NEL MIND

# Experimental Evidence Dossier

Autonoema research | NeuralEdge Limitless
Prepared for consideration alongside an Anthropic Research Residency application

Research architect: Charles Dooner
Evidence review date: 10 November 2026 (GMT)
Project: NEL Mind - experimental experience-dependent cognitive architecture

---

Scope. This document summarises ten reported successful diagnostic tests and a separately documented full Ubuntu reboot-and-restoration observation. Evidence consists of console output and selected test code supplied by the researcher; the work was not independently rerun on the authoring system. Descriptions distinguish confirmed assertions, controlled test conditions, interpretations, and remaining questions.

## Executive summary

The evidence supports a functioning research prototype with mechanisms for accumulating experiences, updating predictions, deriving internal action values, changing selections under controlled conditions, preserving learned state across process boundaries, and running a shared cognitive architecture across simulated environments. A controlled relational transfer test also found a different first action in World B following training in World A, relative to an untrained control.

On 2 November 2026, dated Beast-interface logs recorded a checkpoint-based continuation from 20 to 22 stored experiences after a separately confirmed host reboot. This establishes observed memory-count continuity across an actual Ubuntu restart. It does not by itself prove that every form of learned knowledge remained usable or that the hardware robot was physically operated.

Important limitations: experiments were predominantly controlled or simulated; one live endogenous-value diagnostic produced strict unique maxima on only 2 of 18 autonomous evaluation cycles; repeated failed actions remain an unresolved behavioural issue. Cross-world transfer was demonstrated in one structured relational task, not across arbitrary worlds.

## Evidence at a glance

| Area | Observed result | Interpretive boundary |
|---|---|---|
| Experience → selection | Known action value 0.08333; unknown 0.16667; selected action changed | Controlled synthetic action setting |
| History-dependent values | Identical current state, differing experience histories, differing selections | Training exposure controlled; no optimality claim |
| Cross-world relation | World B: control chose L (preserved); trained mind chose M (changed) | One benchmark; causation pathway not isolated |
| Persistence | Experience counts 20 → 22 after system reboot | File provenance and OS boot timings established; deeper learning state needs checking |
| Multi-world architecture | Shared loop and shared memory passed maze/filesystem diagnostics | Not proof of general transfer or autonomous competence |

## 1. Experimental inventory and traceability

A “PASS” below means the stated test assertions completed successfully in the researcher-provided execution output. The ten test cases were executed across nine command invocations, since two multi-world cases were run together. They are not ten independent scientific replications.

| ID | Experiment / source file | Recorded outcome |
|---|---|---|
| T01 | test_nelmind_learned_state_process_restore.py | PASS; separate-process restore and continued transformation learning |
| T02 | test_endogenous_value_history_divergence.py | 1 passed; distinct action choices after distinct histories |
| T03 | test_key_learning_live_endogenous_value_diagnostic.py | PASS; 2/2 strict-value maxima selected; 16 ties |
| T04 | test_nelmind_checkpoint_restart.py | PASS; checkpoint pending-work/order integrity |
| T05 | test_beast_runner_checkpoint_sessions.py | PASS; mock Beast session continuation and failure handling |
| T06 | test_experience_changes_action_selection.py | PASS; chosen action changed after one observed consequence |
| T07 | test_action_transformation_live_prediction_differentiation.py | PASS; experienced action 0.2 vs unknown 1.0 |
| T08 | test_relational_cross_world_first_action_transfer.py | 1 passed in 0.41 s; control L vs trained M |
| T09 | test_stage3b5_multi_world_generic_orchestration_diagnostic.py | PASS; generic loop over two runtime adapters |

| ID | Experiment / source file | Recorded outcome |
|---|---|---|
| T10 | test_unified_world_cognition_shared_memory.py | PASS; one mind stores both environment experiences |

Source filenames and selected terminal command lines are listed in Section 6. Raw local logs remain on the researcher's Ubuntu workstation; this PDF is an interpretive index, not a replacement for machine-readable artifacts.

