Global Debate Evaluation Standard (GDES): A Strategic Framework for Restoring Epistemic Integrity to Democratic Discourse

Julian Immanuel Ohm M.Sc.

Version 1.1

Executive Summary

The architecture of contemporary public discourse is structurally unsound. Across the world's leading democracies, a "crisis of credibility" has metastasised into a "crisis of delivery," where the inability to agree on basic facts precludes the possibility of agreeing on effective solutions. Trust in foundational institutions—government, media and business—has eroded to historic lows, fuelled by a perception that these entities are either incompetent or actively dishonest.

Simultaneously, the digital information ecosystem has been fractured by algorithmic amplification. False information travels faster, farther and deeper than the truth, driven by the human predilection for novelty and emotional arousal. This environment overwhelms human cognitive capacity, forcing citizens to rely on heuristic shortcuts rather than analytical reasoning.

The **Global Debate Evaluation Standard (GDES)** is proposed herein as a structural intervention—a rigorous, standardised framework designed to restore clarity, transparency and analytical depth to public discourse. By decomposing complex arguments into three quantifiable dimensions—*Value Significance (VAL), Impact (IMP)* and *Plausibility (PLA)*—and utilising a normalised scoring algorithm, GDES provides a universal grammar for evaluating the strength of competing claims.

1 Part I: The Crisis of Modern Discourse

To understand the necessity of the Global Debate Evaluation Standard, one must first analyse the structural failures within our current discourse environment. We stand at a critical juncture where the mechanisms designed to aggregate public will and translate it into policy are buckling under the weight of information overload and partisan fragmentation.

1.1 The Erosion of Trust and the "Delivery" Crisis

Democracies function on a tacit social contract: citizens grant legitimacy to institutions in exchange for security, prosperity and responsive governance. When this contract is perceived

as broken, trust evaporates. The 2025 *Edelman Trust Barometer* reveals a profound global shift toward grievance, where citizens feel the system is biased against them, fuelling hostile activism and a rejection of traditional authority.

This is often identified as a *delivery problem*. When democratic governments fail to address existential challenges—such as climate change or wealth inequality—citizens increasingly view the deliberative process as an obstacle to progress. The perception is that democracy is "all talk and no action." Into this vacuum step populist actors who promise simple solutions to complex problems, often by bypassing democratic norms and attacking the concept of objective truth.

1.2 The Information Environment: A Cognitive Minefield

The crisis is exacerbated by a digital architecture that privileges engagement over accuracy. Research confirms the structural disadvantage of truth in the current ecosystem. A seminal study published in *Science* analysed the diffusion of news on Twitter and found that false stories travelled significantly farther, faster, deeper and more broadly than the truth. Falsehoods were found to be roughly 70 % more likely to be retweeted than factual stories. The drivers of this phenomenon are rooted in human psychology: false news is often more novel and elicits higher–arousal emotions such as surprise or fear. In a marketplace optimised for virality, nuanced policy debate is competitively disadvantaged.

1.3 Cognitive Bottlenecks: Why We Need "Engineering Math" for Debate

The human brain is not designed for the scale of modern information consumption. Cognitive psychology defines strict limits on our working memory—our ability to process simultaneous variables. While early research suggested a limit of "seven plus or minus two," modern studies suggest the effective limit for active processing is closer to four "chunks" of information. When a citizen is presented with a complex policy debate involving economic forecasts, ethical trade–offs and statistical data, their working memory is rapidly saturated. Once this buffer is exceeded, the brain abandons analytical processing (System 2) and reverts to emotional heuristics (System 1).

We solve similar problems in other fields using formal languages. Engineers do not build bridges using metaphors; they use standardised notation and physics calculations to manage complexity. Civic conversation lacks this "engineering math." GDES supplies this missing notation, compressing intricate concepts into unambiguous scores that fit within human cognitive limits.

2 Part II: The GDES Framework

The Global Debate Evaluation Standard moves beyond binary notions of "true" or "false" to provide a nuanced, three–dimensional profile of any claim. It acknowledges that in complex policy debates, truth alone is often insufficient to determine action; we must also weigh values and consequences.

