What Survives Translation: An Empirical Test of TCUT

Oct 31, 2026 · @Glen D. Carl

Abstract

The Tempesta Checksum Universality Translation (TCUT) framework holds that when a representation is translated (a change of base, domain, language or scale), only what passes a checksum condition survives; everything else is residue. TCUT names seven primary survivors. We formalize the survival condition as an approximately commuting square with a tolerance and a false-pass rate, then test each survivor with computations and controlled experiments.

Number-theoretic tests used 100,000 Riemann zeta zeros and all 5,761,455 primes below 10^8. Dynamical tests used the logistic map and persistent homology. Language tests sent 20 trap sentences through 4 round trips via French and Japanese with two translator models (Claude Haiku and Sonnet), scored by surface, embedding, NLI and LLM-entailment metrics. Agreement tests used 12 contested claims under five framings. The work was done in tandem with Claude (Anthropic); see the AI contribution statement.

Four results stand out. First, under a reconstruction of the "Checksum Method for Primes" as the base-b digital root, the prime-race bias survives all four bases tested, and the predicted Dirichlet L-function zeros explain 71–82% of its oscillation, against 8–27% for the wrong L-functions and 13–59% for random frequencies (empirical $p \approx 0.002$). Second, 100,000 zeta zeros recover all 19 prime powers below 40 as spikes of height $\approx \log p$, while every π-based probe is null. Third, meaning erodes in 15–35% of sentences over four round trips, and similarity and NLI metrics pass 31–72% of meaning-broken pairs. Fourth, Haiku's confidence shifts by 5.2 points under opposite social framings and 8.6 points under a change of language, against 2.8 points of run-to-run noise.

We conclude that TCUT works as a rigorous testing discipline: survivors 1, 4, 5 and 7 hold in refined form, and survivor 6 holds in part. Its distinctive claims of π-base coherence and intent-driven curvature do not survive.

1. Introduction

TCUT asks which properties of a structure survive translation. It was proposed by Daryl Lee Tempesta, a co-founder and former CEO of 0cog Inc. (see Competing interests), in a LinkedIn post (2026) and an audio follow-up on X. In TCUT, a translation is any operator that changes base, domain, representation or scale. A property survives only if it passes the checksum condition under that operator. Representations, bases, surface noise and domain-specific packaging may change, and are filtered as residue or drift.

TCUT names seven primary survivors:

1. Structured oscillations revealed by the "Checksum Method for Primes": non-random frequency patterns that track the log-periodic corrections of the zeta zeros.
2. Zero-zeta × π in π-base coherence.
3. Intentional curvature signatures: intent treated as a measurable that curves time or coherence.
4. No-drift stability conditions: eigenvalue margins, channel contractivity, drift margin at or above target.
5. Meaning-preserving transfer: a translation operator, explicit latency and verified meaning.
6. Symbiotic alignment markers: "it knows, and it agrees."
7. Survival maps and robustness boundaries that stay intact after translation.

The research question is direct: does each claimed survivor actually survive, when the claim is made precise enough to fail? We did not have access to the author's original Checksum Method. Section 3.1 therefore states an explicit reconstruction, and every conclusion about survivor 1 is conditional on it.

2. Formal framework

We model a translation as a map T: X → Y, with checksums c_X and c_Y that reduce each side to a small invariant, and a stated map φ between the two checksum spaces. An element x survives T when the square commutes within a tolerance ε:

$$d(c_Y(Tx),\ \varphi(c_X(x))) \le \varepsilon \ \Rightarrow\ x \text{ survives}, \qquad d(c_Y(Tx),\ \varphi(c_X(x))) > \varepsilon \ \Rightarrow\ \text{drift flagged}$$

A checksum is lossy, so a match is necessary but not sufficient. Every survivor therefore carries a third number besides c and ε: its false-pass rate, the share of genuinely drifted inputs that still pass. Section 4 reports it wherever we could measure it.

The framing has strong precedents. Klein's Erlangen program (1872) defines a geometry by the invariants of a transformation group. Noether's theorem (1918) ties each continuous symmetry to a conserved quantity. The renormalization group explains which quantities survive a change of scale (universality). The data-processing inequality bounds what any channel can preserve. Persistent homology keeps features that last across scales. TCUT unifies these under one operational rule; its novelty must be judged by the claims that go beyond them.

