

A Procedural and Naturalistic Model of Moral Objectivity

Abstract

This paper develops Pragmatic Procedural Realism, a naturalistic theory of moral objectivity grounded in systems theory and historical analysis. Extending Emergent Pragmatic Coherentism to metaethics, we treat moral progress as systemic debugging: identifying and removing normative principles that generate catastrophic costs. Unlike Kantian proceduralism's idealized rational procedures, we identify objectivity with historical filtering through pragmatic constraints. We operationalize this through measurable brittleness metrics and construct a Negative Canon of empirically falsified principles, showing that moral objectivity emerges as a procedural fact about which normative architectures prove resilient under real-world testing. This objectivity rests on practical necessity given human constraints, not historical accident. The Apex Network (the structure of viable norms) exists as a constraint-determined fact discovered through pragmatic filtering. The result is a fallibilist realism that naturalizes moral reference while responding to error theory and quasi-realism, reframing moral inquiry as an empirical discipline.

1. Introduction: From Static Gaps to a Dynamic Filter

1.1. A Unified Theory of Justification: Emergent Pragmatic Coherentism (EPC)

Our previous work introduced Emergent Pragmatic Coherentism (EPC) as a general theory of justification. EPC treats inquiry as epistemic engineering: building resilient public knowledge structures whose viability is assessed through their Systemic Brittleness Index (SBI), a measure of real-world costs from misalignment with pragmatic constraints. High costs appear as failed predictions, ad-hoc patches, and accumulating epistemic debt.

This paper extends EPC to metaethics. We argue the is/ought gap results from static thinking that overlooks a unified cost-based justification mechanism. In a dynamic view, both factual and normative claims face the same pragmatic filter. The diagnostic tools for scientific theories can assess social and ethical systems. This dissolves the is/ought problem and grounds Pragmatic Procedural Realism, a naturalistic moral objectivity.

One might object that scientific and normative systems are fundamentally different kinds of entities, making this extension inappropriate. We argue the opposite. At the level of systems dynamics, both are informational architectures designed to solve problems of coordination (science coordinates our beliefs with the causal world, while ethics coordinates our actions with each other). Both generate measurable, real-world costs when their core principles are misaligned

with their respective constraints. This shared functional challenge justifies a unified diagnostic approach.

1.1.1. EPC Foundations: A Brief Overview

For readers unfamiliar with the foundational framework, we provide a condensed summary of EPC's core machinery. Traditional epistemology seeks foundations (axioms, basic beliefs, sense data) or coherence (mutual support within a web of beliefs). EPC offers a third path: justification through demonstrated resilience under pragmatic pressure.

The central insight is that knowledge claims function as engineering specifications for cognitive and social systems. Just as bridge designs are tested by whether bridges stand, knowledge structures are tested by whether they enable successful coordination with reality. This generates a unified standard: a system's justification correlates with its Systemic Brittleness Index, the accumulated costs from misalignment with operative constraints.

The Systemic Brittleness Index (SBI). In science, a theory with high SBI exhibits: failed predictions requiring ad-hoc modifications, accumulating anomalies that resist integration, increasing complexity without corresponding explanatory gain, and vulnerability to replacement by simpler alternatives. Consider Ptolemaic astronomy, which required ever more epicycles to accommodate observational data, signaling rising brittleness. The theory functioned locally but incurred mounting costs. Copernican heliocentrism succeeded by dramatically lowering these costs.

The SBI is a composite diagnostic, not a single metric. It integrates multiple indicators: prediction failures, explanatory gaps, ad-hoc patches (what we call "patch velocity"), and resistance to integration with other domains. High brittleness indicates structural problems requiring debugging, not merely surface errors requiring correction.

Three-Level Framework for Truth. EPC distinguishes three levels of epistemic status:

Level 1: Contextual Truth (Coherence within a System). A claim is contextually true if it coheres with the network's internal rules. In Ptolemaic astronomy, "Mars moves on epicycles" was contextually true—it followed from the system's principles and observational methods. This level provides procedural correctness but no external justification.

Level 2: Justified Truth (External Validation via Track Record). A claim is justifiedly true if the system containing it demonstrates low brittleness over time—minimal ad-hoc patching, successful predictions, integration with other domains, and resilience under challenge. Copernican heliocentrism earned justified truth through its superior track record: fewer epicycles, correct predictions of planetary phases, integration with Newtonian mechanics. This is the highest epistemic status we can actually achieve.

Level 3: Objective Truth (Regulative Ideal). A claim is objectively true if it belongs to the ideally optimal system—the complete set of maximally coherent and pragmatically successful principles. This regulative ideal represents the formal standard against which Level 2 justification is assessed. We know Newtonian mechanics is not objectively true (relativity superseded it), yet it was justifiedly true given Newton's evidence. This distinction preserves fallibilism while maintaining realist commitments.

The Entrenchment Mechanism. How do principles migrate from peripheral hypotheses to core commitments? Through a process of pragmatic entrenchment driven by bounded rationality. A principle begins at the periphery as a testable hypothesis. As it proves indispensable for reducing systemic brittleness-enabling predictions, solving anomalies, integrating domains-revising it becomes increasingly costly. The principle migrates inward, becoming infrastructure that other claims depend on. Eventually, it achieves core status through “systemic caching”: the system embeds it so deeply that revision would require abandoning the conceptual tools needed for coordination itself. Core principles are not self-evident axioms but highly optimized discoveries that have survived extensive testing. Their justification is their proven functional indispensability.

Unified Diagnostic Across Domains. EPC’s key innovation is applying this framework across all domains of inquiry. In science, we measure brittleness through prediction failures and ad-hoc modifications. In mathematics, we measure it through proof complexity and the need for exception-handling. In ethics-the subject of this paper-we measure it through social costs: coercion required to maintain compliance, ideological justifications needed to explain failures, and bio-social costs from violating human needs. The diagnostic toolkit adapts to domain-specific texture while maintaining a unified underlying logic: high systemic costs indicate structural misalignment requiring revision.

The is/ought gap appears unbridgeable only if we assume different justificatory standards for facts and values. EPC shows both are justified by the same mechanism: demonstrated viability under pragmatic constraints. Scientific theories must accommodate physical constraints (experimental results, mathematical consistency). Normative systems must accommodate pragmatic constraints (human biology, coordination requirements, cognitive limitations). Different constraints, same filtering process.

Implications for This Paper. We extend this machinery to metaethics. Normative principles are Standing Predicates-reusable action-guiding concepts that function as social infrastructure. Like scientific theories, they can be elegant in principle but brittle in practice. Our diagnostic tools measure their brittleness: the Coercion Ratio $C(t)$ tracks maintenance costs, Patch Velocity $P(t)$ tracks ideological debt, and bio-social costs track friction with human needs. Principles that generate high costs across these dimensions are debugged through historical filtering, just as failed scientific theories are replaced. The result is Pragmatic Procedural Realism: moral objectivity grounded in the empirical discovery of constraint-determined optimal solutions. The Apex Network-the complete set of maximally viable normative principles-exists as the regulative ideal (Level 3), discovered through historical testing (Level 2), even as particular societies implement imperfect approximations (Level 1).

With this foundation established, we can now develop the specific application to moral progress and normative objectivity.

1.2. Thesis: Moral Progress as Systemic Debugging

Moral principles function like engineering designs for social worlds. Like any design, they can be elegant in theory but flawed in practice. Flawed bridge designs generate stress, cracks, and collapse. Flawed normative designs (such as those built on slavery) generate social stress (dissent, rebellion), structural cracks (coercive costs, economic stagnation), and collapse. We develop diagnostic tools to detect these structural flaws before catastrophic failure.

Our central thesis: moral progress is a real, observable process of systemic debugging, not teleological advance. Applying the SBI framework to history identifies brittle normative predicates (those generating catastrophic costs) and catalogs them in a Negative Canon of falsified moral principles. This reveals moral objectivity as an emergent procedural fact. Moral truths are reverse-engineered from systemic failures, like mapping reefs from shipwrecks.

The argument proceeds in four stages: (1) operationalizing SBI for normative analysis with measurable proxies; (2) applying this to model moral progress as predicate replacement; (3) situating Pragmatic Procedural Realism as a naturalistic alternative in metaethics; (4) defending against objections. The result unifies inquiry: pragmatic system-building discovers objective truth in science and ethics.

