target within its own stimulus category — a finer, higher-load decision near the target/non-target boundary — whereas the same stimulus when task-irrelevant carries no such boundary. The prediction (A) is therefore directional and controlled: within the *same stimulus category and the same electrode*, the relevant case should show later high-gamma (70–150 Hz) onset than the irrelevant case. The within-category, within-electrode design is what makes this an S_2 test rather than a generic latency comparison — it holds stimulus identity fixed, making task-imposed constraint the targeted variation rather than a category artifact. A clean disconfirmation of the *mapping* (as distinct from a null statistic) would be the effect reversing in direction, or vanishing once stimulus identity is controlled — either of which would show the latency difference was a category artifact, not a constraint effect.

A.3 The empirical run

The audit is run in the Tier-1.6 paper *Constraint-Dependent Perceptual Resolution*, a pre-specified reanalysis of the open COGITATE intracranial-EEG dataset (34 analyzable patients) — the dataset the COGITATE adversarial collaboration released after challenging key predictions of both Integrated Information Theory and Global Neuronal Workspace Theory. Three predictions were registered before analysis.

Prediction	Pre-specified claim	Outcome
A	Task-relevant stimuli resolve later than the same category when task-irrelevant.	**Supported** — 25/34; p = 0.007; +14.1 ms
B	Hysteresis across miniblock transitions.	**Not supported** — p = 0.748
C	Task modulation of duration-tracking.	**Inconclusive** — comparison-level only

Prediction A was supported: task-relevant non-target stimuli showed later high-gamma onset than the same stimulus category when task-irrelevant (25 of 34 patients in the predicted direction; Wilcoxon on subject means p = 0.007; subject-level median +14.1 ms; bootstrap 95% CI on the mean [+3.6, +19.8] ms), and the effect held across all four stimulus categories. Prediction B, concerning hysteresis across miniblock transitions, was *not* supported (p = 0.748). Prediction C showed a comparison-level effect but no subject-level effect, and was reported as inconclusive.

A.4 What Part A demonstrates

The point of Part A is not that S_2 "won." It is that the full vertical ran, on real data, with a result that could have come out otherwise — and partly did. One prediction was supported, one failed, one was inconclusive, and **all three are reported as first-class outcomes**. A flagship with a recorded failure inside it is doing the thing the program claims to do: exposing the layer to disconfirmation and keeping the full result, rather than curating the win. The supported result means something precisely because the failed one was published beside it. That is the research stack working as designed — and it is the kind of outcome that counts toward the program's ledger.

The line between the two demonstrations

Part A and Part B run related machinery. The AIP method that Part B applies contains, as its own internal layers, the same update-integrity audit and the same S-signature battery that the research stack formalizes — the connection is structural, not rhetorical. But relatedness of machinery is not equivalence of evidentiary role, and the program is explicit about the difference.

> **Evidence versus application**
>
> **Part A is research evidence.** It is a public, versioned, pre-specified empirical test with a recorded failure; it counts toward the program's research ledger and is subject to the program's falsification standards.
>
> **Part B is application.** It is a diagnostic translation of the same logic into a deployed-systems setting. It demonstrates that the method does real work in the field, but it is not evidence for the law-level program unless it is independently audited, versioned, and entered into the public ledger under the same standards as Part A. A commercial diagnostic result is not a research result.

Stating this line is what lets the two demonstrations sit in one document without contaminating each other. AIP is linked to UCT — a reader can now see that the audit method has commercial reach and that the research has applied consequence — but neither is absorbed into the other. The research is not validated by the commercial work; the commercial work requires no acceptance of the research program. The line is the membrane that makes the link safe.

The membrane is not one-way by accident; it has a stated crossing. **An applied audit becomes research evidence only if** its protocol is public and versioned, its analysis is independently reproducible, its predictions are pre-specified where the case allows, and its outcome is entered into the public ledger with failures, nulls, inaccessible layers, and inconclusives recorded under the same rules as Part A. Until those conditions are met, an applied result stays applied. That is the conversion rule, and it is deliberately strict: it is what keeps "the method works in the field" from quietly becoming "the theory is true."

Part B — The applied stack, worked: a reverse-audit of Apple Card

The AI Integrity Protocol is the coherence-audit logic translated into a third-party diagnostic for deployed AI and algorithmic socio-technical systems. It runs four layers: a Structural Inventory, an Update Integrity Audit (the UIS corruptor set), an S-Signature Test Battery, and an Integrity Ledger. Part B walks one worked application of that method — reverse-audit RA-001, applied to a public, recognizable algorithmic consumer-credit ecosystem at Tier D (public-record access only). The failures at issue are structural — in routing, record-handling, and update pathways — not claims about model behavior; the method audits the system around the model as much as the model itself.