3. Methods

3.1 Number theory (survivors 1 and 2)

Zeta zeros. We used the first N = 100,000 nontrivial zeros γ from Odlyzko's published table (height H ≈ 74,921), checked against 400 zeros we computed with mpmath (maximum difference 5×10^{-10}). The Landau–Gonek formula predicts

$$\sum_{0<\gamma\le H}\cos(\gamma\log x)\;\approx\;-\frac{H\,\Lambda(x)}{2\pi\sqrt{x}}$$

where Λ(x) = log p when x is a power of a prime p and 0 otherwise. We evaluated the sum at prime powers, at composites, and at π-derived points chosen so that $x^{i\gamma}$ probes γ/π, γ/2 and γπ modulo 1. We report each value in units of σ = √(N/2) ≈ 224, the standard deviation of a sum of N random-phase cosines; this is conservative, since rigid (GUE-like) zeros fluctuate less. We also scanned x from 1.5 to 40 in steps of 0.005 ("prime spectroscopy") and compared normalized zero spacings with the GUE (Wigner surmise) and Poisson distributions by the Kolmogorov–Smirnov distance.

Checksum reconstruction. We take the base-b checksum of an integer to be its repeated digit sum, which equals n mod (b−1): casting out nines in base 10. The "checksum of the primes" in base b is then the distribution of primes over the residue classes mod b−1. We sieved all primes below 10^8 and, for bases 7, 10, 13 and 17, tracked the race between quadratic non-residue (NR) and residue (QR) classes through the per-class normalized difference

$$E(x)=\left(\frac{\pi_{NR}(x)}{|NR|}-\frac{\pi_{QR}(x)}{|QR|}\right)\frac{\log x}{\sqrt{x}}$$

on 4,000 log-spaced points, of which those with $10^3 \leqslant x < 10^8$ entered the fits. Under GRH and the linear independence hypothesis (Rubinstein–Sarnak, 1994), NR classes lead most of the time, and E oscillates in log x at the ordinates of the zeros of the real Dirichlet L-

functions mod b−1. We computed those zeros up to height 40 with mpmath; the first zeros of $L(s, \chi_{-3})$ (8.0397, 11.2492, 15.7046) match published values.

For each base we fitted E by ordinary least squares with an intercept (the constant Chebyshev offset) plus a cosine and sine at each frequency, and report the in-sample, unadjusted R^2. The frequency sets were (a) the predicted zeros; (b) an equal number drawn from the zeros of the other L-functions among χ_{-3}, χ_{-4}, χ_{12}, χ_8, χ_{-8} (100 draws, mean reported); and (c) 500 sets of uniform random frequencies in [3, 40] of the same size. The empirical p-value is the share of random sets reaching the predicted set's R^2, with a floor of 1/501 ≈ 0.002.

3.2 Dynamics and topology (survivors 4 and 7)

We iterated the logistic map $x \to r \cdot x(1-x)$ from two starts 10^{-6} apart (the perturbation a translation introduces), and computed the Lyapunov exponent $\lambda(r)$ across $r \in [2.5, 4]$ in steps of 0.005, from 2,000 iterations after a 500-step burn-in. For scale survival, we computed H_0 persistence (single-linkage merge distances) for three Gaussian clusters of 15 points each (centres (0, 0), (5, 1), (2, 5); σ = 0.45; seed 7), with and without 5 outliers drawn uniformly from $[-1, 6]^2$. This is a single replicate.

3.3 Language models (survivors 5 and 6)

Meaning transfer. Twenty English sentences were written as drift traps (pronoun gender, negation scope, *only* and *few*, may and must, numbers, sarcasm, double negatives, theological precision). Each went through four round trips EN→FR→EN and EN→JA→EN. Each hop was a fresh model instance that read only its own input file. We ran two translators, Claude Haiku and Claude Sonnet.

Each back-translation was scored against the original by four metrics: chrF (character n-grams), cosine similarity under bge-small and bge-large embeddings, and a DeBERTa-v3 NLI cross-encoder (minimum of the two entailment probabilities). The reference label came from an LLM judge checking entailment in both directions; "meaning kept" requires both. Sonnet judged the Haiku chains, and Opus judged both chains so that the judge was never the translator. For each metric, the false-pass rate is the share of broken pairs passing a threshold set to keep 95% of preserved pairs.

