LEAP Paper 5 C28 — Verification and Original Protocol

Meshaal Obaid Mouawad
21 November 2026

1 Evidence records

DOI1 — core replication materials: https://doi.org/10.7910/DVN/W2RJ51.
DOI2 — documentary condition register: https://doi.org/10.7910/DVN/4Q8BFP.
Original REF_A and REF_B were two role-conditioned sessions of the same coding model. The original assigned role descriptions were not verified personal credentials.

2 Original assessment form schema

Both original assessment CSVs have exactly fifteen column keys, in this order:

1. `case_id`
2. `reviewer_id`
3. `release_applicability`
4. `counted_entity`
5. `identity`
6. `time_window`
7. `eligibility`
8. `numerator`
9. `denominator`
10. `unit`
11. `allowed_output`
12. `source_spans`
13. `uncertainty`
14. `started_utc`
15. `finished_utc`

This is the exact header sequence from the frozen forms, not reconstructed field names.

3 Original case-identifier reconciliation

The following IDs are identical to the originals in REF_A/REF_B. Display aliases are editorial labels and are not keys in the archival CSV.

Original case ID	Display label	Pair
QUANTSTATS_01	QuantStats_01	Unresolved
QUANTSTATS_02	QuantStats_02	Unresolved
QUANTSTATS_03	QuantStats_03	Unresolved
QUANTSTATS_04	QuantStats_04	Unresolved
FFN_01	ffn_01	Unresolved
FFN_02	ffn_02	Unresolved
FFN_03	ffn_03	Unresolved
FFN_04	ffn_04	Point
RISKFOLIO_01	Riskfolio-Lib_01	Point
RISKFOLIO_02	Riskfolio-Lib_02	Point
RISKFOLIO_03	Riskfolio-Lib_03	Different
RISKFOLIO_04	Riskfolio-Lib_04	Point

Original case ID	Display label	Pair
B_pyportfolioopt_01	PyPortfolioOpt_01	Point
B_pyportfolioopt_02	PyPortfolioOpt_02	Point
B_pyportfolioopt_03	PyPortfolioOpt_03	Unresolved
B_pyportfolioopt_04	PyPortfolioOpt_04	Point
B_skfolio_01	skfolio_01	Unresolved
B_skfolio_02	skfolio_02	Point
B_skfolio_03	skfolio_03	Point
B_skfolio_04	skfolio_04	Point
B_vectorbt_01	vectorbt_01	Unresolved
B_vectorbt_02	vectorbt_02	Unresolved
B_vectorbt_03	vectorbt_03	Unresolved
B_vectorbt_04	vectorbt_04	Unresolved
empyrical_1	empyrical-reloaded_01	Unresolved
empyrical_2	empyrical-reloaded_02	Unresolved
empyrical_3	empyrical-reloaded_03	Unresolved
empyrical_4	empyrical-reloaded_04	Point
finquant_1	FinQuant_01	Unresolved
finquant_2	FinQuant_02	Unresolved
finquant_3	FinQuant_03	Unresolved
finquant_4	FinQuant_04	Unresolved

4 Citation-integrity verification

The frozen forms yield 128 case–path references to 127 distinct existing packet files; 212 numbered line-span occurrences; 180 ranges originally in bounds; 32 original REF_A `case_context.json` end points one to three lines past the file end. The 212-entry ledger provides the original start/end, source length, actual file checksum, validity flag and non-destructively corrected end for every occurrence. All documentary-text line-span references were valid. Eighteen additional file-only citations (across sixteen cases) point to existing files and do not contain line numbers. They are listed separately in the eighteen-entry file-level ledger; they are not part of the 212 numbered citations.

Source verification	Count
Paired case records	32
Numerically matching point-reference pairs	11
Both unresolved	20
Different resolution	1
Concordant native-input recalculations passing tolerance	11
Conditional numerical pairs verified	5
Original numbered citations	212
Numbered citations originally within file bounds	180
Documented context-file endpoint corrections	32
Separate existing file-only citations	18

5 Original instructions: REF_A

The source file is preserved byte-for-byte in the package; the PDF below renders its complete original wording for reading. Subject: Paper 5 — independent reference reconstruction, Specialist A

Please complete the attached Reference_Reviewer_A.zip independently. Your task is to reconstruct what each documented KPI is allowed to mean and return under its pinned release and supplied input settings. There are 32 cases. This phase establishes the reference; it does not ask you to judge whether the implementation has a defect.

