The surrogate band and its three readings

One sentence of methods prose — "the 95% confidence interval calculated from the `h_max − h_min` of the surrogates" — is three different tests on identical surrogate ensembles; the standard-error reading tightens the band by `√n_surr` and turns a nominal 5% test into one that fires on a third to a half of calibrated monofractal nulls

Evan Tabak Atlas (ORCID 0009-0007-7374-2338)

Draft — methods note, in the allied register. It prices one decision convention, not a research programme and not anybody's conduct: the lineage engaged here introduced the IAAFT surrogate test that this note treats as the correct null, published the prior art that anticipates its own low jaw, and its standardised instrument keeps full power under every reading measured below. Every quantitative claim traces to a named, immutable results record, cited by DOI and never by repository path: record v57 (the cognition card — the three readings on 987 public lexical-decision sessions and on ten calibrated null families at three lengths), 10.5281/zenodo.22759701; record v60 (the birdsong card — the same finding on a different lineage member's own published recording), 10.5281/zenodo.22759707; record v38 (the t_MF robustness battery — the convention's first measured false-positive rates), 10.5281/zenodo.22759652. The companion methods paper is "The width and the null" (papers/08), concept DOI 10.5281/zenodo.22180476, whose §6 quotes the published `t_MF` formula that this note takes as its starting point. Target venue: a short methods note or comment, alongside papers/08. Owner actions, not this note's: posting, and whether to open the joint pre-registration §7 offers. This note's own deposit: Zenodo `10.5281/zenodo.22760449` (preprint; DOI reserved 2026-09-15; published 2026-09-15 on the owner's sign-off).

Abstract

The multifractality test this literature has run since 2010 compares a series' singularity-spectrum width to the widths of its IAAFT surrogates and rejects when the original lies outside *"the 95% confidence interval"* of the surrogate widths. That phrase is not one test. On one and the same ensemble it admits at least three: a **band** (the surrogate mean plus `1.96` SD of the surrogate widths), a **standard error of the surrogate mean** (the mean plus `t × SD/√n_surr` — the construction the successor instrument `t_MF` standardised), and a **rank** or exchangeability test (the original must out-rank its surrogates at `p < 0.05`). The readings differ by a factor of `√n_surr` — about **5.5** at the **30** surrogates the 2010 precedent fixed — and the difference is not cosmetic. On 987 public lexical-decision session series, the reconstructed 2010 rule flags **51/987** under the band reading and **63/987** under the rank reading — the nominal rate — while the *same* thirty surrogates read through the surrogate mean's standard error flag **478/987**, and the successor instrument's published rule flags **495/987** significant-positive plus **252** significant-negative. On identical ensembles at the same length, the standard-error reading fires significant-positive on **8/24** realisations of a fitted fGn, **8/24** of a fitted AR(4), **8/24** of a Gaussian monofractal carrying the cohort's own response-time marginal and **10/24** of one carrying a single lapse episode, where the band readings of the same ensembles fire 0–2/24 and the rank readings 0–1/24 — at full power on the canonical cascade. The same behaviour replicates on a different lineage member's own published recording. The convention, not the statistic, is the cost: the sealed flag says so, and the fix is one line of a methods section. Report the band, the standard error and the rank side by side, both tails, and state `n_surr`.

---

1. One sentence, three tests

The surrogate test that carries most multifractality claims in the behavioural, physiological and bioacoustic literatures entered it with a single paper (Ihlen & Vereijken 2010) and a single sentence of instruction. The sentence, as it is quoted by the documented instantiation of that framework on the first author's own toolbox (Kuznetsov & Wallot 2011), reads:

> "Multiplicative interaction … is present when the observed multifractal spectrum (`h_max − h_min`) is greater than the 95% confidence interval calculated from the `h_max − h_min` of the surrogates."

with **30** IAAFT surrogates, *"following the precedent of Ihlen and Vereijken, 2010"* (Bell et al. 2019). Read as English this is unambiguous enough to implement and ambiguous enough to implement three ways, because *"the 95% confidence interval"* of a sample of thirty numbers is not one object:

