Response to the Third-Party Verification of Package Revision 5

Site: followtheepicenter.com **Analyst:** Jon Bray **Date:** 10 Nov 2026 **Supersedes:** package rev. 5 → rev. 6 (`forensic_acoustic_package_r6.zip`)

The verification of revision 5 reproduced the source hashes, the 20,000-draw Monte Carlo, the geometry and the 16,384-assignment sensitivity results, and raised two remaining issues. Both are accepted; checking the first one turned up a third problem of my own. The corrections are logged as C32–C34 in `docs/results_v7.md`, Appendix A. The corrected verification report ("Independent R5 Recheck — pre-send precision review") was received while this response was being prepared; the cross-check against its calibration table is in §2.

1. The impulsiveness criterion — accepted, and it was worse than stated

Both halves of the criticism are right. The criterion written into revision 5 (CHECKLIST G6, results §3) was never implemented in `first_onset()`, which still accepted the first k=5 crossing inside a hand-chosen window. And the criterion as worded does not work. Under a stated definition — 2–15 kHz band, sliding 1-ms RMS envelope at sample resolution, half-maximum width of the strongest peak within 25 ms of the onset — the measurements are:

event	half-maximum width (ms)
Cam 6 pick	1.54
Cam 11 pick	**5.78** (auditor: 5.78 by one method, 4.58 by another)
2.MOV pick	6.26
IMG_6368 pick	3.63
Cam 8 pick	2.18
Cam 7 pick	1.29
Canon pick	1.33
Cam 11 precursor at 1.561 s	**3.0**
Cam 6 precursor at 6.984 s	10.8

Two of the seven picks fail the "< 5 ms" test I wrote, and one of the sustained-level crossings I said the test would reject passes it. The "41–47 ms" I quoted for the precursors was a different measure (the span above half-maximum over a 50-ms block envelope), which I did not state and which I should not have compared with a peak width. The Cam 11 discrepancy the auditor asked about (their 4.58 vs 5.78, my 9 ms) is therefore not a discrepancy to resolve: my figure was wrong, and their two values differ only by the envelope definition (block vs sliding RMS), which is exactly why the definition has to be fixed before the number means anything.

What *does* separate the picks from the precursors, by a wide margin, is amplitude contrast — peak of the 1-ms RMS envelope within 25 ms of the crossing, divided by the median 1-ms RMS of the preceding 150 ms:

	ratio
picks (seven recordings)	21 – 236
sustained-level crossings in the widened windows (fourteen, all six phones)	3.9 – 9.2

Revision 6 implements this in `first_onset()` with a threshold of 15 (a factor ≥ 1.4 below the weakest pick, Cam 8 at 21, and ≥ 1.6 above the strongest rejected crossing, Cam 7 at 9.2), widens every search window by 300 ms so the criterion rather than the window rejects the precursors, and prints the rejected crossings next to each pick in Table R1 (C32). Half-maximum width is not used anywhere.

A third correction, found while doing this (C33). The half-millisecond scan grid in `first_onset()` started at the window's first sample, so the pick depended on where the window started — a sub-block phase effect. Widening the windows exposed it: Cam 6 moved +1.11 ms, the others ≤ 0.1 ms. The grid is now anchored to absolute sample multiples and the pick is verified identical from three different window starts. Because Cam 6 is the reference, every true-time difference in the package is 1.0–1.2 ms smaller than in revisions 5–6. Every table has been regenerated; no conclusion changes, but the onset table the auditor reproduced "to the sample" from revision 4 will now differ at Cam 6 (7.08880 → 7.08991 s), and that is the reason.

2. P-values for χ^2 minimised over speaker assignments — accepted

Agreed, including the auditor's correction of their own earlier reading. Under the null (pin true, *some* assignment true), the minimum over 256 assignments is stochastically smaller than $\chi^2(3)$, so sf(χ^2_min, 3) is an upper bound on the tail probability — conservative, not aggressive. Revision 6 adds `scripts/pin_calibration.py`, which simulates Gaussian timing errors (σ as stated) under each of the 256 candidate true assignments, repeats the minimisation on every draw, and reports the calibrated p averaged over a uniform prior on the true assignment together with its range:

pin (v = 830 m/s)	χ^2_min	nominal p (upper bound)	calibrated p	range over true assignment
"shooters position"	2.17	0.54	**0.115**	0.016 – 0.26
"n wave source"	4.53	0.21	**0.028**	0.001 – 0.09
prosecution pin	10.08	0.018	**0.001**	≤ 0.006

Cross-check against the auditor's own calibration. Their table (30,000 draws, uniform speaker prior, revision 5 picks) gives 0.0872 / 0.0120 / 0.000367 for the three pins at 830 m/s. Running `pin_calibration.py` on the revision 5 arrivals (`--arrivals output/results_arrivals_r5.json`) gives **0.087** / **0.012** / < **0.001** — the two implementations agree to the stated precision. The revision 7 values above differ only because the picks moved by 1 ms (C33), which raises the "shooters position" χ^2_min from 2.61 to 2.17.

(780 and 880 m/s in the output file.) So the revision 6 statement that the "shooters position" pin was "comfortably compatible" (p ≈ 0.46) overstated it: calibrated, it is p ≈ 0.08–0.17 — not excluded at 5%, but not comfortable, and below 5% for some true assignments. The "n wave source" pin moves to the 5% boundary. The prosecution pin's exclusion strengthens (calibrated p ≤ 0.003), which is the direction the auditor's nuance predicts. All of it remains conditional on Gaussian independent timing errors, pinned camera positions and c, none of which is independently established — the same caveat as before, now attached to numbers that are calibrated rather than bounded.

Other points in the report

- *Derived-WAV hash mismatches across FFmpeg builds.* Agreed that they do not indicate altered source audio. For the record: `extract.sh` with the `pan=` filter on FFmpeg 6.1 reproduces every derived hash in the manifest bit for bit; a build that writes a different WAV header (LIST/INFO chunk, `WAVE_FORMAT_EXTENSIBLE`) will change the file hash without changing a sample. A reviewer who sees a

mismatch should compare sample data, e.g. `ffmpeg -i x.wav -f s16le - | sha256sum`; the next revision will carry sample-data hashes alongside file hashes.

- *Canon WAVs and geometry.* The two Canon channels as analysed are now at `followtheepicenter.com/archive/canon/` with their manifest hashes. Surveyed camera positions and PA speaker identities still do not exist; they remain the first items in Part V.
- *"We do not reject an otherwise correct location hypothesis merely because its model is incomplete."* Agreed, and the package's wording has been brought to the same standard: the timing neither selects nor rejects the "shooters position" pin; it rejects the prosecution pin only under an idealised error model.

What changed in the package

`first_onset()` (criterion + grid anchoring), `SEARCH` windows (−300 ms), Table R1 (ratios and rejected crossings), every downstream table (−1.0 to −1.2 ms), `pin_calibration.py` and its output, results §3, §7, §8a, §8 item 4, Appendix A C32–C34, CHECKLIST items H1–H3 (G6 withdrawn). `output/null_big.log` is still the revision 5 run: the association it tests is unchanged (Cam 6 candidate 344 → 343 ms, same event, same 1.9 ms rms) and the hour-long re-run is not justified by a 1 ms reference shift, but the file is labelled accordingly.

Both of the auditor's points were right, and the second was right in the refined form they themselves supplied. If the corrected report shows anything in the above that still does not reproduce, I would like to see it.