Please do not discuss interpretations or results with the other specialist, view another review, inspect metric implementations, or seek native metric outputs before submitting your own completed review. Use the supplied documentary extracts, source-span receipts, case_context.json and raw input records. If a definition cannot be established from these materials, record the uncertainty explicitly. Send technical access questions only to the coordinator.

1. Extract the ZIP and keep the supplied case folders and input files unchanged. Open reference_review.csv and save your working copy as reference_A_completed.csv, in UTF-8 CSV format. Preserve the header, all 32 original case_id values and exactly one row per case. Use reviewer_id REF_A consistently.

2. For each case, read case_context.json first. It records the actual release, target, selected assets/portfolio, time/data selection and parameters. Then read the original documentary extracts and their span receipts. Some sources are

project-authored API docstrings; assess whether they establish a sufficient applicable definition. Do not presume a unique reference merely because a function is documented.

3. Fill every reference field:

- release_applicability: identify the supplied version/revision and why its definition applies, or explain unresolved applicability.

- counted_entity and identity: name what is counted/measured and the key identifying each observation, asset, date or portfolio.

- time_window and eligibility: identify the relevant rows/period, filtering, duplicate/missing-record treatment and documented prerequisites. State any unresolved convention.

- numerator and denominator: give the applicable components. For a non-ratio metric, write NOT_APPLICABLE with a short explanation rather than inventing a ratio.

- unit: distinguish fraction/percentage, money, per-period/annualized quantity and relevant output identity.

- allowed_output: provide the expected value, permissible values/range, or reproducible defining formula with numerical convention/tolerance and rationale. Include your numerical reconstruction when the documents and data support it. If no unique result is supportable, write UNRESOLVED and state the competing interpretations or missing information; do not guess a value.

- source_spans: cite exact supplied relative filenames and line/paragraph spans. Link calculations to your working notes/files by case_id.

- uncertainty: write NONE only when you find no material ambiguity; otherwise explain what prevents a unique reference.

- started_utc and finished_utc: record the actual start and finish for that case as ISO 8601 timestamps with Z or an explicit UTC offset. Do not invent or retrospectively estimate unrecorded times. Record interruptions in your notes.

4. You may calculate independently from the definitions and raw data using your usual calculation tools. Keep the calculations you actually used. Do not run the supplied metric implementation or use its outputs as the answer. Keep any working copies separate from the frozen inputs. There is no 20-minute limit for this reference-review phase.

5. Before returning your work, confirm that all 32 cases have one row, reviewer_id is REF_A, required fields are filled, timestamps are actual and timezone-qualified, every supported numerical result is reproducible from your workings, and every unresolved case is identified. An honest unresolved reference is a valid finding.

Please return exactly:

- reference_A_completed.csv — the completed 32-row form with the original column header.

- reference_A_notes.md — your reviewer ID, relevant expertise in a sentence, confirmation of independent completion, case-by-case explanations/calculation conventions, unresolved cases, interruptions and a list of files used.

- reference_A_workings.zip — any calculation scripts, worksheets or detailed workings you actually used. If the notes already contain all workings and no separate files were used, state this explicitly in the notes; the workings ZIP is then unnecessary.

Submit your original files directly to the coordinator. Keep a dated copy and do not reconcile your answers with the other specialist until both independent submissions have been frozen. A separate adjudicator will resolve disagreements later.

Thank you.

6 Original instructions: REF_B

Subject: Paper 5 — independent reference reconstruction, Specialist B

Please complete the attached Reference_Reviewer_B.zip independently. Your task is to reconstruct what each documented KPI is allowed to mean and return under its pinned release and supplied input settings. There are 32 cases. This phase establishes the reference; it does not ask you to judge whether the implementation has a defect.

Please do not discuss interpretations or results with the other specialist, view another review, inspect metric implementations, or seek native metric outputs before submitting your own completed review. Use the supplied documentary extracts, source-span receipts, case_context.json and raw input records. If a definition cannot be established from these materials, record the uncertainty explicitly. Send technical access questions only to the coordinator.