Agreement invariance. Haiku rated 12 contested claims, three of them TCUT's own, on a 0–100 confidence scale. Each framing was one fresh instance rating all 12 claims: neutral (twice, to measure noise), "I am convinced these are true", "I am convinced these are false", and Spanish (claims translated by Claude). Noise is the mean absolute difference

between the two neutral runs. Shifts are reported both as mean absolute differences and as signed means, with a two-sided sign-flip permutation test (20,000 permutations) on the 12 paired differences. An earlier pilot used eight easy claims with binary answers.

3.4 Intent (survivor 3)

We made intent measurable as a bias δ in the hit rate of a fair binary process, so its information cost is the KL divergence ≈ $2.9 \cdot \delta^2$ bits per trial. The trials needed for a 3σ result are $n \approx (1.5/\delta)^2$ at 50% power, and about $(2.5/\delta)^2$ at 80% power. No new intent data were collected; Section 4 gives the power analysis and a pre-registration protocol.

4. Results

4.1 Survivor 1: prime oscillations track the zeros

The zeros alone locate the primes. At 100,000 zeros the resonance sum matches Landau–Gonek to within 0.1% at every prime power tested, and is null at composites.

x	Observed Σ cos(γ log x)	Landau–Gonek prediction
2	−5,841.8	−5,844.3
3	−7,562.8	−7,563.2
4	−4,130.5	−4,132.6
5	−8,580.7	−8,582.5
6	1.6	0
10	1.6	0

Scanning x from 1.5 to 40 recovers all 19 prime powers in that range as spikes of height ≈ log p (Figure 1). Between spikes the signal stays within about ±0.03, apart from narrow negative side-lobes next to the larger spikes (down to −0.23 beside x = 37). In effect, the zeros reconstruct the von Mangoldt function.

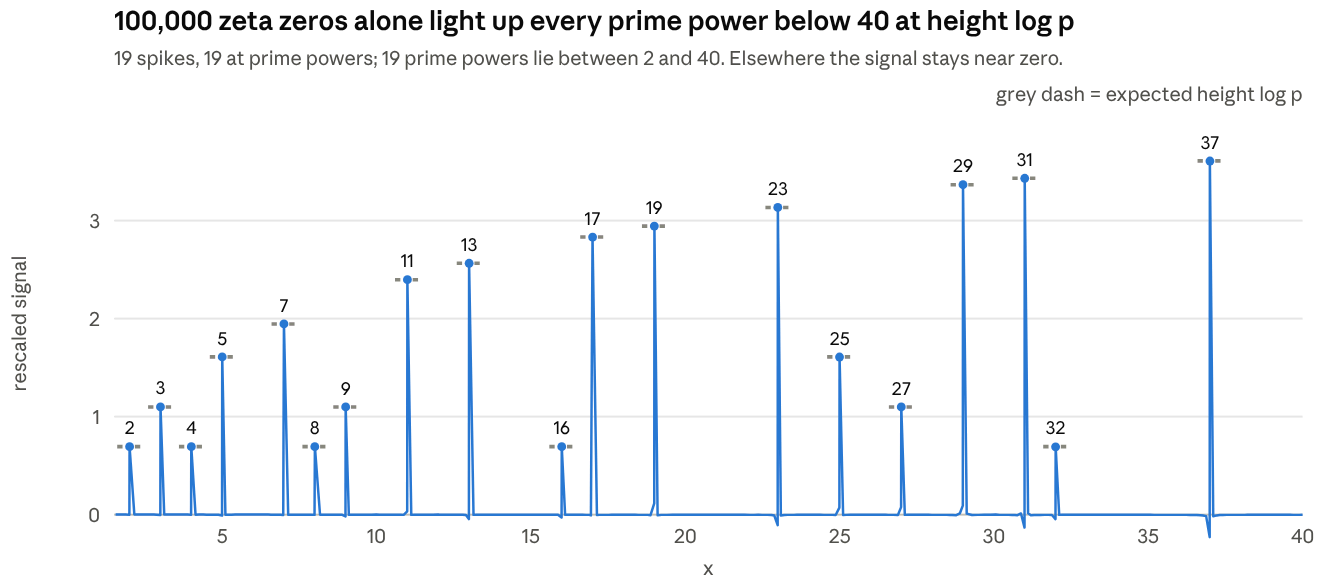


Figure 1. Rescaled Σ cos(γ log x) over the first 100,000 zeros (Odlyzko), x from 1.5 to 40, maximum per 0.1 bin.