1. **SD — a band of the surrogate width *distribution*.** Reject when the original exceeds `mean + 1.96 × SD` of the thirty surrogate widths. This asks: *is the original outside the range the surrogates themselves occupy?*
2. **SE — a confidence interval for the surrogate *mean*.** Reject when the original exceeds `mean + t₀.₉₇₅,₂₉ × SD/√30`. This asks: *is the original at the surrogate mean?*
3. **RANK — exchangeability.** Reject when `(1 + #{surr ≥ obs})/(n_surr + 1) < 0.05`. With thirty surrogates the original must exceed every one of them (`1/31 ≈ 0.032`). This asks: *is the original a plausible draw from its own surrogate ensemble?*

Reading 2 is not a straw man, and this note does not attribute it to the 2010 paper: it is the construction the successor line standardised and printed. The one-sample statistic `t_MF` is given, in the lineage's own words (Bell et al. 2019, from Kelty-Stephen et al. 2013, and quoted in papers/08 §6), as

```
t_MF = (W_MF of original series − Average W_MF of surrogate-data series) /
(Standard error of W_MF of surrogate-data series)
```

and the later methodological statements say the denominator in so many words — *"dividing by the standard error of Δα for the surrogates"* (Mangalam et al. 2023), with `t_MF > 1.98` read as evidence of an intermittent cascade. The same lineage elsewhere writes the statistic with an SD denominator, `t_MF = (Δα_Orig − ⟨Δα_Surr⟩)/σ_{Δα,Surr}` (Mangalam, Likens & Kelty-Stephen 2025), which is reading 1. **Both denominators are in print, from the same programme, for the same statistic.** That is the whole of this note's subject: not a mistake, but an unpinned convention that no reader of a results section can recover, and that changes the answer by a factor of `√n_surr`.

2. The arithmetic of the tightened band

The three readings share a numerator, `W_orig − mean(W_surr)`, and differ only in what they divide it by. Reading 1 divides by `SD(W_surr)`; reading 2 divides by `SD(W_surr)/√n_surr`. The bar is therefore tighter by exactly `√n_surr` — **5.5** at `n_surr = 30`, and 5.7 at the **32** surrogates the successor instrument fixes.

Written as the offset an original needs in order to fire, the band reading demands that the original sit `1.96` SD above the ensemble; the standard-error reading demands only `2.045/√30 ≈ 0.37` SD. Any *systematic* offset of the estimator — a finite-size bias, a marginal effect the surrogates do not reproduce exactly, an estimator's own floor — of roughly four tenths of a surrogate SD is enough. Nothing in that sentence is about multiplicative cascades.

Reading 2 also has no nominal level to speak of. The rank reading's level is exact by construction (it is a permutation-style p-value on an exchangeable ensemble); the band reading's `1.96` is a Gaussian approximation to a 5% two-tailed rate on the width distribution and is close enough in practice. The standard-error reading tests a hypothesis nobody stated — *the original equals the surrogate ensemble's mean* — and as the ensemble grows the test gets *stricter*, not better calibrated. Record v38 stated this in its fourth headline: a one-sample `t` against the surrogate mean's standard error tests "is the original exactly at the surrogate mean", and with 32 surrogates any systematic offset beyond about 0.36 SD fires it, in either direction. What follows is that statement, measured on real cohorts and on calibrated nulls.