2.1 The Three Dimensions

Every argument is decomposed into three distinct, quantifiable dimensions, each scored on a scale from 0 to 10:

1. **Value Significance (VAL)** – Why does this matter? It measures the normative weight of the argument by identifying the moral principle, ethical goal or societal priority at stake (e.g. Liberty, Equity, Safety).

   - **0–3 (Marginal):** matters of convenience, slight preference or trivial utility.
   - **4–7 (Moderate):** affects standard of living, economic efficiency or secondary goals.
   - **8–10 (Critical):** involves fundamental rights, existential threats, human life or core democratic principles.

2. **Impact (IMP)** – How large is the consequence? It measures the magnitude, scope and duration of the effect if the argument's premises are realised. While *Value* looks at the quality of the goal, *Impact* looks at the quantity of the effect.

   - **0–3 (Local/Temporary):** minor inconvenience or benefit to a small group; reversible effects.
   - **4–7 (Regional/Sectoral):** significant economic shift, change in law or medium–term consequence.
   - **8–10 (Systemic/Irreversible):** national or global transformation, generational consequence, mass loss of life or irreversible environmental change.

3. **Plausibility (PLA)** – How likely is this to be true? It measures the evidentiary strength and logical coherence of the claim and anchors the argument in reality.

   - **0–3 (Speculative/False):** relies on fallacy, debunked data, conspiracy theory or no evidence.
   - **4–7 (Possible/Probable):** supported by some evidence, correlation or expert opinion but lacking consensus.
   - **8–10 (Certain/Proven):** supported by scientific consensus, rigorous meta–analysis, historical fact or physical laws.

2.2 The Normalised Scoring Formula

GDES utilises a multiplicative formula to calculate the overall *Strength Index* (STR). This formula ensures that a failure in any one dimension—for example, zero evidence—collapses the entire argument:

$$\text{STR} = \frac{\text{VAL} \times \text{IMP} \times \text{PLA}}{10}$$

The result is a *Strength Index* between 0 and 100. The logic of the formula is straightforward:

- **Zero means zero:** if an argument is factually false (PLA = 0), it has strength 0, no matter how serious the *Impact* is. Similarly, if an argument is true but affects something nobody cares about (VAL = 0), it also has strength 0.

- **Proportionality:** doubling the stakes (*Impact*) or the moral weight (*Value*) proportionally increases the strength of the argument.

- **Scaling:** dividing by 10 maps the 0–1000 raw product onto an intuitive 0–100 index, keeping the score easy to read and compare.

2.3   Strength Tiers: Interpreting the STR Index

The Strength Index is based on a multiplicative model (VAL × IMP × PLA). This creates a skewed distribution: many combinations yield low scores, while high scores are rarer because all three components must be high at the same time.

To keep the index readable in public and political discourse, GDES groups scores into tiers. These tiers act as a translation layer from a numeric index to a typical level of decision weight: how much an argument should usually matter when weighing competing claims.

Strength tier	Strength Index (0–100)	What it means
Strong	$100 - 50$	Likely to meaningfully influence the decision
Solid	$50 - 20$	Should be weighed seriously
Moderate	$20 - 5$	Relevant support, rarely decisive alone
Marginal	$5 - > 0$	Low weight, unlikely to change the conclusion
No support	0	Not usable as justification (fails an essential component)

2.4   Illustrative Scenarios

The following table summarises several example scenarios. It lists each scenario along with its Value, Impact and Plausibility scores and the resulting strength (STR).

Scenario	VAL	IMP	PLA	STR
Critical Crisis	10	9	9	81
Solid Policy	8	6	7	33.6
Fear Mongering	10	10	1	10
Trivial Fact	1	1	10	1

The strength values above are calculated using STR = (VAL × IMP × PLA)/10.

- In a *critical crisis*, high values across all three dimensions yield a very high strength score, demanding urgent attention and action.

- A *solid policy* achieves moderate scores across the board, resulting in a respectable but lower overall strength; such arguments represent standard policy successes.

- *Fear mongering* combines high values and impact with extremely low plausibility; despite the dramatic stakes, the overall strength remains weak because the evidence does not support the claim.