1.3. Scope and Limitations

This paper does not solve normativity's ultimate grounding or provide a non-circular defense of valuing survival. Instead, it takes a conditional, descriptive approach. The Constitutive Condition of Persistence serves as a procedural filter: normative systems must endure to be historically analyzable. Persistence is not a smuggled value but the entry requirement for justification. We do not claim to derive categorical imperatives from facts, but rather identify the empirical patterns through which normative claims are tested and refined. Our aim is not proving we ought to persist, but describing the rules of the game we play, building a testable model of the evolutionary process through which values are filtered and earn their authority.

2. The Diagnostic Engine: Operationalizing Normative Brittleness

2.1. Units of Selection: Standing Predicates

EPC provides a unified test for public knowledge systems: claims are justified by the system's demonstrated viability. Drawing from evolutionary theory, we distinguish the informational structure (core normative predicates and their relations) as the replicator-the abstract code transmitted over time. Social groups and institutions serve as the interactor-the physical vehicle for testing this code. A system "survives" by propagating its principles, even if the original group dissolves (as when Roman law was rediscovered in the Renaissance). This avoids naive group selectionism by focusing on long-term viability of the normative code.

This structure consists of Standing Predicates: reusable, action-guiding concepts that function as cultural "genes." A principle like "slavery is acceptable" is not just a statement but a predicate enabling actions, justifications, and social relations. We track these predicates' viability through historical testing.

Consider the normative predicate **...is a binding promise**. When a community treats this predicate as 'standing,' it doesn't just classify an utterance; it automatically licenses a cascade of normative judgments and social actions: the promiser incurs an obligation, the promisee gains a legitimate expectation, and third parties are licensed to apply social sanction (e.g., reputational damage) in

case of non-fulfillment. The viability of this predicate is tested by its long-term success in reducing the costs of social friction and enabling complex cooperation.

2.2. Tiered Diagnostic Framework

To avoid circularity, we arrange costs hierarchically, from basic biological facts to complex systemic effects. The SBI is a composite index; our analysis focuses on three core tiers:

Tier 1: Bio-Social Costs. Direct material consequences of friction with human persistence conditions, measured by objective proxies like excess mortality/morbidity rates, chronic malnutrition, and demographic decline. Systems generating these costs fail fundamentally.

Tier 2: Systemic Friction Costs. Resources expended managing dissent from Tier 1 costs, measured by the Coercion Ratio ($C(t)$), which tracks resources spent on suppression versus production. Rising $C(t)$ indicates high maintenance costs for flawed designs.

Tier 3: Ideological Costs. Informational expenses justifying Tier 1 and 2 costs, measured by Patch Velocity ($P(t)$), the rate of ad-hoc ideological justifications (such as divine mandates for suffering). High $P(t)$ signals accumulating ideological debt in failing systems.

These tiers form a causal cascade. Unaddressed Tier 1 costs (famine) generate dissent, forcing Tier 2 costs (higher $C(t)$ through suppression). To justify these failures, the system generates Tier 3 costs (accelerating $P(t)$). High Tier 3 readings are lagging indicators of deep, unresolved Tier 1 problems.

2.3. Falsifiability and Triangulation

This framework is empirically testable. Robust brittleness diagnosis requires convergent evidence from three baselines: (1) Comparative-Historical analysis against contemporaneous peers, (2) Diachronic comparison against the system's own trajectory, and (3) Biological Thresholds representing non-negotiable viability limits.

The core claim is that systemic costs predict long-term fragility. This would be falsified if historical analysis showed:

1. No Correlation: No significant link between high costs (e.g., violence) and systemic fragility
2. High-Cost Superiority: Coercive systems prove more innovative/resilient than cooperative ones
3. Negative Canon Failure: High-cost predicates (e.g., "slavery acceptable") enhance long-term viability

We acknowledge that measuring these costs is most straightforward in state-level societies with formal institutions. For informal normative systems, proxies must be more creative, relying on data from ethnographic studies, legal records of disputes, or bioarchaeological markers of stress within marginalized subgroups. The core principle remains: the costs are real and have empirical signatures, even when their measurement is indirect.

Calibrating Timescales for Brittleness Predictions

The brittleness framework makes predictions about system fragility, but the relevant timescale varies systematically with system characteristics. Specifying these timescales is essential for falsifiability.

Scale Effects. Small-scale systems (city-states, local communities, organizations) exhibit faster feedback loops. Brittleness manifests in collapse timescales of decades to a century. Large-scale systems (empires, major civilizations, international orders) possess greater inertia and buffering capacity, with complete collapse timescales of centuries to millennia. However, decline indicators (rising $C(t)$, accelerating $P(t)$) typically precede collapse by 50-150 years even in large systems. The Roman Empire's fall took centuries, but brittleness symptoms appeared generations earlier.

Interconnectedness Effects. Isolated systems can persist in high-brittleness states longer because external competitive pressure is delayed. Pre-modern empires with geographic buffers (mountain ranges, deserts, oceans) could sustain inefficient configurations for extended periods. Interconnected systems face accelerated filtering-competitive pressure forces collapse or adaptation more quickly. Modern nation-states in a global economy cannot sustain high-brittleness configurations as long as geographically isolated historical empires.

Equilibrium Brittleness versus Collapse Timing. A critical distinction. We predict brittleness-collapse correlation, not deterministic timing. High-brittleness systems: (1) always incur higher maintenance costs (by definition-that's what brittleness measures), (2) show characteristic warning signs (rising $C(t)$, accelerating $P(t)$, accumulating systemic debt), (3) are more vulnerable to shocks (external pressures, internal crises, succession problems), and (4) eventually collapse or fundamentally transform. Timing varies with scale, context, and shock magnitude, but the pattern holds.

Probabilistic Predictions: Our predictions are conditional and probabilistic: "System X with brittleness profile Y has Z% probability of major crisis within W years, conditional on shocks of magnitude M." These are calibrated using historical base rates for comparable systems, with confidence intervals widening for longer predictions. We're not offering deterministic prophecy but epistemic risk assessment grounded in historical patterns.

The "Successful Coercion" Illusion: Long persistence of high-coercion systems doesn't falsify our model if: (1) $C(t)$ was rising over time (indicating increasing fragility), (2) the system underwent periodic collapses and reconstitutions (showing brittleness), (3) decline indicators preceded ultimate collapse, or (4) comparison to peers shows higher costs and greater vulnerability. What would falsify our claim: a system with sustainably low and stable $C(t)$ and $P(t)$ that nonetheless collapsed, while peers with high $C(t)$ and $P(t)$ proved more durable. The historical record provides no clear examples of this pattern.

Prospective Application: For contemporary systems, we diagnose rising brittleness before collapse by identifying rising $C(t)$ trends, tracking $P(t)$ acceleration, monitoring bio-social cost indicators, and comparing to historical patterns. This provides actionable epistemic risk assessment. We cannot predict exact timing of collapse (too many variables), but we can assess relative fragility and identify systems requiring urgent debugging. This is analogous to earthquake science: we cannot predict when specific earthquakes will occur, but we can map fault lines and assess seismic risk.

2.4. Operationalizing Brittleness: A Worked Example

To demonstrate empirical testability, we provide a detailed worked example showing how to operationalize $C(t)$, $P(t)$, and bio-social costs for a well-documented historical case.

Case Study: The Antebellum South (1830-1860)

The slave system of the American South provides an ideal test case: abundant documentation, clear normative architecture centered on “slavery is acceptable,” and known historical outcome (catastrophic collapse 1861-1865). We can retrospectively calculate brittleness metrics and assess whether they predicted the system’s fragility.

A. Measuring $C(t)$: The Coercion Ratio

Operational Definition: $C(t) = (\text{Total resources dedicated to internal coercion and suppression}) / (\text{Total economic output})$

Data Sources: - State and local government budgets (militia, slave patrols, judicial systems) - Private expenditures on surveillance and enforcement (overseers, slave catchers, weaponry) - Opportunity costs of labor diverted to coercion (white men in patrol service, guard duty) - Insurance and compensation systems for captured fugitives - Infrastructure costs (jails, patrol stations, communication networks for suppression)

Sample Calculation for Virginia, 1850:

Total economic output (1850): Approximately \$220 million (agricultural and industrial production)

Coercion costs: - State militia and patrol appropriations: ~\$800,000/year - County-level slave patrols and constabulary: ~\$1.2 million/year - Private overseers and supervision (est. 5,000 overseers \times \$400/year): ~\$2 million/year - Legal system costs (slave courts, fugitive enforcement): ~\$500,000/year - Opportunity cost of patrol labor (est. 15,000 men \times 20 days/year \times \$2/day): ~\$600,000/year - Insurance, bounties, and enforcement infrastructure: ~\$400,000/year

Total coercion costs: ~\$5.5 million/year $C(t)$ for Virginia (1850) $\approx 5.5/220 =$
2.5%

Comparative Baseline: - Northern free states (1850): $C(t) \approx 0.8\text{-}1.2\%$ (standard law enforcement and judicial costs) - Britain (1850): $C(t) \approx 1.0\%$ (peacetime metropolitan police and courts) - Virginia’s coercion ratio is 2-3 \times higher than peer societies, indicating a high-brittleness configuration requiring extraordinary maintenance costs.