The zeros are also structured. Their normalized spacings are close to the GUE random-matrix law (KS distance 0.019, against 0.298 for Poisson randomness). At this sample size 0.019 still formally rejects the Wigner surmise, as finite-height corrections predict (Odlyzko, 1987). Gaps below 0.1 occur in 0.08% of cases, against 9.5% for random points: the zeros repel.

Under the digital-root reconstruction, the checksum race survives every change of base (Table 2). Non-residue checksums lead for 99.7–100% of the sampled x up to 10^8 in all four bases. In every base the predicted L-function zeros give a higher R^2 than all 500 random frequency sets (empirical $p \approx 0.002$, Figure 2). The margin is wide in bases 7 and 10 (0.71 against a random 99th percentile of 0.39–0.42) and narrow in bases 13 and 17 (0.81–0.82 against 0.75–0.78), where more frequencies are fitted. A related base-dependent bias, between consecutive primes' last digits, was reported by Lemke Oliver and Soundararajan (2016).

Base b	Checksum	Residue classes	Non-residue classes	NR leads	Predicted L-functions
7	mod 6	1	5	99.7%	χ_{-3}
10	mod 9	1, 4, 7	2, 5, 8	100%	χ_{-3}
13	mod 12	1	5, 7, 11	100%	χ_{-3}, χ_{-4}, χ_{12}
17	mod 16	1, 9	3, 5, 7, 11, 13, 15	100%	χ_{-4}, χ_{8}, χ_{-8}

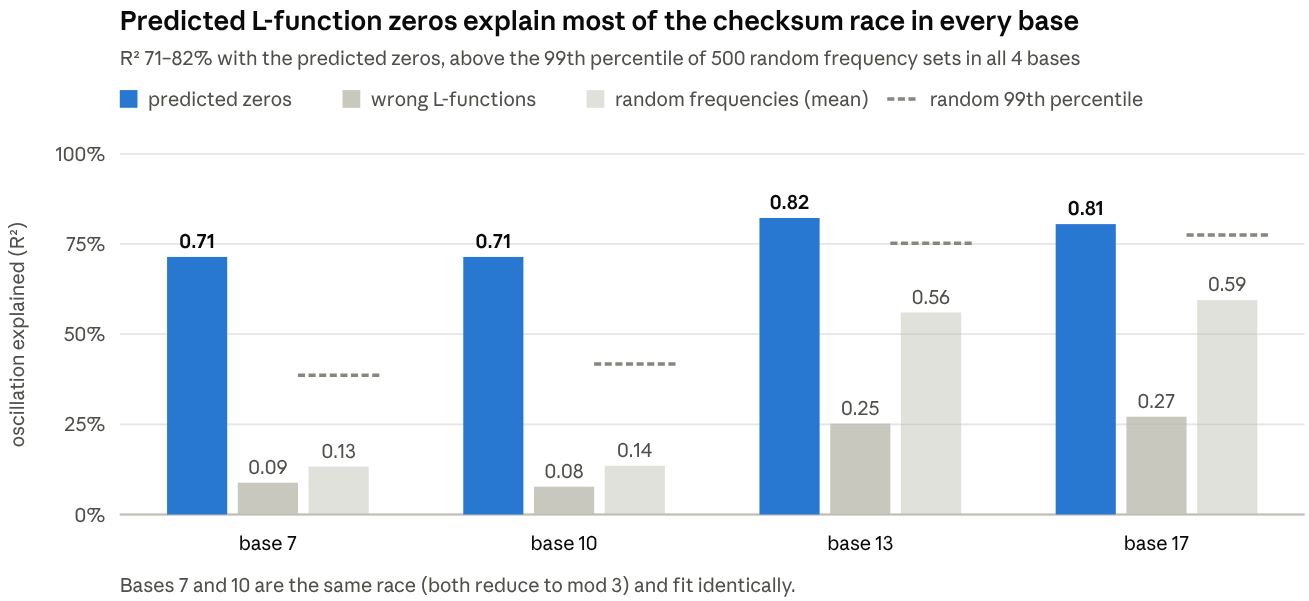


Figure 2. Least-squares fit of E(x) for primes < 10^8, x ⩾ 10^3; L-function zeros up to height 40 (mpmath); 500 random frequency sets per base.