- A *trivial fact* can be certainly true but of negligible importance and consequence; its strength is correspondingly tiny.

3 Part III: The GDES Methodology (Step–by–Step)

Implementing GDES is an active process of engineering a debate. It requires breaking down the chaotic flow of information into structured components.

Phase 1 – Framing and Decomposition

1. **Frame the Decision Question.** Define the specific question being debated. Ambiguity at this stage leads to incoherence later. For example: "Should the government implement a wealth tax?" (a binary choice).

2. **Decompose Arguments into "Cards."** Complex rhetoric often hides weak points. Break the position down into individual, standalone claims or *Argument Cards*. For example, a politician's speech might contain three distinct cards: (1) economic growth, (2) fairness and (3) implementation cost.

3. **Filter Opinion vs. Argument.** Check if the card has a basis in Value, Impact or Plausibility. If a claim expresses a preference without a reason—for instance, "I just don't like it"—it is an opinion. It receives a score of 0 and is removed from the GDES calculus.

Phase 2 – The Scoring Process

1. **Assign Value Significance (VAL).** Identify the underlying value. Ask: "If we assume this claim is true, how much does the principle behind it matter?" Assign a score from 0 to 10. High values reveal moral weight; if two sides have high VAL scores on different issues (e.g. Liberty vs. Safety), you have identified a value conflict.

2. **Assign Impact (IMP).** Estimate the consequence. Ask: "If this claim is true, what is the scope of the effect?" Assign a score from 0 to 10. Treat the largest conceivable consequence in the debate as a 10; scale everything else relative to that.

3. **Assign Plausibility (PLA).** Verify the truth. Ask: "What is the quality of evidence supporting this claim?" Assign a score from 0 to 10. Use peer–reviewed studies for scores of 8–10, expert consensus for 6–7, and treat unsourced assertions as 0–3.

Phase 3 – Calculation and Aggregation

1. **Calculate Strength (STR).** Apply the formula $\mathrm{STR} = (\mathrm{VAL} \times \mathrm{IMP} \times \mathrm{PLA})/100$ for each card.

2. **Aggregate the "Debate Deck."** Sum the STR scores for all arguments supporting the pro and all arguments supporting the contra positions:

$$\mathrm{Total_{pro}} = \sum \mathrm{STR}_i, \qquad \mathrm{Total_{contra}} = \sum \mathrm{STR}_j.$$

3. **Comparative Analysis.** Compare the totals to determine the winner of the debate based on current evidence.

   - *Clear winner:* if one side exceeds the other by a significant margin (e.g. more than 20 %), the evidence heavily favours that position.

   - *Close call:* if scores are within a narrow margin, the debate turns on the subjective *VAL* settings. This reveals the decision is fundamentally a political choice of values, not a technical one.

Phase 4 – Dynamic Iteration

GDES is a living document, akin to a flight log. Update scores as new information emerges:

- **Update PLA:** when new studies or data emerge, adjust plausibility scores accordingly.

- **Update IMP:** if forecasts change the expected scope of outcomes, adjust impact scores.

- **Update VAL:** if public sentiment or moral priorities shift, adjust value scores.

Reasons for Disagreement. Even a structured framework will not eliminate disagreement entirely. When interlocutors assign different Value scores, the divergence is often personal or philosophical rather than empirical. In such cases it can be illuminating to run the calculation twice, using each party's value assumptions. If both versions still point toward the same conclusion, the dispute over values is immaterial; if not, at least the source of the disagreement is explicit.

By contrast, disagreements over Impact or Plausibility are factual. The proper remedy is to look for more studies, data and examples, or to agree on a standard of evidence that would be convincing. Sometimes the science is not yet clear. GDES can still support provisional recommendations in those cases: "If further research proves X, then Y would be a good proposal, but more study is needed," or "We will try X, but if Y does not materialise we need to stop this policy." Pilot programmes and experiments may be appropriate, but the criteria for success and willingness to adapt must be specified up front.

4 Part IV: Applications and Case Studies

4.1 Case Study: The Wealth Inequality Debate

Consider the question: "Should the government implement a progressive wealth tax?" Using GDES, we decompose the debate into pro–redistribution and contra–redistribution arguments and score each card.