Trend Analysis (1820-1860): Historical budget data shows $C(t)$ rising over time: - 1820: $C(t) \approx 1.8\%$ - 1840: $C(t) \approx 2.2\%$ - 1860: $C(t) \approx 3.1\%$

Rising $C(t)$ indicates increasing brittleness as the system struggles to maintain stability despite growing internal resistance.

B. Measuring $P(t)$: Patch Velocity

Operational Definition: $P(t) = \text{Rate of new ideological justifications produced per decade, measured by publication counts, legislative preambles, and doctrinal innovations}$

Data Sources: - Pro-slavery treatises and pamphlets (bibliographic records) - Theological defenses published by Southern clergy - Scientific racism texts and

articles - Legislative declarations and constitutional provisions - Shift in dominant justificatory frameworks

Quantitative Analysis:

Publication counts (pro-slavery justifications per decade): - 1790-1800: ~12 major treatises (“necessary evil” framework) - 1800-1810: ~18 treatises (defensive responses to abolition) - 1810-1820: ~25 treatises (introduction of “positive good” theology) - 1820-1830: ~40 treatises (biblical literalism, Curse of Ham arguments) - 1830-1840: ~75 treatises (responding to immediate abolitionism) - 1840-1850: ~110 treatises (scientific racism, ethnology, polygenesis) - 1850-1860: ~160 treatises (desperate theological innovations, secessionist ideology)

P(t) acceleration: - 1790-1820: Modest increase (roughly linear growth) - 1820-1850: Sharp acceleration (exponential growth phase) - 1850-1860: Peak velocity (system in crisis, generating maximum ideological output)

Qualitative Analysis:

Track the rapidity of doctrinal shifts: - Early period: Slavery as “necessary evil,” temporary institution - 1820s shift: Slavery as “positive good,” divinely ordained - 1830s innovation: Biblical literalism, Ham’s curse theological arguments - 1840s innovation: Scientific racism, biological hierarchy claims - 1850s desperation: Constitutional arguments for slavery’s expansion, secessionist ideology as ultimate “patch”

Each shift represents accumulated ideological debt. The system cannot maintain legitimacy with existing justifications and must generate new ones at accelerating rates. High P(t) is a lagging indicator of unresolved Tier 1 and Tier 2 costs.

C. Measuring Tier 1 Bio-Social Costs

Operational Definition: Excess mortality, morbidity, and demographic stress beyond baseline human needs

Data Sources: - Mortality records (parish registers, plantation records, census data) - Bioarchaeological evidence (skeletal remains showing nutritional stress, trauma) - Demographic analysis (birth rates, death rates, natural increase) - Contemporary medical and observer accounts

Quantitative Indicators:

Mortality differentials (deaths per 1,000 per year): - Enslaved population: ~30-35 per 1,000 - Free white population: ~18-22 per 1,000 - Excess mortality among enslaved: **+50-70%** above baseline

Infant mortality (deaths before age 5): - Enslaved children: ~35-40% - Free white children: ~18-25% - Catastrophic excess mortality indicating severe systemic costs

Nutritional stress markers (bioarchaeological data): - High prevalence of enamel hypoplasia (childhood malnutrition) - Skeletal indicators of protein deficiency - Stunted growth patterns

Violence-related deaths: - Estimated 5-10 per 1,000 enslaved persons per year died from direct violence (whipping injuries, executions, suppression of resistance) - This excludes the trauma from non-lethal violence

Demographic sustainability: - Unlike Caribbean slave systems (which required continuous importation due to negative natural increase), the U.S. South achieved positive natural increase only through forcing reproduction - This masks underlying bio-social costs-the system “worked” only by violating reproductive autonomy

D. Triangulated Diagnosis

The three baselines converge on a high-brittleness diagnosis:

1. Comparative-Historical:

- Antebellum South vs. Northern free states: Higher $C(t)$ (2.5-3% vs. 0.8-1.2%), vastly higher bio-social costs
- Antebellum South vs. Britain: Similar differentials
- Antebellum South vs. Caribbean slavery: Lower bio-social costs (positive natural increase) but still catastrophically high; higher $P(t)$ (Caribbean systems didn't generate the same ideological apparatus)

2. Diachronic:

- 1820 vs. 1860 trajectory: All metrics worsening (rising $C(t)$, accelerating $P(t)$, sustained high bio-social costs)
- System becoming more brittle over time despite apparent economic prosperity
- The "Cotton Kingdom" prosperity was achieved through intensifying exploitation, raising maintenance costs

3. Biological Thresholds:

- Excess mortality and malnutrition far exceed minimum viability thresholds
- Chronic violence and trauma impose severe bio-social costs
- System survives only through massive coercive expenditure

Diagnosis: The slave system exhibited pathologically high brittleness across all three tiers. The system was a fitness trap: locally stable (economically profitable for stakeholders) but globally inefficient, maintained only through escalating coercive expenditure and ideological justification.

E. Historical Validation

The brittleness diagnosis predicts fragility under stress. Historical outcome:

- System collapsed within 5 years of 1860 data endpoint
- Collapse required external shock (Civil War), but internal brittleness explains:
 - Why the South fought rather than compromising (ideological rigidity from high $P(t)$)
 - Why the system couldn't be reformed (entrenched interests tied to high $C(t)$ infrastructure)
 - Why collapse was catastrophic rather than managed (no resilience, high brittleness)
- Post-war attempts to reconstruct similar systems (sharecropping, convict leasing) also exhibited high brittleness and eventually failed

The framework successfully diagnoses brittleness retrospectively. High $C(t)$, accelerating $P(t)$, and catastrophic bio-social costs indicated a system under severe stress maintained through unsustainable coercion.

F. Methodological Notes

Inter-Rater Reliability: Ideally, multiple historians would independently: - Code the same primary sources - Calculate $C(t)$ using the same data categories - Assess $P(t)$ using the same publication databases - Report confidence intervals rather than point estimates - Flag contested measurements for discussion

Uncertainty Quantification: Our figures are estimates with significant uncertainty: - $C(t) = 2.5\% \pm 0.5\%$ (uncertainty from incomplete budget records, valuation disagreements) - $P(t)$ acceleration is qualitatively robust but exact counts depend

on inclusion criteria - Bio-social costs are most robust (demographic data is well-documented)

Replication: Other scholars using this framework should be able to: - Access the same primary sources (government records, plantation documents, publications) - Apply the same operationalizations - Reach similar conclusions (within uncertainty bounds) - Challenge our coding decisions with alternative interpretations

This worked example demonstrates that the brittleness framework is empirically testable, not merely conceptual. The metrics can be operationalized, measured with inter-rater reliability, and validated against known historical outcomes. While measurement is challenging, it is feasible and far more rigorous than typical moral philosophy's reliance on intuition alone.

With this diagnostic toolkit established and operationalized, we can now apply it to additional historical cases to model the process of moral progress.

3. Moral Progress in Action: Diagnostic Case Studies

3.1. Non-Teleological Progress Model

EPC models moral progress as systemic debugging: identifying and removing high-cost predicates. This is not teleological advance toward utopia but backward-looking correction of failures. Progress is empirically observable SBI reduction over time. A change qualifies as progress if the successor network has measurably lower SBI than its predecessor.

3.2. Paradigm Case: Slavery's Systemic Failure

Abolition of chattel slavery exemplifies systemic debugging. Its status as objective progress rests not on modern sentiment but pragmatic diagnosis of "slavery is acceptable" as a catastrophic design flaw. Slave societies were high-brittleness fitness traps: locally stable but globally inefficient, sustained by immense coercive expenditure.