3. The three readings on 987 lexical-decision sessions

Record v57 runs the reconstructed 2010 pipeline — an analytic Morlet CWT (`ω₀ = 6`, L1), scales `2..N/16` at quarter-octave spacing, the structure function `⟨|W_a|^q⟩` over `q = 0.1..3`, the Legendre width `h_max − h_min`, against **30** IAAFT surrogates — on 987 trial-order lexical-decision session series (median `N = 830`) from 506 subjects of the Semantic Priming Project, the named public equivalent of the 2010 paper's non-public response series. The three readings are computed on *the same* thirty surrogates per series. `t_MF` is computed beside them, on its own 32-surrogate Chhabra–Jensen ensembles, and read the same three ways.

reading of "the 95% CI"	2010 rule	`t_MF`
SD (band)	**51/987** (5.2%)	93/987 (9.4%)
RANK (exchangeability)	**63/987** (6.4%)	69/987 (7.0%)
SE (surrogate-mean standard error)	**478/987** (48%)	**495/987** (50%) positive, **252** (26%) negative

Under exchangeability the cohort sits at the nominal rate: **51/987** under the band reading (exact-CP `[0.039, 0.067]`) and **63/987** under the rank reading, with the deficit tails at 25 and 21 series, and with the IAAFT ensembles handing back essentially the whole width (cohort mean 0.170 against a surrogate mean of 0.158). Under the standard error the same thirty surrogates give **478/987**, and the successor instrument's published rule gives **495/987** significant-positive plus **252** significant-negative. The published claim these numbers were run against — *"the major portion of these response series had multiplicative interactions between temporal scales"* — is reached by the standard-error convention and by nothing else in the panel. The cohort's own trimming (the project's `keep` flag, 200–3,000 ms) moves nothing: 53 / 471 / 69.

The point is not which of these cohorts is "right". It is that a reader given only a results section cannot tell which test produced them, and the spread between the readings is an order of magnitude on the same data with the same surrogates.

4. The price, on calibrated nulls

The cohort numbers alone do not say which reading is miscalibrated; the nulls do. Record v57's P3 stage builds ten calibrated families from the cohort itself — a fitted fractional Gaussian noise, the reference session's own IAAFT and phase-randomised twins, a fitted Yule-Walker AR(4), a Gaussian monofractal rank-remapped onto the reference session's exact response-time marginal (excess kurtosis 8.5 — the pitfall the framework's first author later named), an iid Student-`t`, a monofractal carrying one lapse episode, one carrying an isolated cusp, one carrying the cohort's measured break-delimited block shifts and practice trend, and one carrying a coherent sine at the measured low-frequency variance fraction — and runs 24 seeds of each through every instrument and every reading at `N = 830`, 1,660 and 4,096.

At the cohort's own length, the published standard-error rule fires significant-positive on **8/24** realisations of the fitted fGn, **8/24** of the fitted AR(4), **8/24** of the monofractal carrying the cohort's own response-time marginal and **10/24** of the monofractal carrying one lapse episode; the response-time-marginal cell rises to **11/24** at `N = 4,096`. Read as bands, the same ensembles fire 0–2/24 across both instruments (the 2010 rule's own SD reading and `t_MF`'s); read as ranks, 0–1/24. The standard-error reading is also two-sided in practice: at `N = 830` it reads 7 to 14 of 24 realisations of the nine non-cusp families significant-*negative*, and the isolated-cusp family — a single non-scaling event in a monofractal — reads **24/24** significant-negative, the estimator's high side arriving with the opposite sign.

None of these families contains a multiplicative cascade. A test that fires on a third to a half of them is not a 5% test, and it is not a test whose positives license the inference the literature draws from them. The record's sealed flag for this cell, `convention_is_the_cost`, is **true**: the same statistic on the same ensembles read as ranks is specific at every length.

5. What switching costs: power, stated honestly

The reason to state the cost is that switching is not free everywhere, and the allied register requires saying where.

On the canonical positive control the three readings are nearly indistinguishable. The binomial multiplicative cascade at `N = 830` is detected by `t_MF` under the standard error **24/24**, under the band **22/24** and under the rank reading **20/24**; the rank reading reaches **22/24** at `N = 1,660` and **24/24** at `N = 4,096`. On this cell — the field's own canonical cascade, and the one on which the MFDFA width has near-zero power (papers/08 §5, record v34) — switching from the standard error to exchangeability costs four of twenty-four realisations at the shortest length and nothing at the longest.

