

Measured results. Visible evidence. Clear limits.

20 October 2026. This publication summarizes selected internal engineering results, preserved negative tests, and an early compatibility head-render preview. It does not establish external replication, generalized intelligence, clinical efficacy, or a released Evelyn assistant.

NewBrain: different tests, different conclusions

Study	Recorded result	What it actually establishes
Original12 after 256 updates	2/12 to 2/12	Two gained passes offset two losses; no net improvement.
Short conditional teaching	2/3 to 3/3 to 3/3	Three rehearsed answers correct after training and independent-process reload; generalization untested.
Longer complete answers	2/3 to 2/3	The first content token improved, but source/version tokens were omitted after an early end marker.
Five-question pilot	2/5 to 0/5 to 0/5	Regression after training, preserved after reload; no acquisition claim.
Numerical derivative check	80/80 and 4/4	Sampled coordinate and directional checks passed; not evidence of better learned answers.
One-step weight update	11,914 / 11,914	Saved parameter values matched one expected update; zero bit mismatches.

Interpretation: A correct sampled numerical check, a precisely matching one-step update, and a three-question taught-answer success are distinct findings. None overrides the unchanged Original12 benchmark or the separate unsuccessful longer and five-question tests.

EVELYN HART | ACTUAL CAPTURED RESULTS

Three visible head views.

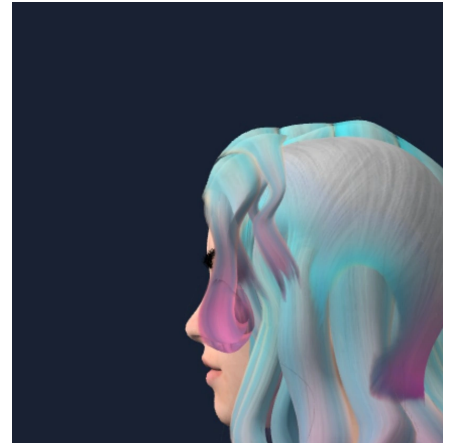
These are crops of screenshots from the recorded compatibility rendering test, not promotional concept art and not AI-generated substitute portraits. The images use an existing licensed head candidate with modified diffuse materials and reduced textures.



Front



Three-quarter



Side

What the review supports: Saved front, three-quarter, and side views were visibly nonblank, showed the face and eyes, differed by camera angle, and were reported as matching direct canvas captures. The earlier blank and graphics-failure records were preserved.

What remains unverified: Blinking, generated speech, mouth movement, lip synchronization, joined voice-and-avatar operation, visitor-memory restart, original PBR material fidelity, and a publicly installable Evelyn release.

Attribution: Base Female Character by BlackProject (CC BY 4.0). The compatibility presentation uses modified materials, reduced textures, and cropped renderer screenshots. License: <https://creativecommons.org/licenses/by/4.0/>

Evelyn also has a separate educational reference library prepared with 52 historical cards and 51 source entries. That reference collection is not proof of generated conversational accuracy or memory persistence.

For the original screenshot context, including the modified-material compatibility notice, see <https://kiralabs.org/evidence.html#evelyn-captures>.

What the reports actually cover.

NewBrain methods and limitations

Numerical-gradient test: a frozen two-probe method passed 80 coordinate and four direction comparisons, rejecting two deliberately wrong-gradient controls. All ten parameter groups met the predeclared signal requirement. Maximum absolute comparison error was about $2.08\text{e-}9$. No training or generated-answer assessment occurred in that diagnostic.

Single-step training check: a separately reviewed update matched 11,914 of 11,914 expected parameter values and rejected three incorrect arithmetic controls. This evaluates the selected numerical step, not whether a new unfamiliar fact was learned.

Later learning behavior: the verified short 3/3 restart involved three explicitly rehearsed synthetic examples; longer output stayed at 2/3 because the model ended after the first correct token rather than giving the required source and version. The separate five-question study fell from 2/5 to 0/5. Main Original12 stayed 2/12.

Patient Horizon and Avatar Builder

A proposed lung demonstration would connect versioned synthetic observations to an anatomical view and separately verified physiology. Recorded software checks include four synthetic request-planner checks and a body engineering result in which 512 weighted skin vertices moved through five commanded poses and returned to reference positions. They do not establish physiological lung simulation, clinical treatment prediction, or hospital integration.

Read the downloadable source reports

Archived original PDF - Measured Skin Movement and Restoration:

<https://kiralabs.org/reports/2026-10-measured-skin-movement-and-restoration.pdf>

Archived original PDF - How Our Tests Work and What the Results Mean:

<https://kiralabs.org/reports/2026-10-test-methods-and-results-explained.pdf>

NewBrain current experiment summaries: <https://kiralabs.org/newbrain.html#newbrain-longer-answer-2026-10-20>

Evelyn evidence gallery and original captures: <https://kiralabs.org/evidence.html#evelyn-captures>

Document scope: Kira Labs internal source reports and development screenshots; not peer-reviewed replication. NewBrain Original12 and Qwen comparisons are not represented as improved. Rendered compatibility head assets are not voice or conversation demonstrations.