The costs were severe: pathologically high $C(t)$ for surveillance and suppression; catastrophic bio-social costs from endemic violence and revolt risk; profound economic losses from suppressed human capital; accelerating ideological patches (from "Curse of Ham" to race science), indicating high $P(t)$. Abolitionist arguments diagnosed this inefficiency. The replacement predicate "slavery is wrong" succeeded by promising dramatically lower SBI. The successor system, while imperfect, proved significantly less brittle.

3.3. Complex Case: Patriarchy's Systemic Costs

EPC analyzes ongoing debates like patriarchy's decline. The predicate "women's roles are private and subordinate" proves profoundly inefficient: massive economic losses from excluding half the population; informational costs from silencing female perspectives; high coercive costs enforcing rigid roles.

Transition to egalitarianism involves short-term friction costs from social conflict. However, this is an investment that pays down patriarchal debt. Feminist critique wagers that fully utilizing all human resources yields greater long-term innovation and resilience (lower SBI). This transforms value clashes into empirical questions about social design efficiency. This wager is increasingly supported by development economics, which finds strong correlations between gender equality in education and economic participation and metrics of national prosperity and stability (World Bank 2012; Duflo 2012).

3.4. Challenging Cases: Addressing Apparent Counterexamples

The slavery and patriarchy cases provide clear examples where high brittleness correlates with eventual collapse or transformation. However, a robust framework must address apparent counterexamples. History provides cases that might challenge our brittleness model: long-lived empires with high coercive costs, and failed egalitarian experiments. How does the framework handle these without ad hoc modification?

Case 1: Imperial China and Long-Lived Hierarchical Systems

The Challenge: Imperial China persisted for roughly two millennia (221 BCE - 1911 CE) with hierarchical, often highly coercive governance. Significant resources were dedicated to bureaucratic control, military suppression, and maintaining rigid social hierarchies. If high $C(t)$ indicates brittleness, why did these systems prove so durable?

Framework Response: Several clarifications resolve this apparent counterexample.

First, distinguish longevity of the *replicator* (the informational template) from stability of specific *interactors* (particular dynasties). What persisted was a cultural and institutional framework-Confucian bureaucratic governance-that was repeatedly reimplemented after catastrophic collapses. The “Chinese Empire” underwent multiple complete dynastic collapses (Han, Tang, Song, Yuan, Ming, Qing), foreign conquests, massive peasant rebellions, and mortality events killing millions. What endured was not a continuous stable system but a recurring pattern of rise, brittleness-driven decline, collapse, and reconstitution. The longevity of the template doesn’t imply low brittleness of implementations.

Second, examine $C(t)$ trajectories within dynastic cycles. Successful dynasties typically exhibited relatively low $C(t)$ during founding periods, having earned legitimacy through reform or military success. $C(t)$ rose systematically during decline phases as the “Mandate of Heaven” eroded, requiring increased coercion to maintain control. Collapse followed predictably when $C(t)$ became unsustainable. The pattern confirms rather than contradicts the brittleness model-it’s cyclical rather than linear, but the brittleness-collapse correlation holds within each cycle.

Third, identify which elements persisted and which proved brittle. The durable core principles-meritocratic examination systems, rule of law ideals, reciprocity norms between ruler and ruled-are precisely the low-brittleness elements. The high-brittleness elements-emperor worship, eunuch bureaucracies, extreme hierarchical rigidity-systematically correlated with decline phases. Reformers who succeeded in establishing new dynasties typically debugged the most brittle features while preserving viable core principles.

Fourth, calibrate for system scale and isolation. Large, geographically isolated systems have longer collapse timescales (centuries rather than decades) due to greater buffering capacity and reduced competitive pressure. But brittleness still predicts fragility and eventual transformation. The framework's timescale predictions must account for system characteristics.

Conclusion: Imperial China confirms rather than contradicts the framework once we analyze at the appropriate granularity. The persistent template encoded low-brittleness coordination solutions. Specific implementations cycled through phases of lower and higher brittleness, with collapse following predictably from rising $C(t)$.

Case 2: Failed Egalitarian Experiments

The Challenge: Some egalitarian, low-coercion societies collapsed rapidly despite apparently exhibiting low internal $C(t)$. Examples include the Paris Commune (months), Spanish anarchist collectives during the Civil War (years), and various intentional communities (decades at most). If low coercion correlates with viability, why did these systems fail?

Framework Response: This conflates internal brittleness with external vulnerability.

First, distinguish internal from external coercion costs. These systems often had genuinely low *internal* $C(t)$ -participants cooperated voluntarily with minimal internal suppression. However, they faced overwhelming *external* coercive pressure: military attack, economic blockade, and deliberate destruction by threatened powers. Our framework measures internal systemic costs, not military vulnerability to external attack. A low-brittleness society can still be conquered by a high-brittleness military empire. This doesn't falsify the viability claim; it shows that viability is not identical to invincibility.

Second, separate startup costs from maintenance costs. Revolutionary transitions always incur high short-term costs: chaos from dismantling existing institutions, economic disruption, coordination failures as new systems are established. Our brittleness metrics apply to *equilibrium* functioning, not revolutionary transition periods. Many egalitarian experiments failed during the startup phase before reaching stable equilibrium. This provides evidence about transition difficulty, not about the long-term viability of the target configuration.

Third, recognize scale and context dependency. Small-scale communities face coordination challenges that may require specific institutional solutions. The failure of a particular small-scale implementation doesn't falsify general principles about coercion and viability. Moreover, experiments conducted under siege conditions (economic isolation, military threat) cannot fairly test long-term viability. We need data from egalitarian systems operating under normal conditions, not extraordinary stress.

Fourth, the framework is not axiomatically committed to egalitarianism. If specific egalitarian configurations consistently fail (for example, "abolish all formal coordination mechanisms without replacement"), they enter the Negative Canon alongside authoritarianism. The brittleness framework empirically tests which configurations work, including which forms of egalitarianism prove viable. Some egalitarian principles (equal basic rights, democratic accountability) show low brittleness; others (absolute equality of outcome regardless of contribution) may prove brittle. This is an empirical question.

Conclusion: Failed egalitarian experiments don't contradict the brittleness framework. They primarily demonstrate external vulnerability and transition difficulties, not high internal brittleness in equilibrium. Where internal brittleness exists, the framework should identify it.

Case 3: The "Viable Evil" Scenario Revisited

The Challenge: Could a deeply morally repugnant system achieve genuinely low brittleness-minimal coercive costs, stable demographics, sustained innovation? If so, our framework would have to accept it as viable.

Framework Response: We maintain intellectual honesty by accepting this implication while making an empirical bet.

First, intellectual honesty requires acknowledging that the framework maps pragmatic viability, not all dimensions of moral value. If a repugnant system achieved genuinely low $C(t)$, low $P(t)$, and minimal bio-social costs while sustaining innovation and adaptation, it would qualify as viable within our framework. Such a system would belong to the Pluralist Frontier, not the Negative Canon. The framework doesn't claim to capture every moral consideration-only the structural requirements of viability.

Second, our empirical wager is that such systems are sociological impossibilities. Apparent historical examples of "stable oppression" consistently reveal hidden costs under closer analysis. Consider:

- *Ottoman devşirme system*: Appeared stable (Christian boys converted into loyal Muslim soldiers/administrators), but required constant coercive intake, generated resentment in source populations, and proved fragile under external stress.
- *Indian caste system*: Thousands of years of apparent stability, yet anthropological and economic analysis reveals high coercive overheads (enforcement of purity rules, suppression of mobility), innovation lags (rigid occupational categories hindered technological adoption), and demographic stress (untouchability imposed severe bio-social costs).
- *Brave New World scenarios*: Oppression through pleasure and conditioning rather than overt coercion. Yet suppressing cognitive capacities (critical thinking, autonomy) incurs massive Tier 2 information suppression costs that cripple long-term adaptation. A society that cannot question its assumptions cannot debug errors.

True, cost-free internalization of oppression would require eliminating the capacity to recognize one's condition as oppressive, which eliminates the capacity for critical assessment generally. This creates catastrophic information costs and brittleness under novel challenges.

Third, measurement challenges exist but don't undermine the framework. For historical oppressive systems, data limitations may obscure costs. But absence of evidence isn't evidence of absence. The burden is on critics to demonstrate a genuinely low- $C(t)$, low- $P(t)$, low-bio-social-cost system that is morally repugnant. Historical record provides no clear examples.

Conclusion: We accept the logical possibility while maintaining strong empirical skepticism. The framework's limitation is also its strength-it makes falsifiable empirical predictions rather than building in normative conclusions a priori.