That statement applies to the binomial cell only. On the weaker cascade families the rank reading loses real power. At `N = 830` the stochastic-multiplier twin matched to the binomial's own `σ` is detected by the standard-error reading **16/24** and by the rank reading **8/24**; the latent msm-Cox log-cascade is detected **11/24** and **2/24**. At 1,660 and 4,096 the same pattern holds (σ-twin 19/24 against 8/24, and 20/24 against 8/24; msm-Cox 13/24 against 0/24, and 16/24 against 2/24). Half the standard-error reading's apparent power on weak cascades is the same anti-conservatism that produces its false positives, and cannot be separated from it by any measurement in these records — but the other half is not nothing, and a study whose cascades are weak and whose series are short will lose detections by switching. The honest summary is that the rank reading buys calibration at a price that is zero on the strong deterministic control and real on the weak stochastic ones, and that the price should be paid openly rather than avoided silently.

Record v38 measured the same convention from the other side and in the counterparty's favour, which is why it is cited here: `t_MF` detects the marginal-encoded deterministic binomial **48/48** — it does *not* share the MFDFA width's blind spot — while its decision convention reads an iid heavy-tailed monofractal significant on **35/48** realisations and an isolated-singularity monofractal on **45/48**, all but one of those significant-*negative*. The power is the statistic's; the false positives are the denominator's.

6. The replication on the counterparty's own data

Record v60 audits a different member of the same lineage on its own published recording: a thrush-nightingale song analysed as a 1000-Hz dB amplitude envelope, with the published rule read as the SD band (the figure's own units are "standard deviations from the mean of the IAAFTs") against 100 IAAFT surrogates. The published positive reproduces — **21/23** songs beyond their band — and the record says so first.

Two facts about the null itself carry the convention finding to a second dataset, a second estimator, a second surrogate budget and a second species. First, on that pipeline's marginal — an envelope that is a **59%**-zero series, because the pauses between notes are set to zero — the IAAFT null is only partially converged, and it is not self-consistent: an IAAFT surrogate of the reference song, read against its *own* IAAFT ensemble through the paper's rule, passes **7/24** times. A null that reads its own draws as positive **29%** of the time is not a 5% test, whatever reading is layered on top of it. Second, the standard-error reading of those same ensembles fires on **11–24 of 24** draws of every calibrated null family the record builds — from the fitted fGn, the AR(16) knee, iid heavy tails, a burst, a quasi-periodic line, a coherent sine and the reference's own IAAFT and phase-randomised twins through to the isolated cusp. The SD reading of the very same draws fires 0–7/24 on the eight families that are neither the cusp nor the note-shuffled song — the two that the record reads 24/24 under every reading, and that are its actual finding about the songs.

The two records share no data, no estimator, no surrogate count and no domain. They agree on the convention.

7. The one-line fix, and a joint pre-registration

The fix does not require a new statistic, a new null, or a re-analysis of anything. It requires one line of a methods section and one row of a table:

> **Report the band, the standard-error and the rank readings of the same surrogate ensemble, side by side, both tails, and state `n_surr`.**

Three numbers where there was one. They come from the ensemble that has already been computed, at no additional surrogate cost, and they answer three different questions that the phrase *"outside the 95% confidence interval of the surrogates"* has been carrying at once. Where they agree, the verdict is robust to the convention and the paper is stronger for having shown it. Where they disagree — as they do by an order of magnitude on 987 real cognitive series — the disagreement *is* the result, and the reader is entitled to it. State `n_surr` because the standard-error reading's bar is a function of it: two studies running the identical pipeline at 30 and at 100 surrogates are running tests that differ by a factor of 1.8 in strictness, with nothing in either methods section to say so.

This note makes no recommendation about which reading should be primary, and it cannot: that choice depends on the estimator's systematic bias at the study's own `N`, which is measurable per pipeline and is not measured here for any pipeline but the two in records v57 and v60.