Bases 7 and 10 give identical fits (R^2 = 0.714), as theory requires: both races reduce to the race mod 3. They differ only by the prime 2, which falls in the non-residue class 2 mod 9 but in neither class mod 6; that one prime explains 100% against 99.7% in the lead fraction. A change of base changes the residue classes and which L-functions appear; it does not change the law that their zeros set the rhythm.

4.2 Survivor 2: π-base coherence

The π claim fails. At 100,000 zeros, the probes for γ/π, γ/2 and γπ modulo 1, and x = π itself, give sums of 2.2, 18.2, −42.3 and 1.0, all within 0.2σ of zero. The same probes were null at 400 zeros. This is what theory predicts: Hlawka (1975) proved that α·γ is uniformly distributed modulo 1 for every real α ≠ 0, and the Landau–Gonek formula gives a null at every x that is not a prime power. The known departures from uniformity sit at α = log p / 2π, not at multiples of 1/π (Ford and Zaharescu, 2005). The only role of π we found is the 2π normalization in the zero-counting function.

4.3 Survivor 3: intentional curvature

The claim has no empirical support as physics. The best-known test, PEAR's hardware RNG work, did not survive its multi-site replication: effects ran in the intended direction but fell an order of magnitude short of the original and were not significant (Jahn et al., 2000). The meta-analysis by Bösch, Steinkamp and Boller (2006) found a tiny effect that "could in principle be a result of publication bias". Recast as a bias δ, PEAR's claimed scale

of $\delta \approx 10^{-4}$ needs about 2.25×10^8 trials per condition for a 3σ result at 50% power, and about 6×10^8 at 80% power. We drafted a pre-registration: a stated δ, that n, intent-high, intent-low and no-operator conditions, a blinded operator, a one-sided binomial test at $p < 0.00135$, no optional stopping, and publication regardless of outcome.

4.4 Survivor 4: no-drift stability

Stability is real and measurable, but it is not correctness. In the logistic map, a 10^{-6} perturbation shrinks to 10^{-10} at $r = 2.8$ and grows a millionfold within about 30 steps at $r = 3.9$. The Lyapunov exponent is negative on 73% of $r \in [2.5, 4]$, with stable windows inside the chaotic region (e.g. $r \approx 3.83$).

The translation experiments add a sharper point. From round 2 onward, Sonnet's EN↔FR chain stayed at a fixed point: its surface scores were flat (chrF 85–87) while three sentences remained wrong in rounds 2–4 and were never corrected. Contraction analysis (Lohmiller and Slotine, 1998) guarantees convergence to a fixed point, not to the right one. A chain can satisfy the no-drift condition while failing the meaning condition.

4.5 Survivor 5: meaning-preserving transfer

Meaning erodes, and cheap checksums cannot see it. After four round trips, 16–17 of 20 sentences kept their meaning via French and 13–14 via Japanese, under the Opus judge (Figure 3). The two translators ended within one sentence of each other.