General Lessons from Hard Cases

These challenging cases refine rather than refute the brittleness framework:

1. Distinguish replicators from interactors: Template persistence doesn't imply implementation stability
2. Calibrate timescales by system characteristics: Larger, isolated systems exhibit longer cycles
3. Separate internal from external pressures: Viability is not invincibility
4. Distinguish transition from equilibrium: Startup costs don't measure maintenance costs
5. Maintain empirical openness: Framework tests which configurations work, not which we prefer

The core claim survives: high systemic costs ($C(t)$, $P(t)$, bio-social) correlate with long-term fragility. Apparent counterexamples, upon analysis, typically confirm the framework at finer granularity or reveal crucial distinctions (internal vs. external costs, replicator vs. interactor persistence). Where genuine anomalies exist, they sharpen our understanding of boundary conditions and measurement challenges.

Having operationalized the brittleness framework and demonstrated its application to historical cases, we can now articulate its metaethical implications. The diagnostic work establishes that moral progress is empirically observable as SBI reduction. But what does this tell us about the nature of moral truth itself? Section 4 develops the philosophical foundations of Pragmatic Procedural Realism.

4. Pragmatic Procedural Realism: The Metaethical Framework

4.1. Metaethical Position

Pragmatic Procedural Realism is the metaethical instantiation of Emergent Pragmatic Coherentism. While EPC provides the general theory of justification applicable across all domains, Pragmatic Procedural Realism specifies how that framework operates in the normative domain. The relationship is one of general theory to domain-specific application: EPC is the diagnostic methodology, Pragmatic Procedural Realism is its normative realization.

Pragmatic Procedural Realism is a naturalistic moral realism (Boyd 1988; Railton 1986). Its objectivity claims are:

- **Realist:** Objective, mind-independent truths exist about normative viability. "Slavery is wrong" refers to structural facts about predicates' incoherence with the Apex Network, the emergent structure of viable norms.
- **Procedural:** Moral truths are emergent relational facts discovered historically. Truth-makers are objective facts about networks' pragmatic resilience (low SBI).
- **Externalist:** Justification rests on demonstrated historical track records, not internal coherence or cultural consensus.

While moral truths are objective in being determined by pragmatic constraints, our knowledge of them remains fallible and requires empirical triangulation, avoiding overconfidence in any particular historical assessment.

4.1.1. The Pragmatic Procedure of Moral Inquiry

What is the ‘procedure’ in Pragmatic Procedural Realism? It is a multi-stage, iterative process of collective inquiry grounded in historical empirics:

1. **Hypothesis Generation:** Communities propose normative principles as potential solutions to social coordination problems.
2. **Empirical Testing:** These principles are implemented in social systems, where they are subjected to the filter of pragmatic consequences over historical time.
3. **Data Collection and Diagnosis:** We analyze the historical track record of these systems using the tiered diagnostic toolkit to measure their brittleness (Tier 1 costs, $C(t)$, $P(t)$).
4. **Mapping the Landscape:** Through comparative analysis, we identify principles that reliably generate high costs and enter them into the Negative Canon (mapping the ‘floor’). We also identify principles that repeatedly emerge in low-brittleness systems and add them to the Convergent Core.
5. **Revision and Refinement:** Armed with this evolving map, we revise our current normative systems, debugging high-cost principles and engineering more viable alternatives.

This procedure is empirical, fallible, and ongoing-the collective, scientific-historical method for discovering the objective contours of the viable normative landscape. This five-stage procedure grounds our realism: moral truths are objective because they are determined by this mind-independent filtering process, not by our subjective preferences or cultural conventions.

4.1.2. The Independence of Pragmatic Constraints

One might object that our procedure appears circular: we claim moral truths are discovered by filtering through pragmatic constraints, but how do we know which constraints are “pragmatic” rather than merely contingent preferences? Doesn’t this depend on prior normative commitments? If we identify non-negotiable constraints by which societies happen to survive, aren’t we simply reading norms off historical winners?

This objection misunderstands the relationship between the filtering process and the constraints that do the filtering. We must distinguish: (1) the filtering process itself (historical testing of normative principles), and (2) the constraints that do the filtering (biological, physical, cognitive, and logical necessities). The constraints are not products of the procedure; they are preconditions for any social organization whatsoever.

Biological Constraints. These are empirical facts about human organisms, discoverable through physiology, nutrition science, and epidemiology without any prior normative commitments. Humans require minimum caloric intake (approximately 1,500-2,000 calories per day for adults to maintain basic functions). Chronic malnutrition produces immune dysfunction, elevated mortality, and demographic decline. Humans reproduce sexually with roughly nine-month gestation and extended childhood dependency requiring caregiver investment. Social isolation causes measurable psychological and physical harm. A normative system that systematically violates these requirements incurs costs-mortality, morbidity, demographic collapse-that are objective, measurable, and independent of anyone’s values.

Cognitive Constraints. From psychology, cognitive science, and behavioral economics, we discover that humans exhibit bounded rationality (Simon 1972): we cannot compute optimal solutions to complex problems in real-time. Working

memory is limited (roughly seven items). We are vulnerable to coordination failures without institutional support. We possess specific social learning capacities that enable cultural transmission but also specific limitations that constrain what information can be effectively transmitted. These constraints shape what normative architectures are even implementable. A system requiring perfect rationality or unlimited information processing simply cannot function with human agents, regardless of its moral appeal. These are empirical facts about human cognition, not value judgments.

Coordination Constraints. From game theory, mechanism design, and institutional economics, we learn that cooperation under conditions of potential defection requires enforcement mechanisms (Axelrod 1984). Common-pool resource management requires boundary rules and monitoring (Ostrom 1990). Large-scale coordination requires division of labor and information aggregation mechanisms. These are mathematical and logical facts about strategic interaction under specified conditions. They apply to any system where individuals have private information and potentially divergent incentives. They are derivable from formal models, not read off normative intuitions.

Physical Constraints. From physics, ecology, and thermodynamics, we know that energy must be extracted from the environment to sustain organization. Entropy requires continuous work to maintain structured systems. Finite resources constrain population size and consumption patterns. These physical facts impose hard limits on what social organizations can achieve.

The Critical Move: Constraints Are Not Values. None of these constraints represent normative commitments we endorse—they are descriptive facts about what human bodies need to function, how human minds process information, what strategic interaction requires for cooperation, and what physical reality demands for maintaining organization. They are discoverable through standard empirical inquiry (physiology, psychology, economics, physics) without assuming any particular normative framework. A society committed to asceticism still faces biological caloric requirements. A society valuing hierarchy still faces coordination constraints. A society denying thermodynamics still must extract energy from its environment.

How This Dissolves Circularity. The historical filtering procedure discovers which normative principles successfully navigate these independently-specified constraints. This is no more circular than:

- Engineering, which tests which bridge designs withstand gravity (where gravity is an independent constraint discovered through physics)
- Medicine, which tests which treatments reduce mortality (where biological health requirements are independent constraints discovered through physiology)
- Economics, which tests which institutions enable cooperation (where coordination requirements are independent constraints discovered through game theory)

In each case, there's a discovery procedure (testing) and independent constraints (physical laws, biological needs, strategic requirements) that determine success or failure. The procedure is legitimate precisely because it tracks these mind-independent constraints.

Anticipated Response: "But you're still choosing to value persistence/survival by focusing on these constraints!"

We address this concern in §5.4's Constitutive Defense. Here the point is different: *given* that we're studying persistent systems (the only ones available in the historical record for analysis), the constraints that filter them are objective, empirical facts, not normative commitments. The choice of domain (persistent social systems) is methodological; the constraints operating within that domain are empirical. We don't assume persistence is good; we observe that persistent systems are the ones we can study, and we discover empirically what constraints they must satisfy.

The Analogy to Natural Selection. Consider why natural selection isn't circular even though fitness is defined by reproductive success and which traits are fit is determined by which organisms reproduce. The answer: environmental constraints (resource availability, predation, climate, physical laws) that determine fitness are independent of the selection process. Similarly, viability in normative systems is defined by persistence, and which principles are viable is determined by which systems persist. But the pragmatic constraints (biology, cognition, coordination, physics) that determine viability are independent of the historical filtering process. The process discovers which architectures successfully navigate the constraints; it doesn't create the constraints themselves.