1. Extract the ZIP and keep the supplied case folders and input files unchanged. Open reference_review.csv and save your working copy as reference_B_completed.csv, in UTF-8 CSV format. Preserve the header, all 32 original case_id values and exactly one row per case. Use reviewer_id REF_B consistently.

2. For each case, read case_context.json first. It records the actual release, target, selected assets/portfolio, time/data selection and parameters. Then read the original documentary extracts and their span receipts. Some sources are project-authored API docstrings; assess whether they establish a sufficient applicable definition. Do not presume a unique reference merely because a function is documented.

3. Fill every reference field:

- release_applicability: identify the supplied version/revision and why its definition applies, or explain unresolved applicability.

- counted_entity and identity: name what is counted/measured and the key identifying each observation, asset, date or portfolio.

- time_window and eligibility: identify the relevant rows/period, filtering, duplicate/missing-record treatment and documented prerequisites. State any unresolved convention.

- numerator and denominator: give the applicable components. For a non-ratio metric, write NOT_APPLICABLE with a short explanation rather than inventing a ratio.

- unit: distinguish fraction/percentage, money, per-period/annualized quantity and relevant output identity.

- allowed_output: provide the expected value, permissible values/range, or reproducible defining formula with numerical convention/tolerance and rationale. Include your numerical reconstruction when the documents and data support it. If no unique result is supportable, write UNRESOLVED and state the competing interpretations or missing information; do not guess a value.

- source_spans: cite exact supplied relative filenames and line/paragraph spans. Link calculations to your working notes/files by case_id.

- uncertainty: write NONE only when you find no material ambiguity; otherwise explain what prevents a unique reference.

- started_utc and finished_utc: record the actual start and finish for that case as ISO 8601 timestamps with Z or an explicit UTC offset. Do not invent or retrospectively estimate unrecorded times. Record interruptions in your notes.

4. You may calculate independently from the definitions and raw data using your usual calculation tools. Keep the calculations you actually used. Do not run the supplied metric implementation or use its outputs as the answer. Keep any working copies separate from the frozen inputs. There is no 20-minute limit for this reference-review phase.

5. Before returning your work, confirm that all 32 cases have one row, reviewer_id is REF_B, required fields are filled, timestamps are actual and timezone-qualified, every supported numerical result is reproducible from your workings, and every unresolved case is identified. An honest unresolved reference is a valid finding.

Please return exactly:

- reference_B_completed.csv — the completed 32-row form with the original column header.

- reference_B_notes.md — your reviewer ID, relevant expertise in a sentence, confirmation of independent completion, case-by-case explanations/calculation conventions, unresolved cases, interruptions and a list of files used.

- reference_B_workings.zip — any calculation scripts, worksheets or detailed workings you actually used. If the notes already contain all workings and no separate files were used, state this explicitly in the notes; the workings ZIP is then unnecessary.

Submit your original files directly to the coordinator. Keep a dated copy and do not reconcile your answers with the other specialist until both independent submissions have been frozen. A separate adjudicator will resolve disagreements later.

Thank you.

7 Reproduction file map

- `supplements/C28_32_Case_Decision_Index.csv`: original-ID keyed complete case index.
- `supplements/C28_212_Original_Source_Span_Validation_and_Corrections.csv`: all numbered citations with original and corrected line endpoints.
- `supplements/C28_18_File_Level_Citations.csv`: all unnumbered original file citations.
- `supplements/C28_32_Case_Recorded_Conditions.csv`: preserved DOI2 case conditions.
- `supplements/C28_21_Nonconcordant_Case_Topic_Summary.csv`: source-linked nonconcordant synthesis.
- `protocol/REF_A_Original_Instructions.md` and REF_B equivalent: original assignments.
- `C28_Original_Assessment_Evidence_Catalog.pdf`: full A/B outputs, uncertainty notes and source strings.

8 Scope of the resulting record

The recorded outcomes describe two separately executed sessions of one coding model under different role-conditioned instructions. The twenty jointly unresolved cases and one different-resolution case are the observed submissions, not an independently adjudicated score of correct abstention. The full original evidence remains preserved in DOI1 and DOI2.