A joint pre-registration, offered in the allied register. papers/08 §6 already offers an adversarial collaboration on cascade constructions, lengths, surrogate budgets and the decision rule, and its §10 reporting standard is where the line above belongs; this note adds one item to the offer, narrower and more immediately decidable than any of them: **fix the reading of the interval, on data the lineage owns.** The successor line has public datasets whose published per-series numbers reproduce end-to-end (record v57 reproduces 166 published block-level widths at `r = 0.885`, slope 1.01, with the condition summaries inside the published SDs) — so the denominator question can be settled on the lineage's own series, with its own estimator, at its own `n_surr`, by running the three readings in parallel and reporting all three. That is a one-afternoon protocol whose outcome binds both sides. Whether to open it is an owner decision; this note only offers it.

8. Scope and limitations

1. **The 2010 pipeline here is a reconstruction.** The 2010 paper is closed-access and its full text was unreachable in the executing session; its estimator, scale range, `q` range, surrogate count and decision rule are those the documented instantiations state for the framework, with locators sealed in record v57's pre-registration. Whether the 2010 runs themselves used the band, the standard error or a rank is not recoverable from here — which is the note's point, not a rebuttal of it. All three readings travel with every number in the record.
2. **The cohort is an equivalent, not the originals.** The 2010 response series are not public; the Semantic Priming Project's lexical-decision sessions are a two-choice decision task with an interleaved priming design and a 3,000-ms deadline, and every cohort number above is scoped to them.
3. **The nulls are 24-seed panels.** Every calibrated rate above carries an exact Clopper–Pearson interval in the record; a rate of 8/24 is `[0.156, 0.553]`. The claim "a third to a half" is a claim about a set of families, not about any one cell's point estimate.
4. **Nothing here audits an empirical finding.** No published result about cognition, posture or birdsong is overturned by this note. What is priced is a decision convention; what any particular paper's verdict becomes under the other two readings is a question for that paper's own data.
5. **The counterparty's power result stands.** `t_MF` sees the canonical cascade the MFDFA width cannot (48/48, record v38; 24/24 at the cognitive length, record v57). This note argues about the denominator, not the estimator, and not the null — which is the right one (record v57's family 5: a Gaussian monofractal carrying the cohort's own heavy-tailed marginal reads 0/24 under the 2010 rule's band reading at every length, and at most 1/24 under its rank reading).
6. **Register.** Allied, at maximum strength. No personal material about any author appears here or in the records cited; the object under audit is a pipeline convention.

9. References

- **Bell, C. A., Carver, N. S., Zbaracki, J. A., & Kelty-Stephen, D. G. 2019**, "Non-linear Amplification of Variability Through Interaction Across Scales Supports Greater Accuracy in Manual Aiming: Evidence From a Multifractal Analysis With Comparisons to Linear Surrogates in the Fitts Task," *Frontiers in*

Physiology, 10, 998. doi:10.3389/fphys.2019.00998 *(the explicit* `t_MF` *formula, with the standard-error denominator in words; and the "30, following the precedent of Ihlen and Vereijken, 2010" surrogate count).*