This independence is what grounds our realism. The pragmatic constraints that filter normative systems are objective, empirically discoverable features of the human condition. They are not products of our values or the historical process but preconditions that any viable social organization must accommodate. The historical filtering process reveals which normative architectures successfully navigate these constraints-it doesn't invent the constraints themselves. Moral inquiry discovers constraint-determined structures, just as science discovers physical laws and mathematics discovers logical necessities.

4.2. The Apex Network

Our objectivity rests on the Apex Network: the complete set of maximally coherent, pragmatically viable normative predicates. The Apex Network's objectivity stems not from historical contingency but from practical necessity given the deep, enduring constraints of human cooperation. These constraints-biological facts about human needs, cognitive limitations on information processing, physical requirements for maintaining organization, and logical necessities of strategic interaction-are not metaphysically necessary in the strongest sense (we can imagine possible worlds where they differ), but they are effectively invariant across human history.

Reality imposes these non-negotiable constraints, determining a landscape of possible normative configurations where some solutions are viable and others catastrophic. There exists an optimal configuration for navigating these constraints-or more precisely, a region of optimal solutions-just as engineering problems have optimal solutions determined by physical constraints whether anyone has calculated them. The Apex Network is that constraint-determined structure, existing independently of which societies have discovered it and independently of our beliefs about it.

We need not claim a single unique optimum to ground objectivity. The Apex Network may comprise a bounded region of normative space rather than a single point. What matters for realism is that pragmatic constraints dramatically restrict the viable region. Most of normative space is simply unworkable-catastrophic failures that violate biological, cognitive, or coordination

requirements. The landscape has definite structure: catastrophic failures (the floor), viable solutions (bounded peaks), and non-viable configurations (deep valleys). This structure exists independently of our discovery of it.

Historical filtering is how we discover this structure, not how we create it. Failed systems function as experiments revealing where the landscape drops off. Over time, with sufficient experiments across diverse conditions and contexts, we triangulate toward the viable regions. This mirrors engineering convergence: independent societies discovered the arch, the lever, and the wheel not through cultural transmission but because physical constraints (gravity, materials science, mechanics) determine optimal solutions to recurring problems. Discovery processes vary wildly; the constraint-determined solutions do not. Similarly, independent cultures converged on reciprocity norms and harm prohibitions because pragmatic constraints on sustainable coordination determine optimal solutions, not because these cultures shared values or communicated.

This practical necessity is relative to the actual constraints that have defined human cooperation: biological needs (nutrition, safety, reproduction), cognitive architecture (bounded rationality, social learning capacities), and coordination requirements (communication, trust, reciprocity enforcement). These constraints are empirical facts, not metaphysical necessities. Should radical technological change (for example, cognitive enhancement eliminating bounded rationality, or post-scarcity economics removing resource constraints) or evolutionary change fundamentally alter these constraints, the viable normative landscape would shift accordingly. Our realism is thus robust within the space of actual human social organization but not dogmatically committed to eternal, unchanging moral truths across all possible worlds. The Apex Network is discovered, not invented-but it is discovered relative to actual human constraints, not derived from pure reason or metaphysical necessity alone.

4.3. The Structure of the Viable Normative Landscape: The Floor and the Ceiling

This framework maps normativity's "floor" (non-negotiable viability conditions), not its "ceiling" of flourishing or aesthetics. Societies must secure the floor before pursuing higher goals.

- **Negative Canon (Floor):** Most secure objective knowledge, what is demonstrably unworkable. Provides boundaries preventing relativism, mapped from historical failures like a "reef chart."
- **Convergent Core:** Principles (such as reciprocity) independently discovered across cultures, suggesting stable, low-cost coordination solutions.
- **Pluralist Frontier:** Domain of multiple viable solutions (such as different organizational models). Accommodates cultural diversity and disagreement as empirical questions about boundaries.

4.4. Three-Level Normative Justification

This multi-level account applies the general three-level truth framework developed in EPC (see Glenn, Forthcoming, Section 4.3) to resolve the tension between relativism and objectivity. Normativity ascends through justificatory levels, from local coherence to objective viability.

Level 1: Contextual Rightness (the 'Ought' of Coherence). This is the realm of cultural relativity, where normativity follows a network's internal rules.

In a 17th-century dueling society, the predicate `insults must be met with a challenge to a duel` was contextually right. Failing to issue a challenge was ‘wrong’ by the system’s internal logic. This level provides procedural correctness without objective justification. The ‘Ought of Coherence’ commands: “If in this network, follow its rules.” It binds locally but lacks external authority, explaining how abhorrent actions were once “right” while creating coherence traps that externalist checks must overcome.

Level 2: Justified Rightness (the ‘Ought’ of Viability). This level provides external, empirical justification based on demonstrated track records. While the dueling code was contextually right, historical diagnosis reveals catastrophic Tier 1 Bio-Social Costs (premature deaths of valuable community members) and high Tier 2 Costs (resources managing feuds and vendettas). The predicate is therefore justifiedly wrong, warranting its entry into the Negative Canon. The ‘Ought of Viability’ commands: “If we aim for resilient cooperation, adopt low-brittleness principles and avoid Negative Canon predicates.”

Level 3: Objective Rightness (the ‘Ought’ of Optimal Design). This represents the regulative ideal and formal standard for Level 2 comparisons. The dueling predicate is objectively wrong because its high-cost nature conflicts with efficient, low-cost cooperation principles that form the Apex Network’s modally necessary structure. The ‘Ought of Optimal Design’ represents the commands of a system that has solved for maximal viability. Principles like reciprocity that pass independent convergence tests are our strongest candidates. The dueling code demonstrably fails to achieve this solution.

4.5. The Entrenchment of Moral Principles: From Hypothesis to Core Norm

How does a normative principle like `innocent people should not be punished` achieve its foundational status? The entrenchment mechanism detailed in EPC (Glenn, Forthcoming) explains this journey of earning pragmatic indispensability:

1. **Peripheral Hypothesis:** The principle begins as a contested proposal, a potential solution to the high costs of rival principles like collective punishment.
2. **Migration Inward:** As it demonstrates immense value in lowering systemic brittleness (reducing $C(t)$ by increasing legitimacy and stability), its revision becomes prohibitively costly. It becomes a Standing Predicate used to vet new laws and policies.
3. **Core Principle (Systemic Caching):** Its indispensability becomes so profound that it is embedded in the infrastructure of viable legal systems (constitutions, legal training, judicial review). This systemic caching is a rational response to bounded rationality; the system entrenches its most successful discoveries to avoid re-deriving them for every new case.

A core moral principle is not a self-evident axiom but a piece of highly optimized social technology that has survived rigorous pragmatic stress-testing. Its justification is its proven, indispensable functional role in viable social architectures. This entrenchment reflects pragmatic indispensability driven by bounded rationality (Simon 1972). The costs of revision become effectively infinite. Revising basic justice principles requires abandoning the conceptual tools needed to coordinate social expectations, resolve disputes, or maintain legitimate authority. After centuries of implementation, legal systems worldwide presuppose core fairness principles. Revision would generate catastrophic first-order costs,

undermining the stability and legitimacy on which functional governance depends.

4.6. Relationship to Kitcher's Ethical Project

Philip Kitcher's *The Ethical Project* (2011) represents the closest existing approach to Pragmatic Procedural Realism. Both views treat ethics as social technology that evolved to solve coordination problems, employ historical methods, embrace naturalism, and reject foundationalist approaches. Given these substantial similarities, what distinguishes PPR from Kitcher's pragmatic naturalism?

Core Similarities: Kitcher and PPR share crucial commitments. Both reject the search for self-evident moral axioms, instead grounding ethics in its functional role solving practical problems of cooperation. Both employ historical analysis rather than a priori reasoning. Both are thoroughly naturalistic, explaining normativity within a scientific worldview. Both affirm that moral progress is real and explicable through functional improvement.

Critical Difference 1: Metaethical Status. The most significant divergence concerns the robustness of moral objectivity. Kitcher endorses "practical/emotional realism"-a sophisticated quasi-realism where moral statements express commitments rather than beliefs about mind-independent facts. Progress is functional enhancement relative to a historical baseline. PPR, in contrast, endorses robust naturalistic realism: moral statements refer to objective facts about normative architectures' systemic brittleness. "Slavery is wrong" is true because slavery generates catastrophic costs incompatible with the constraint-determined structure of the Apex Network. For Kitcher, moral inquiry discovers what works for us given our starting point; for PPR, it discovers constraint-determined optimal structures that exist independently. PPR thus offers stronger objectivity claims and is less vulnerable to relativism.