# 2. Experience, prediction and action selection

### T06 - Experience changes action selection

Two initially equivalent actions were assigned information gain 0.5 and internal value 0.1666667. Following one recorded transition for action_experienced, its information gain decreased to 0.25 and value to 0.0833333. The still-unknown alternative retained information gain 0.5 and value 0.1666667, and became the selected action. The output explicitly reported no external reward, externally imposed goal, motor execution or Beast involvement. This is a direct controlled demonstration of experience-sensitive selection through the existing valuation/arbitration path.

### T07 - Transformation evidence affects predictions

Four consistent transformation observations were supplied for a known action. The prediction manager assigned information gain and expected uncertainty reduction of 0.2 for that action versus 1.0 for an unknown action. The test verified that predict() consumed this evidence. Input evidence was supplied by the test; this experiment alone does not establish autonomous discovery or subsequent selection.

### T02 - Distinct histories produce distinct action selection

Two initially equivalent NELMind instances received different controlled histories (18 exposure cycles per mind), then evaluated the same current state. The reported Mind A chose blue with blue/green values tied at 0.423515 (red 0.396160); Mind B chose red with value 0.017544 (blue/green 0.015351). The result supports history-dependent value differences and differing selections. Mind A's blue choice was not uniquely better than green.

### T03 - Live endogenous-value diagnostic: success and limitation

Following 30 test-controlled neutral exposure cycles, NEL was allowed 18 live arbitration cycles. In 2/18 cycles an action had a unique maximum internal value, and both selected actions matched that maximum. The other 16 cycles were ties; notably, cycles 4-18 repeatedly attempted an ineffective red-key action against a blue lock. Consequently the diagnostic supports active use of values when the ranking is unique, not consistently adaptive or successful behaviour.

# 3. Cross-world operation and relational transfer

### T09 - Shared generic runtime machinery

Stage 3B.5 ran five decision cycles in a filesystem validation world and five in a maze validation world. The generic orchestration loop used a common runtime contract without maze- or filesystem-specific logic inside the shared cycle. Filesystem actions opened apple.txt/banana.txt; maze actions moved north/west/south/east. Selected values and arbitration reasons were printed for each cycle. No robot motor intention occurred.

### T10 - Shared memory across environments

A single NELMind instance performed three filesystem cycles followed by three maze cycles. The diagnostic passed assertions that both sets of experiences remained present in the same mind's memory. Logged relationship counts progressed from 3 during initial filesystem processing to 58 by the final maze cycle. The test demonstrates unified memory representation under this setup; cross-world transfer is a separate hypothesis.

### T08 - Relational first-action transfer

| Condition | First World B action | World B transition |
|---|---|---|
| Fresh control mind | try:instrument_l | object_p → object_p |
| World A trained mind (30 cycles) | try:instrument_m | object_p → object_q |

The trained mind had acquired two asserted patterns: a shared-property participant relationship associated with state change, and a nonshared-property relationship associated with state preservation. Its World B pre-action predictions listed 26 pieces of evidence for a preserved outcome for instrument_l, 0 for the unknown instrument_m, and 26 for instrument_n. The fresh control chose L; the trained mind chose M and changed World B's state on its first action. Pytest reported 1 passed in 0.41 s.

Inference boundary: these observations support a specific experience-dependent cross-world first-action difference. They do not isolate the exact contribution of relational knowledge versus novelty, uncertainty, action order or other decision inputs. Follow-up ablation tests and arbitration/value traces are needed before assigning a stronger causal mechanism. Earlier, different cross-world transfer diagnostics reportedly produced "NO SIGNAL"; those should not be silently displaced by this positive result.

## 4. Persistence and restart integrity

### T01 - Separate-process learned-state restoration

A fresh writer process generated experience and saved a temporary learned-state snapshot; a later process restored it. The test reported restored hypotheses/expectations, no recounting of old evidence, correctly counted new familiar observations, continued transformation learning and retained provenance references. Output: PASS. This is separate-process restoration, not an actual OS reboot.

### T04 - Checkpoint restart/order integrity

A controlled checkpoint test verified pending assessments resumed in the intended order; later perceptions did not retroactively alter earlier assessments; stale cached action information was not inadvertently executed after restoration. Output: PASS.