- **Hutchison, K. A., Balota, D. A., Neely, J. H., et al. 2013**, "The Semantic Priming Project," *Behavior Research Methods*, 45(4), 1099–1114. doi:10.3758/s13428-012-0304-z *(§3's 987 lexical-decision session series).*
- **Ihlen, E. A. F. 2014**, "Multifractal analyses of human response time: potential pitfalls in the interpretation of results," *Frontiers in Human Neuroscience*, 8, 523. doi:10.3389/fnhum.2014.00523 *(the framework's first author on the response-time marginal, and on 1,000 trials possibly being too few — §4's family 5 and the P2 stage of record v57).*
- **Ihlen, E. A. F., & Vereijken, B. 2010**, "Interaction-dominant dynamics in human cognition: Beyond 1/f^α fluctuation," *Journal of Experimental Psychology: General*, 139(3), 436–463. doi:10.1037/a0019098 *(the paper that brought the IAAFT surrogate test into this literature; the sentence §1 reads three ways is its decision rule as documented by the instantiations below).*
- **Kelty-Stephen, D. G. 2018**, "Multifractal evidence of nonlinear interactions stabilizing posture for phasmids in windy conditions: A reanalysis of insect postural-sway data," *PLOS ONE*, 13(8), e0202367. doi:10.1371/journal.pone.0202367 *(the public dataset whose published block-level widths record v57 reproduces at r = 0.885 — the data §7's joint pre-registration would run on).*
- **Kelty-Stephen, D. G., Lane, E., Bloomfield, L., & Mangalam, M. 2022/2023**, "Multifractal test for nonlinearity of interactions across scales in time series," *Behavior Research Methods*, 55(5), 2249–2282. doi:10.3758/s13428-022-01866-9 *(the standardised* `t_MF` *instrument).*
- **Kuznetsov, N. A., & Wallot, S. 2011**, "Effects of Accuracy Feedback on Fractal Characteristics of Time Estimation," *Frontiers in Integrative Neuroscience*, 5, 62. doi:10.3389/fnint.2011.00062 *(the documented instantiation of the 2010 framework on the first author's own toolbox; the source of §1's quoted decision sentence and of the 30-surrogate count).*
- **Lee, H., & Kelty-Stephen, D. G. 2017**, "Cascade-Driven Series with Narrower Multifractal Spectra Than Their Surrogates: Standard Deviation of Multipliers Changes Interactions across Scales," *Complexity*, 2017, 7015243. doi:10.1155/2017/7015243 *(prior art from inside the lineage: a genuine cascade can read narrower than its surrogates — the deficit direction that makes "both tails" a requirement rather than a courtesy).*
- **Mangalam, M., Kelty-Stephen, D. G., Sommerfeld, J. H., Stergiou, N., & Likens, A. D. 2023**, "Temporal organization of stride-to-stride variations contradicts predictive models for sensorimotor control of footfalls during walking," *PLOS ONE*, 18(8), e0290324. doi:10.1371/journal.pone.0290324 *(the lineage's own wording of the decision rule quoted in §1: "dividing by the standard error of Δα for the surrogates";* `t_MF > 1.98` *).*
- **Mangalam, M., Likens, A. D., & Kelty-Stephen, D. G. 2025**, "Multifractal nonlinearity as a robust estimator of multiplicative cascade dynamics," *Physical Review E*, 111(3), 034126. doi:10.1103/PhysRevE.111.034126 *(the robustness line; the same statistic written with an SD denominator,* `σ_{Δα,Surr}` *— §1's second printed convention).*
- **Oświęcimka, P., Drożdż, S., Frasca, M., et al. 2020**, "Wavelet-based discrimination of isolated singularities masquerading as multifractals in detrended fluctuation analyses," *Nonlinear Dynamics*, 100(2), 1689–1704. doi:10.1007/s11071-020-05581-y *(the published high side that §4's cusp family instantiates).*
- **Roeske, T. C., Kelty-Stephen, D., & Wallot, S. 2018**, "Multifractal analysis reveals music-like dynamic structure in songbird rhythms," *Scientific Reports*, 8, 4570. doi:10.1038/s41598-018-22933-2 *(§6's replication target, audited on its own published recording as record v60).*

- **Schreiber, T., & Schmitz, A. 1996**, "Improved Surrogate Data for Nonlinearity Tests," *Physical Review Letters*, 77(4), 635–638. doi:10.1103/PhysRevLett.77.635 *(IAAFT: the null every reading above shares).*
- **Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. 1992**, "Testing for nonlinearity in time series: the method of surrogate data," *Physica D*, 58(1–4), 77–94. doi:10.1016/0167-2789(92)90102-S *(the founding statement of the rank/exchangeability reading — reading 3 is not this note's invention but the method's original form).*

10. Reproducibility

Everything above is recomputed from committed record JSON by the repository's own gate. `fenocosm.verify.check_paper_anchors` binds every bolded number in this note to a key of `out/v57/p57_cognition.json`, `out/v60/p60_birdsong.json` or `out/v38/p38_tmf.json` and asserts both that the record still renders that value and that this text still quotes it verbatim; a drift in either direction fails `./verify.sh`. Appendix A is the human-readable form of that registry.