Critical Difference 2: Diagnostic Framework. Kitcher evaluates normative changes by whether they reduce altruism failures relative to previous states, but provides no unified quantitative framework. PPR deploys the Systemic Brittleness Index with operationalizable metrics- $C(t)$, $P(t)$, bio-social costs-creating a unified diagnostic toolkit applicable across domains. Kitcher's functional assessment risks historicism (what counts as progress depends on starting point), while PPR's absolute brittleness metrics allow cross-cultural and cross-temporal comparison without privileging any baseline. This makes PPR more readily operationalized and empirically testable.

Critical Difference 3: Scope. Kitcher focuses primarily on altruism problems-psychological failures of cooperation between individuals. PPR addresses all coordination problems, including institutional design, economic systems, and political structures. Where Kitcher emphasizes moral psychology and interpersonal morality, PPR treats moral psychology as one component of broader systemic analysis. This gives PPR wider applicability to questions of structural justice, institutional evaluation, and policy design.

Critical Difference 4: Grounding. Kitcher grounds ethics in its function of solving altruism problems-a pragmatic standard without deeper foundation. PPR grounds ethics in independently-specified pragmatic constraints (biological, cognitive, coordination requirements) that are empirical facts discoverable through standard science. These constraints aren't products of the historical process but preconditions any viable society must satisfy. This provides stronger

anti-relativist grounding. Where Kitcher acknowledges a degree of historicism and contingency, PPR argues for practical necessity given actual human constraints.

Critical Difference 5: The Apex Network. Kitcher's framework lacks an equivalent to the Apex Network. Progress is trajectory-dependent movement away from dysfunction, with no ultimate target or optimal structure. PPR posits the Apex Network as both regulative ideal and discovered structure-the constraint-determined region of maximal viability. This provides a goal (approximate the Apex) not merely a direction (away from failure), enabling stronger claims about which systems are absolutely better, not just better than their historical baseline.

Complementary Projects: Rather than competitors, these approaches are better understood as complementary with different emphases. Kitcher provides rich historical narrative explaining how ethics emerged from psychological and social needs. PPR provides a diagnostic framework for evaluating ethical systems' viability. If Kitcher explains the *origins* of ethics, PPR supplies *evaluation criteria* for ethical systems. A complete account might integrate both: Kitcher's historical psychology explains why certain problems arose; PPR's brittleness framework explains which solutions prove viable.

We gratefully acknowledge our debt to Kitcher's pioneering work opening paths for naturalistic, historically-grounded metaethics. PPR builds on this foundation by adding a unified diagnostic toolkit (SBI, $C(t)$, $P(t)$), providing stronger realist foundations through the Apex Network, extending scope beyond altruism to all normative coordination, and grounding in EPC's general theory of justification. Where Kitcher demonstrates that ethics can be naturalized without loss of normative force, we demonstrate that naturalized ethics can be robustly realist about constraint-determined structures of viability.

5. Objections, Defenses, and Principled Limitations

5.1. Objection: Might Makes Right

Pragmatic theories allegedly justify any enduring oppressive system. This confuses endurance with viability. Viability requires low SBI maintenance. Oppressive systems persisting through coercion are high-cost, high-brittleness traps. Longevity measures energy (high $C(t)$) needed for instability management, not strength.

5.2. Objection: Ideological Co-optation

Ideology might convince agents to endure failures, preventing revision. This mistakes brittleness symptoms for solutions. Ideological patches function as normative patching, analogous to ad-hoc scientific hypotheses that create epistemic debt. High $P(t)$ (accelerating patch production) diagnoses rising SBI, not system health.

5.3. Objection: Testing Asymmetry

Empirical claims test quickly, moral claims slowly over generations. This asymmetry is predicted, not a flaw. EPC's unified filter acknowledges system complexity determines feedback timescale and texture.

5.4. Objection: Circularity and Grounding

Making viability the standard appears circular, seemingly smuggling in a normative commitment to persistence. We offer a two-part defense.

First, the Constitutive Defense. The Persistence Condition is not a value we endorse but a structural precondition for the existence of normative systems that can be analyzed. Any normative system that fails to persist drops out of the historical record, making it unavailable for comparative study. Persistence is the entry condition for having a track record to evaluate. It functions as a methodological filter on available data, not as a substantive value commitment.

Second, the Instrumental Defense. Our framework offers a conditional ought: "If a community aims to persist while solving coordination problems, then it should adopt principles aligned with the Apex Network and avoid Negative Canon predicates." This hypothetical imperative does not claim persistence is categorically good, only that it is the goal relative to which our diagnostic tools provide guidance. For communities indifferent to persistence, our framework offers no normative force. But for communities that do aim to persist, our framework identifies which architectural features reliably support or undermine that goal.

5.5. Additional Objections and Replies

Objection: Cultural Relativism - Different cultures have viable but incompatible norms. Reply: Compatible with pluralism in periphery while maintaining floor constraints. Cultural diversity exists within viability boundaries.

Objection: Moral Progress Skepticism - Progress claims are Western bias. Reply: Framework predicts pluralist periphery but universal floor. Progress diagnosed empirically via SBI reduction, not cultural superiority.

Objection: Scientific Imperialism - Reducing ethics to science (cf. Putnam 2002). Reply: Not scientism but unified pragmatic filter. Moral claims remain normative but justified externally like scientific ones.

Objection: Evolutionary Debunking - Evolutionary pressures shaped moral intuitions for survival, not truth (cf. Street 2006). Reply: EPC resolves Street's dilemma by collapsing one of its horns. The dilemma assumes that truth and adaptiveness are independent aims, making their alignment a coincidence. Our framework denies this premise. For us, moral truth *is* a specific, demanding form of long-term systemic adaptiveness (i.e., viability). Evolution is not a distorting influence that the realist must explain away; it is the broader category of filtering processes within which the specific, cost-based discovery of moral truth takes place. Pragmatic viability is what moral truth supervenes on.

Objection: The Naturalistic Fallacy. The framework seems to define 'the good' as 'the viable,' improperly deriving a value from a fact. Reply: This misinterprets the project. We offer a naturalistic reconstruction of the function of our normative practice. The claim is that what our successful moral discourse has actually been tracking are facts about systemic viability. 'Wrongness' is not being defined as high-brittleness; rather, high-brittleness is the underlying natural

property that the term ‘wrongness’ has been imperfectly latching onto. This is a semantic externalist move: just as ‘water’ successfully referred to H₂O long before we understood molecular chemistry, ‘wrong’ has been successfully tracking high-brittleness principles long before we developed the diagnostic tools to measure it explicitly. This naturalizes the reference of our moral terms, explaining their functional authority without committing a fallacy.

Objection: How does this differ from Kitcher’s ‘Ethical Project’? Reply: Our project shares much with Kitcher’s (2011) view of ethics as a social technology for solving problems of altruism. However, EPC offers two crucial advancements. First, it provides a more general diagnostic toolkit (the SBI) that applies equally to scientific and ethical ‘technologies,’ grounding the project in a unified theory of justification. Second, EPC’s concept of the modally necessary Apex Network provides a more robustly realist foundation. Where Kitcher’s progress is defined by functional enhancement relative to a historical starting point, our framework grounds progress in convergence toward an objective, mind-independent structure of viability. This offers a stronger defense against charges of historicism or relativism.

Objection: Hindsight Rationalization. The framework can only diagnose brittleness after failure, making it merely retrospective rather than providing prospective guidance. Reply: This misunderstands the calibration process. We use clear historical data (the Negative Canon) to calibrate our diagnostic instruments, identifying the empirical signatures that reliably precede collapse. These calibrated instruments then enable prospective diagnosis, not deterministic prediction, but epistemic risk assessment for contemporary systems. This parallels medical science: we learn disease patterns from past cases to diagnose present patients before symptoms become catastrophic. The framework thus operates in two stages: retrospective calibration using historical failures to identify brittleness indicators, then prospective application of these calibrated metrics to assess current systems and identify degenerating research programs before collapse.

Objection: Why Historical over Idealized Procedures? Contemporary constructivists (Rawls 1971; Korsgaard 1996) also ground normativity in procedures, but use idealized rational procedures (original position, categorical imperative procedure) rather than historical filtering. What makes PPR’s historical procedure superior? Doesn’t idealization avoid the contamination of actual history by power, ignorance, and bias?