B.1 Why this case

The Apple Card consumer-credit program (Apple Inc. / Goldman Sachs Bank USA, reviewed August 2019 through 2024) is a clarifying case because it separates two questions that are routinely conflated. One review looked at *underwriting*: in March 2021, the New York State Department of Financial Services analyzed underwriting data for roughly 400,000 New York applicants and found no fair-lending violation. A later enforcement action looked at *servicing*: in October 2024, the Consumer Financial Protection Bureau identified distinct servicing, dispute-routing, and dispute-investigation failures, ordering a $25 million civil penalty against Apple (File No. 2024-CFPB-0012) and a $45 million penalty plus $19.8 million in consumer redress against Goldman (File No. 2024-CFPB-0011). These belong to the same deployed product ecosystem, but they sit at **different layers** and test different integrity questions: a clean fair-lending finding at the underwriting layer does not speak to the integrity of the servicing and dispute layer, or the reverse.

This is the answer to the question a skeptic puts to any framework: *what does it let me see that I could not already?* A fairness-only review asks whether outcomes discriminate. The structural audit asks a different question — whether the system's records, dispute routing, explanation pathways, and update mechanisms remain coherent under disconfirming evidence — and that question surfaces failures a fairness-only review is not designed to adjudicate.

B.2 The structural findings

Applied to the public record, the audit located concrete update-integrity failures. Some were identifiable at or before the August 20, 2019 launch: an August 16, 2019 presentation to Goldman's board — four days before launch — reported that the disputes system was "not fully ready" due to technological issues (CFPB Goldman order ¶ 4), against a partnership incentive structure that penalized launch delay. Others surfaced later in the dispute-handling and record-preservation pathways. These are read through the Update Integrity Standard's named corruptor set — the same failure modes the research-side standard specifies — and reported per corruptor, with each finding tied to a documented public source. The mini-ledger below shows the structure of that sourcing; the full per-finding ledger is in RA-001.

Structural finding	UIS corruptor	Public source	What remains inaccessible (Tier D)
Disputes system not fully ready at launch.	Constraint freezing	CFPB Goldman order ¶ 4 (recites Aug 16, 2019 board presentation).	Internal readiness assessments and remediation timelines.

Structural finding	UIS corruptor	Public source	What remains inaccessible (Tier D)
Dispute routing and investigation failures.	Selective update	CFPB consent orders, Apple and Goldman, Oct 2024.	Internal routing logic and case-handling logs.
Record-preservation gaps in dispute handling.	Record falsification / loss	CFPB consent orders, Oct 2024.	System-level record schemas and retention policy.

UIS corruptor labels name structural failure modes; they do not imply intent unless the public record independently supports that claim.

Boundary note. The NY DFS fair-lending finding (no violation, underwriting layer) is *not* a structural finding and is deliberately excluded from the ledger above. It is recorded here only to mark what the structural audit does *not* adjudicate: it concerns a different layer (underwriting), reviewed by a different body, on a different question. The ledger makes the discipline visible — each structural finding is tied to a public source and paired with what Tier-D access cannot see. The audit reads structure off the record; it does not infer internal intent.

B.3 The claim discipline

What makes the audit a diagnostic rather than an accusation is the discipline it runs under, and the report states it on every axis:

- **Structural, not legal.** The findings do not constitute a fair-lending violation, a regulatory-compliance finding, or a finding of unlawful discrimination. They are findings about whether specified structural failure modes appeared in the public record.
- **Inaccessible is a finding.** At Tier D, internal systems are not observable. The report has a dedicated Inaccessible-Layer section: what could not be seen is named, not guessed at or quietly assumed.
- **Inconclusive is a finding.** Where the public record does not settle a question, the result is recorded as inconclusive rather than resolved in either direction — a first-class outcome with its own section.
- **The adjudicator holds the decision.** The protocol produces no score and no certification. It produces auditable structural findings; the deciding judgment about what they mean remains with a human reviewer.

B.4 What Part B demonstrates

Part B shows the same audit discipline the research stack runs on the bench, translated into the field: specifying failure modes in advance, testing against them, and reporting what it finds — including what it cannot see and cannot settle. It is worth separating the three kinds of value the demonstration produces, because they are easy to run together and the firewall depends on keeping them apart:

- **Diagnostic value.** It surfaces structural integrity failures — in records, routing, and update pathways — that a fairness-only review is not designed to adjudicate, on a case the public already knows.

- **Applied / commercial value.** That diagnostic reach is real and useful in deployment regardless of whether the law-level program ultimately succeeds, because the method is judged by whether its structural findings reproduce, not by the fate of the kernel that inspired it.
- **Research value.** None — unless the audit is independently reproduced, versioned, and entered into the public ledger under the conversion rule stated above. As a standalone applied demonstration, it does not count toward the research ledger.

Holding those three apart is the whole point of running Part B inside this document rather than letting it stand alone: it shows the method has genuine field reach *and* shows exactly why that reach is not, by itself, evidence for the theory.