### T05 - Beast session checkpoint integrity

A mock Beast interface ran a fresh session from 0 to 2 experiences and a restored session from 2 to 4; tests included checkpoint protection, injected save failure, observation timeout handling and corrupt-file handling. It did not require the physical Beast or establish full hardware control. Output: PASS.

## 5. Verified full Ubuntu reboot: 2 November 2026

Evidence classes: (a) dated per-cycle JSONL session records and (b) systemd journal boot history copied from the researcher's Ubuntu workstation. Unlike T01/T04/T05, this observation covers a host operating-system reboot.

| Time GMT | Source / session | Recorded state |
|---|---|---|
| 15:16:33-15:16:39 | New Beast-interface session | starting_memory_count=0; memory_count 2 → 10 |
| 15:17:08-15:17:14 | Restored session | starting_memory_count=10; memory_count 12 → 20 |
| 15:53:23 | systemd boot history | Previous boot’s last journal entry |
| 15:55:30 | systemd boot history | Next boot’s first journal entry |
| 17:43 | Beast-interface log | File exists; zero JSONL records |
| 18:03:51 | Restored session | starting_memory_count=20; memory_count=22 |
| 19:57:04 | Restored session | starting_memory_count=22; memory_count=24 |

Key finding: the operating-system boot boundary lies after the 15:17 restored session and before the 18:03 restored session. The post-reboot log explicitly reports session_kind="restored" and starting_memory_count=20; a new cycle increased the memory count to 22. At 19:57, a further session resumed with 22 and increased to 24. This is concrete evidence of persisted, subsequently extended experience records across a verified host restart.

Evidence beyond counts: the 18:03 record reports outcome_evidence_count=1 and a non-null predicted_outcome_entity_id for actuator_0_positive, corresponding to an earlier runtime state. This is suggestive of surviving predictive evidence, but no comprehensive before/after learned-model equivalence comparison was performed in this audit.

## Checkpoint files and fixed identifiers

The following paths were reported on the researcher’s workstation; the hashes came from a read-only sha256sum invocation and identify the bytes then observed. They do not prove the files’ contents or creation history without further inspection.

| Checkpoint / last modified (GMT) | SHA-256 |
|---|---|
| beast_before_reboot_20261102T155139887228Z.chk<br>2 Nov 15:17:14; approx. 250 KB | 4f5bd8102f8758af00f3ac2b3d087ba37c64dc11f3387b6c8537d5416278bc26 |
| beast_continuity_validation.chk<br>2 Nov 19:57:04; approx. 299 KB | 7e7c28b2ecf2631ab0c8ba08bb1a9a3d3278d22cb04ff7dcc1d94392e38aa73f |

Provenance caution: the restored session logs explicitly reference checkpoints/beast_continuity_validation.chk. The separately named beast_before_reboot_...chk was found on disk, but the reviewed cycle logs do not establish that it was the checkpoint loaded at 18:03. This report does not conflate the two. The later hash of a mutable checkpoint should not be represented as the hash of its earlier, pre-reboot version.

## System boot identifiers

Boot ending 2 Nov 15:53:23 GMT: 3acc355fdeb9482e9d363e261d1937e8. Next boot beginning 2 Nov 15:55:30 GMT: d32f6d5160b14d808301277bb5097a1f. Times derive from journalctl --list-boots --no-pager.

# 6. Artifact inventory and reproducibility

## Experiment commands and captured outputs

All paths below are relative to ~/nel_mind_v2 unless shown with ~/; PYTHONPATH values are as supplied in the researcher's command transcript. Each reported invocation used output capture to a timestamped text log when available.