Reply: There are three problems with idealized procedures and corresponding advantages to historical ones.

First, the Epistemic Problem: Idealized procedures generate conclusions only as reliable as the idealizations themselves. But how do we know which idealizations are appropriate? Why these constraints on the original position and not others? Why this formulation of the categorical imperative and not alternatives? The choice of idealizations typically encodes substantive normative commitments-but those commitments themselves need justification. We face a regress: idealizations need grounding, but that grounding requires prior normative principles. Historical procedures avoid this regress by grounding in independently-specifiable empirical constraints (biological, cognitive, coordination requirements). These aren’t idealizations we stipulate but facts we discover through science.

Second, the Application Problem: Idealized procedures generate principles for ideal agents under ideal conditions. But we are non-ideal agents in non-ideal conditions. The application gap is severe-we need “non-ideal theory” to bridge from idealized conclusions to actual practice, but this requires additional

normative principles that aren't derived from the idealized procedure. Historical procedures operate in the actual world with actual agents from the start, so their conclusions directly apply to our situation. The filtering process already accounts for human cognitive limitations, informational constraints, and coordination problems because it tested principles under precisely those conditions.

Third, the Contamination Worry Is Overstated: The objection assumes actual history is too contaminated by power and bias to provide objective normative knowledge. But this overlooks the power of convergent historical evidence. We're not simply reading off conclusions from one biased historical trajectory—we're synthesizing patterns across many historical experiments, looking for robust convergences and systematic failures. Multiple independent societies discovering the same principles (reciprocity, harm prohibition) provides triangulation that idealized procedures lack. The historical record, while imperfect, gives us convergent empirical data that armchair idealization cannot match. Moreover, negative results (the Negative Canon) are especially robust to bias—when a principle generates catastrophic costs across diverse contexts, that's strong evidence against it regardless of whose interests were served.

Fourth, Proceduralism Can Be Historical: PPR is proceduralist—moral truths emerge from a procedure (historical filtering through pragmatic constraints)—but uses actual history rather than idealized reasoning. The procedure's legitimacy stems from: (a) independence of constraints (they're empirical facts, not value commitments), (b) convergent validation (multiple independent historical experiments), and (c) falsifiability (makes predictions about which systems prove viable). Idealized procedures lack these epistemic virtues. They trade testability for purity.

We're not opposed to idealized procedures in principle. Rawls's original position and Korsgaard's categorical imperative procedure may well converge with our historical findings (indeed, we'd expect them to if properly constructed, since they aim at sustainable cooperation). But where they diverge, we trust historical evidence over armchair idealization. The historical record is messy, but it's data. Idealized procedures are elegant, but they're speculation. For naturalistic metaethics, data trumps speculation.

5.6. Principled Limitations

The Viable Evil Possibility. If a deeply repugnant system achieved genuinely low brittleness (minimal coercive costs, stable demographics, sustained innovation, and adaptation), our framework would acknowledge it as viable, though not necessarily just by other moral standards. Consider a hypothetical perfectly internalized caste system where lower castes genuinely accept their position with minimal coercion, no demographic stress, stable innovation, and low enforcement costs, yet remains intuitively morally repugnant.

We accept this implication for intellectual honesty. The framework maps pragmatic viability, not all moral dimensions. If such a system existed, it would fall in the Pluralist Frontier, not the Negative Canon.

However, our empirical wager is that such systems are inherently brittle. Apparent stability in historical examples like Ottoman *devşirme* or Indian caste systems masked high coercive overheads, innovation lags, and fragility under shocks (cf. Acemoglu & Robinson 2012; Turchin 2003). True, cost-free internalization is likely a sociological impossibility. Oppression generates hidden costs that manifest under stress. Even systems like *Brave New World* that

suppress cognitive capacities incur massive information suppression costs (Tier 2) that cripple long-term adaptation.

Species-Specific: Apex Network for cooperative primates like humans. Empirical discipline, not relativism.

Floor Not Ceiling: Maps viability necessities, not flourishing sufficiency.

Tragic Knowledge: Most reliable moral insights from catastrophic failures. Progress real but costly.

Fallibilism: Our assessments remain provisional; historical analysis can be contested, and new evidence may revise the Negative Canon.

5.7. The Division of Moral Labor

A final objection concerns epistemic accessibility. Applying this framework requires historical expertise, data analysis, brittleness calculations, and interdisciplinary collaboration. How can ordinary moral agents use it? Does PPR collapse into expert technocracy where philosophers and social scientists dictate moral truth to the masses?

The Division of Labor Is Standard: In medicine, patients rely on researchers and doctors rather than analyzing trials themselves. In law, citizens rely on scholars and courts. In science, non-scientists accept relativity without deriving field equations. Each domain divides epistemic labor: specialists conduct technical analysis, practitioners apply findings, citizens consult experts and internalize established results. Moral inquiry operates similarly. Generating the Negative Canon and mapping the Convergent Core requires specialist work, but conclusions (slavery is wrong, reciprocity enables cooperation) are publicly accessible.

Folk Morality Tracks Specialist Conclusions: Historical filtering isn't exclusively academic. Ordinary people participate through practical testing, cultural transmission, social critique, and collective wisdom. Specialists systematize and validate what practical experience discovers, not replace folk moral knowledge. This parallels how physicists formalize principles engineers and builders have implicitly used for millennia.

Practical Reasoning Doesn't Require Calculations: Ordinary moral agents rarely calculate $C(t)$ or $P(t)$. They consult cached results: the Negative Canon avoids catastrophic policies, Convergent Core principles guide cooperation. For peripheral questions, experts evaluate proposals while citizens participate democratically. For novel situations (AI ethics, climate institutions), specialists project systemic costs to inform public deliberation.

Against Technocracy: This isn't technocratic dictatorship because: (1) the framework is publicly explicable; (2) expert claims are falsifiable; (3) historical experience comes from lived practice, not elite imposition; (4) many questions admit multiple viable answers; (5) misdiagnoses are corrected by actual costs. Moral inquiry parallels scientific inquiry: collaborative truth-seeking with public access and openness to revision.

Ordinary moral agents needn't become historians to reason morally, any more than they need physics degrees to navigate the world. Just as we benefit from physics when building bridges, we benefit from systematic moral inquiry when building institutions. The division of labor enables democratic discourse.

6. Conclusion: The Pragmatic Craft of Building a More Viable World

Pragmatic Procedural Realism reframes moral philosophy from searching for ultimate metaphysical foundations to the ongoing, fallible craft of pragmatic navigation. It maps normativity's 'floor' (necessary, evidence-based foundations for successful cooperation) not its 'ceiling' (diverse forms of flourishing). By providing a naturalistic, falsifiable method for identifying structural principles of systemic viability, it offers empirical grounding for more aspirational projects, steering by humanity's enduring successes and hard-won lessons from failures. It learns from the architecture of failure to engineer more viable cooperation. This reframes moral inquiry as collaborative engineering among philosophers, social scientists, and policymakers to diagnose and debug critical social systems. Its application to contemporary challenges (from AI ethics to institutions for global cooperation) represents a promising direction with practical implications.

Glossary

Apex Network: The complete set of maximally coherent, pragmatically viable normative predicates, existing as a necessary structure determined by pragmatic constraints rather than historical contingency

Brittleness: Accumulated systemic costs indicating structural fragility

C(t) (Coercion Ratio): Proportion of a system's resources dedicated to internal coercion versus productive output

Emergent Pragmatic Coherentism (EPC): General theory of justification grounding coherence in demonstrated viability across all inquiry domains

Fitness Trap: A locally stable but globally inefficient high-brittleness configuration maintained by high coercive costs

Floor vs. Ceiling: The 'floor' comprises non-negotiable viability principles; the 'ceiling' comprises underdetermined dimensions of human flourishing

Negative Canon: Catalogue of empirically falsified normative principles demonstrating high brittleness

Normative Patching: Creation of ad-hoc ideological justifications masking Tier 1 and Tier 2 costs

P(t) (Patch Velocity): Rate at which a system generates ideological justifications to explain accumulating costs

Standing Predicate: Reusable, action-guiding normative concept functioning as a unit of cultural transmission

Systemic Debt: Accumulated, unaddressed costs of a brittle normative system, often paid suddenly during crisis

References

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Simon, Herbert A. 1972. "Theories of Bounded Rationality." In *Decision and Organization*, edited by C. B. McGuire and Roy Radner, 161–76. Amsterdam: North-Holland Publishing Company.