Limitations and next tests

A flagship is strengthened, not weakened, by stating plainly what each demonstration does not establish and what would extend it.

- **Part A** is one signature, one dataset, and one empirical run. It supports ledger inclusion, not program validation. Its next test is replication or extension under the same S_2 audit conditions — an independent reanalysis, or another dataset meeting the same constraint-asymmetry specification. It would be counted against the demonstration if Prediction A failed to replicate under those conditions, or if the independent ΔK audit were shown to be circular or confounded.
- **Part B** is a Tier-D public-record diagnostic. It cannot see internal systems, cannot establish legal liability, and does not count as research evidence unless converted under the rule above. Its next test is reproducibility across cases and reviewers. It would be counted against the demonstration if independent auditors could not reproduce the structural findings from the same public record, or if any finding were shown to depend on inference beyond what Tier-D access supports.

Both next-tests are the per-layer falsification discipline of the *Program Map* applied to this document: a demonstration in good standing is one whose characteristic failure has been named and looked for.

What the pairing shows

Read together, the two demonstrations make one structural point that neither makes alone. The same coherence-audit logic runs on both sides of the program: as pre-specified empirical test on the research side, and as field diagnostic on the applied side. The machinery is shared — update-integrity auditing and the S-signature battery appear in both. But the two are held in different roles, and the document keeps them there: research evidence that counts toward the ledger, and applied translation that does not unless separately audited to the same standard.

That separation, demonstrated rather than asserted, is the credibility move. A program with a commercial wing invites the suspicion that the commercial work is being used to prop up the theory, or the theory to market the product. The pairing answers the suspicion concretely: here is the research stack, with a failure recorded inside it; here is the applied stack, doing real diagnostic work; and here is the line that keeps each

from standing in for the other. The link between UCT and AIP is made visible — a connection a reader would otherwise miss — without either collapsing into the other.

Neither demonstration settles the program. Part A is one run; the verdict on the research is longitudinal, read off the accumulating ledger over many such runs. Part B is one application; the method's standing is read off whether its structural findings reproduce. What the pairing settles is narrower and worth settling on its own: that the program can show its logic working, on real cases, on both sides of its own firewall — and can keep the sides apart while doing it. The pairing is therefore evidence of architectural discipline, not evidence that the law-level program is true.

Referenced artifacts and source base

Every artifact and public event this demonstration relies on, with its role and identifier, so the worked chain can be traced without leaving the document.

Item	Role in this paper	Identifier / source
TN-S_2 — Neutrality Delays Resolution	Formal bound (Part A)	10.17605/OSF.IO/6WRQV
Methods-S_2 — Auditing Constraint Asymmetry	Audit protocol (Part A)	10.17605/OSF.IO/HRKWT
Constraint-Dependent Perceptual Resolution (COGITATE reanalysis)	Empirical run (Part A)	10.17605/OSF.IO/MXYU2
AIP RA-001 — Apple Card (Tier D)	Applied reverse-audit (Part B)	10.17605/OSF.IO/K63UQ
NY DFS — Report on the Apple Card Investigation	Public fairness-review source (Part B; underwriting layer)	NY Dept. of Financial Services, Mar 2021 dfs.ny.gov — rpt_202103_apple_card_investigation
CFPB — Apple Inc. consent order	Public enforcement source (Part B; servicing layer)	File No. 2024-CFPB-0012, Oct 23, 2024 consumerfinance.gov/enforcement/actions/apple-inc
CFPB — Goldman Sachs Bank USA consent order	Public enforcement source (Part B; servicing layer)	File No. 2024-CFPB-0011, Oct 23, 2024 consumerfinance.gov/enforcement/actions/goldman-sachs-bank-usa

Research artifacts (TN-S_2, Methods-S_2, the COGITATE reanalysis) are live OSF deposits. The applied artifact (RA-001) is published through the AI Integrity Protocol and carries its own full per-finding ledger and source index. Regulatory items are public-record documents cited by issuing body and date.

UCT Library

This document is part of the Universal Collapse Theory library maintained by Jeremy C. Jones and HoldingLight LLC. It is the flagship demonstration companion to the Program Map; the research artifacts it walks (TN-S_2, Methods-S_2, the COGITATE reanalysis) are live deposits, and the applied artifact (AIP RA-001) is published through the AI Integrity Protocol. Roadmap: universalcollapse.com/roadmap

AI Disclosure

AI tools were used to assist with manuscript preparation, drafting, organization, and editorial refinement. The underlying theory, structural decisions, analysis, and conclusions are the author's own.

Suggested Citation

Jones, J. C. (2026). Universal Collapse Theory: Two Worked Demonstrations — The Research Stack and the Applied Stack (v1.0). HoldingLight LLC. https://doi.org/10.17605/OSF.IO/Z7HY2