| ID | Source test | Console evidence file |
|---|---|---|
| T01 | tests/test_nelmind_learned_state_process_restore.py | ~/nel_learned_state_evidence_2026-11-10.txt |
| T02 | tests/test_endogenous_value_history_divergence.py | ~/nel_value_history_evidence_2026-11-10.txt |
| T03 | tests/test_key_learning_live_endogenous_value_diagnostic.py | ~/nel_live_endogenous_value_evidence_2026-11-10.txt |
| T04 | tests/test_nelmind_checkpoint_restart.py | ~/nel_checkpoint_restart_evidence_2026-11-10.txt |
| T05 | tests/test_beast_runner_checkpoint_sessions.py | ~/nel_beast_session_evidence_2026-11-10.txt |
| T06 | tests/test_experience_changes_action_selection.py | ~/nel_experience_action_selection_evidence_2026-11-10.txt |
| T07 | tests/test_action_transformation_live_prediction_differentiation.py | ~/nel_prediction_differentiation_evidence_2026-11-10.txt |
| T08 | tests/test_relational_cross_world_first_action_transfer.py | ~/nel_cross_world_transfer_evidence_2026-11-10.txt |
| T09/T10 | tests/test_stage3b5_multi_world_generic_orchestration_diagnostic.py<br>tests/test_unified_world_cognition_shared_memory.py | ~/nel_multi_world_evidence_2026-11-10.txt |

## November 2 JSONL originals

• sandbox/beast_physical_world/logs/beast_continuous_20261102T151631694513Z.jsonl

• sandbox/beast_physical_world/logs/beast_continuous_20261102T151706648570Z.jsonl

• sandbox/beast_physical_world/logs/beast_continuous_20261102T174320345367Z.jsonl (empty)

• sandbox/beast_physical_world/logs/beast_continuous_20261102T180349838054Z.jsonl

• sandbox/beast_physical_world/logs/beast_continuous_20261102T195702615261Z.jsonl

The command-line journal excerpt and checksum output are also part of the 10 November 2026 research conversation. Preserve raw JSONL files, checkpoint snapshots, source revision, Python dependency lockfile/environment, and complete journal export as a reproducibility bundle. Avoid overwriting the mutable checkpoint used by the runner.

# 7. Open questions and next validation

• Adaptive correction: measure whether repeated adverse outcomes eventually reduce the selected action's value in a richer sensory environment; log candidate values and tie-breaking decisions before each action.

• Cross-world mechanism: run controlled ablations (remove relational patterns, hold uncertainty constant, permute instrument order) and repeated seeds to isolate the reason for the first World B action.

• Persistence depth: compare snapshot hashes or canonical learned-state summaries before and after a verified reboot, including prediction tables, relationships and stable provenance IDs; demonstrate their subsequent use.

• Robotic validity: separate Gazebo/mock tests from real hardware sessions, document sensor inputs, commands, environment observations and physical safety controls.

• Reproducibility: record code commit, environment versions, repeated trials, null results and source scripts; invite external reproduction.

## 8. Interpretation for research review

NEL Mind is best described at present as an experimental cognitive architecture under active investigation. Its evidence demonstrates a set of interacting mechanisms rather than a finished general-purpose intelligence. The strongest controlled result is the trace from a recorded consequence to altered predictions, different endogenous value estimates, and a changed chosen action. The strongest persistence observation is checkpoint-based experience continuity across a verified Ubuntu host reboot. The most promising cross-environment result is a trained-vs-control first-action difference in a designed relational transfer task.

These findings justify further research and replication. They do not currently establish broad task competence, causal transfer of abstract knowledge across unrestricted environments, consistently effective self-correction, or full physical-robot autonomy. In particular, repeated failed actions and frequent value ties are measurable, unresolved limitations, not peripheral details.

## Evidence-status legend

| Term | Meaning in this report |
|---|---|
| PASS | The supplied console output reports a passing test/assertion; not independently replicated. |
| Observed | A value or state transition explicitly appears in the researcher’s supplied log. |
| Supported | A bounded interpretation tied to the described instrumentation and test design. |
| Unverified | A broader claim that would require additional logs, comparisons or tests. |

## Closing statement

The purpose of this dossier is to make the project’s present evidence inspectable, reproducible and falsifiable, while preserving the distinction between functioning mechanisms, controlled demonstrations, and open research questions.

Prepared: 10 November 2026
Researcher: Charles Dooner
Project: NeuralEdge Limitless (NEL Mind / Autonoema)