The records themselves regenerate deterministically from their own drivers, with pinned seeds and pinned input checksums, behind the same gate:

```
python3 -m venv venv && venv/bin/pip install -r requirements-dev.txt
./verify.sh                                          # apparatus gate, data-free
venv/bin/python -m fenocosm.p38_tmf                  # record v38  (~5 min)
venv/bin/python -m fenocosm.p57_cognition            # record v57  (~112 min, 4 workers)
venv/bin/python -m fenocosm.p60_birdsong --workers 4   # record v60  (~31 min)
python tools/build_pdf.py                            # -> papers/pdf/*.pdf
```

Record v57's only external inputs are the Semantic Priming Project workbook and the phasmid Dryad files, both pinned by SHA-256 and refused by the loader if the bytes differ; record v60's is the xeno-canto recording XC75409, fetched at run time with its SHA-256, decoder and page licence sealed in the JSON's `provenance` block (the audio itself is never redistributed); record v38 needs no data at all. Everything else in all three is synthetic and seeded. The readings compared in §§3–6 are computed on *identical* surrogate ensembles within each record — the same thirty (or thirty-two, or one hundred) draws, read three ways — which is what makes the comparison a statement about the convention rather than about sampling.

11. Data availability

Each record is deposited as its own citable object and is cited by DOI; the source repository is private and remains private.

record	what it carries here	DOI
`out/v57/`	§§3–5: the three readings on 987 lexical-decision sessions, on ten calibrated null families at three lengths, and on the three cascade positive controls; the sealed `convention_is_the_cost` flag	10.5281/zenodo.22759701
`out/v60/`	§6: the replication on a different lineage member's own published recording — the self-consistency rate of the null and the standard-error reading's rates on every calibrated family	10.5281/zenodo.22759707
`out/v38/`	§5: the convention's first measured false-positive rates, and the `t_MF` power result that stands beside them	10.5281/zenodo.22759652

The companion methods paper, “The width and the null” (papers/08), whose §6 quotes the published `t_MF` formula this note starts from, is deposited at concept DOI 10.5281/zenodo.22180476.

This note is itself deposited as a standalone Zenodo preprint at 10.5281/zenodo.22760449 (Markdown and PDF; DOI reserved 2026-09-15; published 2026-09-15 on the owner’s sign-off), citing the three records and the companion paper by DOI in its metadata.

External data used by the records: the Semantic Priming Project’s trial-level lexical-decision data (Hutchison et al. 2013); the Dryad dataset of Kelty-Stephen 2018 (CC0; original recordings Bian, Elgar & Peters 2016); and xeno-canto recording XC75409 (*Luscinia luscinia*; recordist Tomas Belka), whose own licence, as stated on its page, governs any reuse.

Appendix A — Provenance table (every headline number → its sealed

record key)

