

The Fleet: evidence guide

A short route from the report to the research behind it. Jonathan Simone.

The report contains four figures and a two-cache table. **SOURCE-NOTES.md** and **visual-evidence/README.md** map claims to evidence; **Factory in action** contains five research cases. The [public repository](#) carries the same evidence.

Follow a claim

Report reference	What to inspect	Where inside the evidence ZIP
[1] Fleet operation	Roles; stopped and repaired two-machine comparison	factory/cases/FACTORY-IN-ACTION.md
[2] Policy dissection	V13 disagreements; project modules; seven-transition replay probe	factory/cases/sources/04–06 files
[3] Labeling	Loss consumer; 46 drafts; eight-case enrichment and unchanged original weights	factory/cases/labeling-aggregates.json
[4] Representation	Two caches, twelve retained rows; capacity and imitation limits	factory/proofs/representation/README.md
[5] Learning	Hash-v1 closure, v2 interface, runnable decoder, historical teacher/student result	factory/proofs/README.md and SOURCE-NOTES.md
[6] Counter	Factorial, three combined comparisons, all seven matchups, c61 follow-up	numerical/METHODS.md
[7] Entry and deck	Credited submitted player, dated score, retention review and mechanics	source-evidence/receipts/ and SOURCE-NOTES.md
[8] Memory	2,333-artifact snapshot; audit recovery, reader repair	SOURCE-NOTES.md and factory/cases/sources/10–13 files

Three different checks

Game-result arithmetic: From the extracted ZIP folder, run `python numerical/verify_results.py --output verified-results.json`. Python 3.10+; standard library only. Checks hashes and cohorts, then recounts 28,000 development, 8,400 panel, 16,000 c61-follow-up and 1,203 student records.

Representation arithmetic: Run `python factory/proofs/representation/verify.py`. Standard library only. Checks twelve retained run rows across two caches, paired counts and metadata; recalculates both historical imitation comparisons.

Selection mechanics: Run `python factory/proofs/selection/selection_demo.py`. Requires PyTorch. Seven prescribed-score checks run on CPU, without a trained model, game engine, GPU or network.

These checks answer different questions. They do not replay historical matches, retrain the representation probe or establish tactical quality. Source manifests make the selected excerpts inspectable; they do not supply the private project's complete runtime.

Attribution and boundaries

Tetsutani's unmodified public c61 Grimmsnarl player is the Simulation entry, teacher and comparator. The experimental learner is a different artifact. The Lucario experiment modifies makthanithin's Apache-2.0 community_1084 policy. Roman Rozen's V13 policy supplied material for the disagreement study and the second recorded teacher cache. Public implementations remain credited foundations.

The factory operated through documented research cycles. Repeated retained improvement in an integrated learned player remains future work. The counter's approximately 19–20-point gains apply to two tested Alakazam implementations; broader benefit is unresolved. Representation agreement is not playing strength. Detailed assumptions, source identities, formulas and limits are in **SOURCE-NOTES.md**, **CONTRIBUTIONS.md**, **REPRODUCING.md** and **NOTICE.txt**.