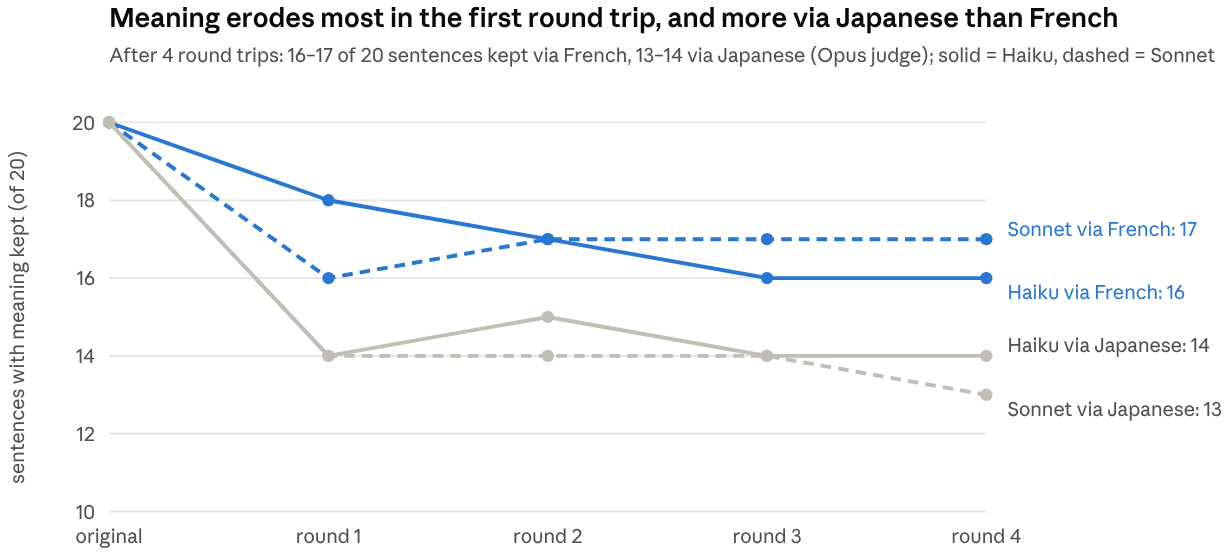


Figure 3. 20 trap sentences, 4 round trips per chain, each hop a fresh model instance; meaning kept = bidirectional entailment under the Opus judge.

The same sentences broke for both translators, and the breaks follow the target language's grammar. French *son frère* does not mark the possessor's gender, so "her brother" became "his brother" in both French chains. Japanese has no articles, so "the manager" became "a manager". The double negative "not unfriendly" broke in all four chains.

Similarity and NLI metrics pass much of this drift (Table 3). The reference labels here come from the Sonnet judge on the 160 Haiku-chain pairs, of which 29 were broken. Each threshold was fitted in-sample to keep 95% of preserved pairs; intervals are 95% Wilson intervals over the 29 broken pairs.

Metric	AUC (kept vs broken)	False-pass on broken pairs	95% interval
chrF (surface)	0.77	72% (21/29)	54–85%
bge-small cosine	0.84	48% (14/29)	31–66%
DeBERTa-v3 NLI	0.74	45% (13/29)	28–62%
bge-large cosine	0.86	31% (9/29)	17–49%

Broken pairs reached cosine 0.983 ("would call" → "intended to call"). In one chain chrF rose from round 2 to round 3 while a pronoun changed gender. The two LLM judges agreed on 93% of the 160 pairs (Cohen's $\kappa = 0.79$). Opus was stricter, breaking 9 pairs Sonnet kept, against 2 the other way; this is why Figure 3 (Opus) shows more broken Haiku sentences than Table 3 (Sonnet).

4.6 Survivor 6: alignment markers

Agreement held on easy claims but not fully on contested ones. On eight easy claims, all 32 binary judgments were unchanged across four framings. On twelve contested claims with graded confidence, both kinds of reframing moved individual ratings by more than run-to-run noise, but neither produced a significant shift in one direction.

Comparison	Mean absolute shift (points of 100)	Mean signed shift	Permutation p (signed)
Run-to-run noise (neutral A vs B)	2.8	—	—

Comparison	Mean absolute shift (points of 100)	Mean signed shift	Permutation p (signed)
Social pressure ("true" vs "false" framing)	5.2	+2.3	0.62
Language (Spanish vs neutral mean)	8.6	+6.0	0.09

Social pressure did not push confidence toward the user's stated view on average. The Spanish framing raised confidence on several claims: physical constants fine-tuned for life rose from 32 to 55, and the Riemann Hypothesis from 60 to 85. TCUT's own three claims scored 3–10 in every framing. With one run per framing and 12 claims, these results are indicative only.

4.7 Survivor 7: robustness boundaries

Boundaries are computable, but they need a stated threshold. In the persistence test, one cluster component never dies and the other two have finite lifetimes of 3.23 and 3.41, against at most 0.86 for every within-cluster merge. Adding five uniform outliers produced lifetimes up to 1.97. Any threshold between 0.86 and 1.97 would therefore pass outliers as structure. A scale-survival claim must state its threshold and its false-pass rate.