number (as used)	record	JSON key / path
√n_surr tightening factor **5.5** at n_surr = 30	v57 (10.5281/zenodo.22759701)	p57_cognition.json → protocol.iv_rule.n_surr (= 30)
the 2010 rule's **30** IAAFT surrogates	v57	protocol.iv_rule.n_surr
t_MF 's **32** surrogates	v57	protocol.tmf.n_surr
cohort, 2010 rule, SD reading **51/987**	v57	headline.cohort_iv_sd_k / headline.n_sessions
cohort, 2010 rule, RANK reading **63/987**	v57	headline.cohort_iv_rank_k
cohort, 2010 rule, SE reading **478/987**	v57	headline.cohort_iv_se_k
cohort, t_MF published rule **495/987** positive	v57	headline.cohort_tmf_se_k
cohort, t_MF published rule **252** negative	v57	headline.cohort_tmf_se_neg_k
fitted fGn, SE reading **8/24** at N = 830	v57	P3.cells["fgn_fitted@830"].tmf_se_k
fitted AR(4), SE reading **8/24** at N = 830	v57	P3.cells["arma_matched@830"].tmf_se_k
monofractal + the cohort's RT marginal, SE reading **8/24** at N = 830	v57	P3.cells["monofractal_marginal@830"].tmf_se_k
monofractal + one lapse episode, SE reading **10/24** at N = 830	v57	P3.cells["monofractal_lapse@830"].tmf_se_k
monofractal + the RT marginal, SE reading **11/24** at N = 4,096	v57	P3.cells["monofractal_marginal@4096"].tmf_se_k
monofractal + one cusp, SE reading **24/24** significant-negative	v57	P3.cells["monofractal_singularity@830"].tmf_se_neg_k
binomial cascade, 2010 rule SD reading **0/24** at N = 830	v57	P2.cells["binomial_cascade@830"].iv_sd_k
binomial cascade, t_MF SE **24/24** at N = 830	v57	P2.cells["binomial_cascade@830"].tmf_se_k
binomial cascade, t_MF SD **22/24** at N = 830	v57	P2.cells["binomial_cascade@830"].tmf_sd_k
binomial cascade, t_MF RANK **20/24** at N = 830	v57	P2.cells["binomial_cascade@830"].tmf_rank_k
binomial cascade, t_MF RANK **22/24** at N = 1,660	v57	P2.cells["binomial_cascade@1660"].tmf_rank_k
binomial cascade, t_MF RANK **24/24** at N = 4,096	v57	P2.cells["binomial_cascade@4096"].tmf_rank_k
σ-equivalent twin, t_MF SE **16/24** at N = 830	v57	P2.cells["sigma_binomial_equiv@830"].tmf_se_k
σ-equivalent twin, t_MF RANK **8/24** at N = 830	v57	P2.cells["sigma_binomial_equiv@830"].tmf_rank_k
msm-Cox latent cascade, t_MF SE **11/24** at N = 830	v57	P2.cells["msm_logG@830"].tmf_se_k
msm-Cox latent cascade, t_MF RANK **2/24** at N = 830	v57	P2.cells["msm_logG@830"].tmf_rank_k

number (as used)	record	JSON key / path
the songs' positive reproduces, **21/23**	v60 (10.5281/zenodo.22759707)	`headline.reproduction.k_pass` / `.n_songs`
the envelope's **59%**-zero marginal	v60	mean of `P1.rows[].zero_frac` over the 23 songs
an IAAFT surrogate of the reference passes its own IAAFT ensemble **7/24**	v60	`headline.families.iaaft_target.k` / `.n`
— the same rate as a fraction, **29%**	v60	`headline.families.iaaft_target.rate`
the SE reading fires on **11–24 of 24** draws of every calibrated null	v60	min/max of `P3.cells[].pass_se_k` over the 24-seed cells
`t_MF` detects the deterministic binomial **48/48**	v38 (10.5281/zenodo.22759652)	`headline.binomial_pooled_k` / `.binomial_pooled_n`
iid Student-`t` monofractal, two-sided significance at `2.04` on **35/48**	v38	`headline.student_fp.n_sig_two_sided` / `.n`
isolated-singularity monofractal, two-sided significance at `2.04` on **45/48**	v38	`headline.singularity_fp.n_sig_two_sided` / `.n`

Numbers quoted in this note without bold — the exact-Clopper–Pearson intervals, the cohort's mean widths, the budget tables, the phasmid reproduction's `r = 0.885`, the trimming sensitivity cell, and the per-cell rates of §5's longer-length rows — are in the same records' README tables and JSON, and are not separately gated.