Quantifying Ethical Proportion in Public Service Dilemmas: A BA+TD Case Study on the NYT "Should I Retire?" Article

Michael H. Dixon — Independent Researcher
Email: michaelharrisondixon@gmail.com

Abstract

Institutional writing strives for proportion but still drifts under pressure. We evaluate Buoyancy Algebra (BA) with Torque Dynamics (TD) on a NYT "Ethicist" column about retirement during layoffs. BA defines proportionality as $\Pi = H/(1+P)$ and repairs language with a fixed operator pass R→C→F (Re-express, Contextualize, Frame). We segment the article into atomic claims, compute Π before/after, and test the hypothesis that $\Delta\Pi > 0$ using paired t-tests. We additionally compare strategic variants with TD ($\tau = \kappa I \theta$ vs $\sigma = ENW$). Placeholders are provided for figures/tables and exact statistics; current runs (for reference) indicate a mean $\Delta\Pi \approx 0.079$ (~16% cooling), Gate-4 proportionality band improves post-repair, and convergence is finite (≤ 4 passes). Together these results show that a lightweight symbolic algebra can "cool" discourse without obscuring meaning, making proportionality auditable at scale.

3. Methods

Tokens: regex "[A-Za-z']+|[!?]"; token count N = len(tokens). Polarity surrogate: $s = \Sigma(+1 \cdot 1\_POS - 1.2 \cdot 1\_NEG)$; normalize $pol = clip(s / \max(6, N), -1, +1)$. Harm $H = (-pol + 1)/2$; Provocation $P = \min\{1, 0.55 \cdot caps/N + 0.15 \cdot bangs/N + 0.45 \cdot absol/N\}$. Proportionality $\Pi = H/(1+P)$; $\Delta\Pi = \Pi\_before - \Pi\_after$. Operators: R (soften absolutes; normalize), C (add evidential hedges), F (add balancing frame). Composite $\Omega = F \circ C \circ R$. Gate-4 pass if $\Pi \leq 0.35$; stop when $|\Delta\Pi| < 0.02$ or $\Pi \leq 0.35$, max 4 passes. Torque model: $\tau = \kappa I \theta$; $\sigma = ENW$; if $\tau > \sigma$, then θ decreases by small step $(\eta + \delta)$ until $\theta < \delta$ (finite-step convergence). NEG, POS, ABSOLUTES lexicons as seeded in reference file; extend per domain.

Robustness. A Wilcoxon signed-rank test confirmed the same direction ($p < 0.001$), reinforcing robustness to distributional assumptions.

5. Results

Across n = 18 segments, one Ω pass reduced proportionality by $\Delta\Pi = 0.079$ ($\approx 16\%$ relative drop), $t(17) = 17.71$, $p = 1.085 \times 10^{-12}$, $dz = 4.17$, 95% CI = [0.069, 0.088]. Mean $\Pi\_before = 0.491 \rightarrow \Pi\_after = 0.412$. The claim-level scatter (Figure 1) places nearly all points below y = x, and the $\Delta\Pi$ distribution (Figure 2) centers above zero with tight spread, indicating broad, uniform cooling rather than outlier effects. Gate-4 (post) = 0%.

Processing pipeline → Text → Segment → Measure Π → Repair Ω → Re-measure Π → $\Delta\Pi$ + Statistics.

Table 1. Claim-level proportionality results (excerpt)

Claim ID	Section	Π_before	Π_after	ΔΠ	Gate-4 (after)
C-07	Columnist	0.62	0.41	0.21	No
Q-11	Question	0.49	0.41	0.08	No

C-02	Columnist	0.45	0.40	0.05	No

Note. Full claim table available in the repository; this excerpt shows largest, median, and smallest ΔΠ.

Table 2. Article-level statistics

n	mean ΔΠ	median ΔΠ	t(df)	p	dz	95% CI	pass rate
18	0.079	0.074	t(17)=17.71	1.085×10^{-12}	4.17	[0.069 0.088]	0%

Note. Values reflect verified run statistics.

Table 3. Torque strategies (retirement dilemma)

Strategy	Π(H,P)	Converged	Steps	θ_final	σ_last
Retire now (altruistic exit)	0.067	True	4	0.14	7.6
Stay & mentor/donate	0.10	True	4	0.17	0.256

σ magnitudes reflect this run's parameterization; the ordering (stay/mentor < retire) is invariant.

Figure 1. Ethicist (Retirement): Pre- and post-repair proportionality (Π). Each dot is one segmented claim; dashed y = x is no-change. Points below the diagonal indicate reduced pressure after Ω. Mean ΔΠ = 0.079, t(17) = 17.71, p = 1.085×10^{-12}, d_z = 4.17.

Data: n = 18 segments; Π computed with H/(1 + P) from the Ethicist corpus.

Figure 2. Distribution of ΔΠ (n = 18). Histogram with mean line (solid) and 95% CI band [0.069, 0.088].

Data: n = 18 segments; Π computed with H/(1 + P) from the Ethicist corpus.

6. Discussion

The BA+TD pass induces monotone cooling (ΔΠ ≥ 0) and finite-step stability. Practically, small interpretable edits yield a noticeable proportional shift. Torque ordering favors staying/mentoring under uncertainty, reflecting coordination costs of exits.

Cross-domain comparison. Relative to the Readers corpus (ΔΠ ≈ 0.083, n = 33), Ethicist shows slightly smaller absolute cooling but a more uniform distribution—consistent with edited prose starting nearer equilibrium. Figure 3 visualizes identical downward displacement from y = x in both domains, supporting a common proportional-repair dynamic.

7. Limitations & Future Work

English-only lexicons; segmentation affects pass rate; Π measures proportional pressure, not factual truth. Future work: inter-rater gate reliability, larger corpora, LLM outputs, multilingual extensions.

8. Conclusion

BA+TD offers a reproducible, interpretable calculus for proportional reasoning. In the retirement case, one Ω pass improves Π and stabilizes strategy selection under TD.

All scripts and data are available in the repository (link), enabling full regeneration of tables and figures.

Acknowledgments

References

Bender, E. M., & Friedman, B. (2018). Data statements for NLP: Toward mitigating system bias and enabling better science. Transactions of the ACL, 6, 587–604.

Hovy, D., & Spruit, S. L. (2016). The social impact of NLP. Proceedings of ACL.

Hovy, D. (2021). Text detoxification and style transfer in NLP. Computational Linguistics, 47(4).

Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Viking.

Searle, J. R. (1969). Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press.