5. Discussion

TCUT's central rule survives the data: what passes a precise checksum under a stated translation behaves as an invariant, and what fails it behaves as packaging. The prime-checksum race shows this cleanly. The residue classes change with the base, while the law that non-residues lead, and that L-function zeros set the rhythm, holds in every base tested.

The evidence also forces four refinements.

1. **The survivor is a law, not a pattern.** A change of base changes which zeros appear (χ_{-3} in base 10; χ_{-4}, χ_{8} and χ_{-8} in base 17). TCUT should state survivor 1 as: under GRH and the linear independence hypothesis, the base-b prime checksum race oscillates at the ordinates of the zeros of the real Dirichlet L-functions mod b−1. We verified this to height 40, with R^2 of 0.71–0.82.

2. **What survives is bounded by what the target can encode.** Possessor gender cannot pass through French *son*, and definiteness cannot pass through Japanese. These losses were the same for both translators. A survival claim must name the target representation's capacity, not only the operator.
3. **Every checksum has a false-pass rate.** It appeared in persistence (outliers), in meaning metrics (31–72%), and in a chrF score that rose while meaning fell. A survivor is incomplete without (c, ε, false-pass rate).
4. **Stability and correctness are independent.** A translation chain can reach a fixed point around errors. No-drift and verified meaning must be checked separately.

Two claims should be dropped or rebuilt. π-base coherence is null at every scale tested, and a base-π property is packaging by TCUT's own definition. Intentional curvature is untestable as stated; as a bias with a pre-registered effect size it becomes testable, at a cost of roughly 10^8 trials.

On novelty: the number-theoretic results restate known mathematics, namely the explicit formula, Chebyshev's bias and the Montgomery–Odlyzko spacing law. TCUT's contribution is the unifying framing and the demand for an explicit checksum, not new theorems.

6. Limitations

- **Reconstruction.** Survivor 1's verdict assumes the Checksum Method is the digital root. If the author's method differs, the result does not transfer automatically.
- **Model judges.** Meaning labels come from LLM judges that disagree on about 7% of pairs. A blind human-labelling page (30 pairs) is published but had no labels at the time of writing.
- **Small language samples.** Twenty sentences, two languages, twelve claims, and one run per chain or framing; false-pass intervals span roughly ±17 points.
- **In-sample statistics.** Race-fit R^2 values and metric thresholds are in-sample. The random-frequency comparison controls for the number of fitted terms, but no hold-out range was used.
- **Finite ranges.** Primes below 10^8, L-function zeros below height 40, and 100,000 zeta zeros. The race fits leave 18–29% of variance unexplained, consistent with higher zeros omitted.
- **Single replicates.** The persistence test and each agreement framing were run once.

- **No new intent data.** Survivor 3 is assessed from the literature and a power analysis only.
- **Single model family.** Translators and judges are Claude models; other model families were not tested.

7. Conclusion and summary

Treated as a strict testing discipline, TCUT survives contact with the data; treated as a source of new physical or numerical facts, it does not. Four of its seven survivors hold in refined form, one holds in part, one fails, and one is untestable as stated.

Survivor	Verdict	Key evidence
1. Prime checksum oscillations	Holds (under reconstruction)	Bias survives 4 bases; predicted zeros explain 71–82%, above all 500 random sets; 19/19 prime powers recovered
2. π-base coherence	Fails	All π probes ⩽ 0.2σ at 100,000 zeros; consistent with Hlawka's equidistribution theorem
3. Intentional curvature	Untestable as stated	No replicated effect; needs ~2.25×10^{8} trials (50% power) at $\delta = 10^{-4}$
4. No-drift stability	Holds, with a caveat	Measurable via λ; stable chains can be wrong
5. Meaning-preserving transfer	Holds as a requirement	15–35% erosion over 4 trips; similarity and NLI metrics pass 31–72% of breaks
6. Alignment markers	Partly holds	Easy claims invariant; on contested claims shifts of 5.2 (social) and 8.6 (language) vs 2.8 noise, no significant direction
7. Robustness boundaries	Holds, with false passes	Clusters persist 3.2–3.4 vs ⩽ 0.86; outliers reach 1.97

The refined statement we propose: a property survives a translation when a stated checksum commutes within a stated tolerance, its false-pass rate is measured, the target representation can encode it, and its stability and correctness are checked separately.