Pro–Redistribution Cards.

- **Social Stability:** VAL = 9, IMP = 8, PLA = 7.5, STR = 54. Extreme inequality erodes democracy.

- **Public Funding:** VAL = 8.5, IMP = 6, PLA = 6.5, STR = 33.15. Revenue for infrastructure and health.

Total (Pro): 87.15.

Contra–Redistribution Cards.

- **Innovation Incentives:** VAL = 7.5, IMP = 7, PLA = 3.0, STR = 15.75. Taxes punish success and slow growth; the supporting evidence is weak.

- **Philanthropy:** VAL = 6, IMP = 4, PLA = 4.0, STR = 9.6. Private allocation is more efficient.

Total (Contra): 25.35.

Result. Pro (87.15) vs. Contra (25.35). The analysis shows the Pro side has structurally stronger arguments. The contra side relies on high–value arguments (innovation incentives) that currently lack sufficient plausibility to compete. To prevail, opponents of the wealth tax must either bolster the evidence supporting their position (raising PLA scores) or diminish the impact of inequality.

4.2 GDES in the Boardroom

GDES can transform corporate governance by turning abstract company values into operational criteria. Consider a profitable but environmentally damaging product launch.

- **Profit argument:** VAL = 6 (short–term gain), IMP = 4 (quarterly bump), PLA = 9 (likely). STR = $(6 \times 4 \times 9)/10 = 21.6$.

- **Sustainability argument:** VAL = 9 (core value), IMP = 8 (reputation/regulator risk), PLA = 9 (certain). STR = $(9 \times 8 \times 9)/10 = 64.8$.

Since the sustainability argument mathematically outweighs the profit argument, a CEO who ignores this calculation implicitly downgrades the company's stated values.

4.3 Journalism: The "Nutrition Label" for News

GDES allows journalists to report objectively without creating false equivalence. Instead of "he said, she said," a GDES–compliant report can append a scoreboard. For example:

- **Claim:** "The election was stolen."

- **GDES Score:** VAL = 10 (democracy) $\times$ IMP = 10 (regime change) $\times$ PLA = 0 (no evidence) $\Rightarrow$ STR = 0.

This visualisation allows the media to report the claim while simultaneously contextualising its truth value, moving journalism from access to assessment.

5 Part V: Conclusion

The crisis of democracy is not one of capacity; it is a crisis of processing. We possess the data, the resources and the technology to solve the great challenges of our time. What we lack is a shared language for evaluating choices that is immune to the distortions of outrage and the fog of misinformation.

The Global Debate Evaluation Standard offers such a language. By enforcing a rigorous separation of *Value*, *Impact* and *Plausibility*, and by holding arguments accountable to the logic of (VAL $\times$ IMP $\times$ PLA)/10, we can drain the swamp of performative rhetoric. Implementing this standard—in our legislatures, our newsrooms, our classrooms and our algorithms—is a necessary next step in the evolution of democratic governance.

Citations

1. Edelman Trust Barometer 2024.

2. MIT News, "Study: Twitter false news travels faster than true stories."

3. Soroush Vosoughi *et al.*, *Science* (2018).

4. George A. Miller (1956), "The magical number seven, plus or minus two," *Psychological Review*.

5. Daniel Kahneman (2011), *Thinking, Fast and Slow*.

6. Edelman Trust Barometer 2025 (Key Findings).

7. Edelman Trust Barometer 2025 (Press Release).

8. PBS NewsHour on the Vosoughi study.

9. Nelson Cowan (2001).

10. *Journal of Democracy*, "Delivering for Democracy."

Part VI: Changelog

- **Version 1.1: Strength Index and tiers.** Strength was rescaled from (VAL $\times$ IMP $\times$ PLA)/100 (0–10) to a *Strength Index* defined as (VAL $\times$ IMP $\times$ PLA)/10 (0–100), improving readability for general audiences while keeping the underlying multiplicative logic unchanged. In addition, Strength tiers were introduced to improve interpretability: motivated by the skewed distribution of a product–based score, they provide a reader–friendly mapping from the numeric index to typical decision weight in public and political discourse.