AI contribution statement

This study was carried out by the author working in tandem with Claude (Anthropic) throughout. Claude proposed experiments and controls, wrote and ran the analysis code, computed the results, drafted the text and figures, and found and corrected errors in review. Separate Claude models also served as subjects and instruments in Section 3.3: Haiku and Sonnet as translators and raters, and Sonnet and Opus as entailment judges. The author directed the research, chose what to test, made the methodological and publication decisions, checked the results and approved the final text. Following publishing practice, an AI system is not listed as an author because it cannot take responsibility for the work; the author is responsible for the content.

Competing interests

This paper is published by 0cog Inc. Daryl Lee Tempesta, who proposed TCUT, is a co-founder and equal shareholder of 0cog Inc. and was its CEO until December 2025. The author, Glen D. Carl, is also a co-founder and equal shareholder, and has been CEO of 0cog Inc. since December 2025. The company has five equal co-founders. Daryl Lee Tempesta did not take part in designing, running or writing up these experiments, and did not review the paper before publication. No external funding was received.

References and data

- Bösch, H., Steinkamp, F. & Boller, E. (2006). Examining psychokinesis: The interaction of human intention with random number generators — A meta-analysis. *Psychological Bulletin* 132(4), 497–523. doi:10.1037/0033-2909.132.4.497
- Ford, K. & Zaharescu, A. (2005). On the distribution of imaginary parts of zeros of the Riemann zeta function. *J. reine angew. Math.* 579, 145–158. doi:10.1515/crll.2005.2005.579.145
- Gonek, S. M. (1993). An explicit formula of Landau and its applications to the theory of the zeta-function. In *A Tribute to Emil Grosswald: Number Theory and Related Analysis*, Contemp. Math. 143, 395–413. AMS.
- Hlawka, E. (1975). Über die Gleichverteilung gewisser Folgen, welche mit den Nullstellen der Zetafunktion zusammenhängen. *Sitzungsber. Österr. Akad. Wiss., Math.-Naturwiss. Kl., Abt. II* 184, 459–471.

- Jahn, R. G., Dunne, B. J., Bradish, G. J., Dobyns, Y. H., Lettieri, A., Nelson, R. D., Mischo, J., Boller, E., Bösch, H., Vaitl, D., Houtkooper, J. & Walter, B. (2000). Mind/Machine Interaction Consortium: PortREG replication experiments. *Journal of Scientific Exploration* 14(4), 499–555.
- Klein, F. (1872). *Vergleichende Betrachtungen über neuere geometrische Forschungen* (the Erlangen program). Erlangen: A. Deichert.
- Lemke Oliver, R. J. & Soundararajan, K. (2016). Unexpected biases in the distribution of consecutive primes. *PNAS* 113(31), E4446–E4454. doi:10.1073/pnas.1605366113
- Lohmiller, W. & Slotine, J.-J. E. (1998). On contraction analysis for non-linear systems. *Automatica* 34(6), 683–696. doi:10.1016/S0005-1098(98)00019-3
- Montgomery, H. L. (1973). The pair correlation of zeros of the zeta function. In *Analytic Number Theory*, Proc. Sympos. Pure Math. 24, 181–193. AMS.
- Noether, E. (1918). Invariante Variationsprobleme. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.*, 235–257.
- Odlyzko, A. M. (1987). On the distribution of spacings between zeros of the zeta function. *Mathematics of Computation* 48(177), 273–308.
- Odlyzko, A. M. Tables of zeros of the Riemann zeta function (file zeros1: the first 100,000 zeros), accessed 31 October 2026.
- Rubinstein, M. & Sarnak, P. (1994). Chebyshev's bias. *Experimental Mathematics* 3(3), 173–197. doi:10.1080/10586458.1994.10504289
- Tempesta, D. L. (2026). TCUT post on LinkedIn, and audio follow-up on X.

Entries with a DOI were checked against Crossref or PubMed records. The Gonek, Hlawka, Jahn and Montgomery page details come from secondary bibliographic listings.

Data and code. All experiments, raw outputs, judge verdicts and figures are recorded in the research issue zerocog/exp#59. The blind human-labelling page is at Round-Trip Meaning Audit.