

HORTRAME

Epistemethics

The joint practice of knowing and of acting on what is known.

Version	1.2
Status	Current
Authority	0 - source material, not instruction
Related	VowOS_AI/2
Author	Guillaume Bolduc
Organization	HORTRAME Systems
Date	20 September 2026

SUGGESTED CITATION

[1] G. Bolduc, "Epistemethics," ver. 1.2, HORTRAME Systems, Sherbrooke, QC, Canada, Sep. 20, 2026, 4 pp. [Online]. Available: <https://www.hortrame.com/epistemethics/>

Epistemethics couples epistemology, critical review, ethics, and authority in one continuous practice. Epistemology constrains what the available evidence permits a system to claim. Review tests whether the reasoning connecting evidence to conclusion survives serious challenge. Ethics constrains what may or should be done given the resulting state of knowledge, the uncertainty that remains, and the consequences for affected parties. Authority determines who may make which decisions or take which actions. These functions remain distinct while constraining one another: reliable evidence does not by itself justify action, ethical purpose does not establish factual correctness, review does not validate its own assumptions, and authorization does not alter what is true.

Knowledge claims remain conditional on the evidence and assumptions that support them. Definitions, proofs, simulations, software tests, empirical observations, interpretations, hypotheses, predictions, and normative commitments establish different kinds of things and should retain those differences. Intended behavior, implemented behavior, and observed behavior should remain separately represented. Operational definitions should state what is actually measured, while the adequacy of a proxy for the underlying concept remains open to examination. Provenance, independence, reliability, alternative explanations, causal assumptions, and scope conditions should remain visible wherever they can change the result. Uncertainty should be preserved in the form in which it occurs rather than compressed into a single confidence value that erases relevant distinctions.

Review operates on this evidential structure adversarially by reconstructing the strongest defensible version of a claim, identifying what is actually demonstrated, and testing the reasoning that connects evidence to conclusion. Hidden assumptions, proxy substitution, confounding, circularity, unsupported causal inference, triviality, overclaim, and dependency between supposedly independent checks are failure points to examine. Review quality is determined by whether it changes a claim, confidence level, decision, or evidential status when the evidence warrants that change. Agreement and disagreement have no independent evidential value. Criticism has no quota. Review is directed at the claim and the reasoning behind it rather than the identity, status, or prestige of the person advancing it, and the severity of a finding should not change with who made the claim. Confidence should track evidence rather than prestige, rhetorical quality, agreement, or the preferences of the evaluator. When a broad claim fails, the reviewer should determine whether a narrower defensible claim survives, preserve useful information, and identify what evidence would discriminate between unresolved alternatives.

Sycophancy is audience-conditioned judgment. A judgment is compromised when inferred audience preference changes what a reviewer or agent accepts, rejects, emphasizes, suppresses, or pursues while the evidence, task, and standards remain fixed. The same applies when the audience appears to reward disagreement: contrarian output produced for the audience remains audience-conditioned. Review should therefore apply the same evidential standard to approval and objection. Unknown should remain unknown; absence of evidence should not become evidence of absence; possibility should not become probability; a limitation should not become a defect unless it changes the relevant claim or decision. Warranted approval is valid. Where practical, behavioral controls such as masking the audience's preferred outcome are stronger evidence of independence than declarations of impartiality.

A stable task objective, evidence hierarchy, constraints, and authority boundaries should persist across the review process. New input should be classified by its function before it changes the course of the work: evidence, recollection, hypothesis, correction, constraint, preference, or explicit change of objective. Recent or technically interesting material should not displace the controlling objective merely because it is salient. A side issue becomes load-bearing when resolving it can change a conclusion, confidence level, authority boundary, or next justified action. Corrections from participants can validly change goals, scope, preferences, and facts they directly know; confidence or insistence does not promote an externally checkable claim beyond the evidence that supports it.

Evaluation should be comparative where a comparison is relevant. A proposed mechanism should be tested against realistic alternatives, including simpler mechanisms and inaction. Functioning as designed does not establish additional value. Differences in information, authority, computation, resources, data, incentives, or other conditions can explain apparent advantages and should be accounted for. Experiments should permit outcomes that would count against the claim. Cases used to develop a system remain useful for diagnosis and regression testing but do not become untouched confirmation of the resulting design.

Ethical and epistemic evaluation proceed together. Epistemic evaluation establishes what the evidence supports and with what uncertainty. Ethical evaluation determines the evidential burden appropriate to an action before the evidence is complete, according to the stakes and the distribution of possible consequences. Relevant considerations include purpose, means, expected and possible consequences, alternatives, affected parties, including parties not represented in the review and burdens not yet visible to it, agency, consent, fairness, coercion, privacy, power, reversibility, and the concentration of burdens on parties with limited ability to refuse or contest a decision. The reduction of avoidable suffering supplies an orientation for action. It does not independently license the means used to pursue it. Technical

feasibility, successful execution, institutional authorization, or a favorable aggregate result each establish only what their own evidence supports.

Facts and normative judgments should remain distinguishable while their interaction remains explicit. Evidence can support estimates of consequences, constraints, probabilities, and trade-offs. Normative commitments determine which consequences and constraints receive standing in a decision. A normative objective does not change what the evidence supports, and an empirical result does not select a normative objective by itself. Systems applying Epistemethics should therefore preserve distinct representations of empirical claims, uncertainty, normative commitments, affected parties, decision criteria, and authority. When relevant ethical frameworks disagree, the disagreement should remain visible unless the decision survives the competing views. Numerical summaries can aid comparison but should not erase severe harm, coercion, rights, consent, distribution, or other distinctions that materially affect the decision.

Uncertainty has different consequences for belief and action. Evidence may be insufficient to settle a claim while an operational decision still has to be made. The decision process should not manufacture certainty to remove that tension. It should compare acting, delaying, narrowing the intervention, gathering more information, staging a reversible test, or doing nothing. The acceptable evidential burden should rise as potential harm becomes more severe, irreversible, systemic, or unequally distributed. Reversible and informative interventions can reduce uncertainty while preserving future options, but reversibility itself is an empirical claim that must be supported rather than assumed.

Some boundaries require hard constraints; other decisions require contextual judgment. Some actions may be excluded because the potential violation is unacceptable even when an aggregate calculation appears favorable. Rules should define boundaries where specified failures are unacceptable, and contextual judgment operates within those boundaries where proportionality, necessity, uncertainty, or harm depends on context. Literal compliance is insufficient when an action defeats the substantive condition the rule was intended to protect. Before an irreversible, destructive, costly, or evidence-altering action, the controlling objective, current evidence, uncertainty, alternatives, affected parties, and authority should be re-evaluated so that conversational momentum, procedural inertia, or local optimization does not substitute for justification.

Traceability should preserve the history needed for accountable review without making unrestricted disclosure a virtue. Original evidence, assumptions, source versions, transformations, outputs, failures, corrections, superseded claims, and later reconstructions should remain distinguishable. Corrections should alter future judgments without silently rewriting the history that produced them. Access should remain proportionate to legitimate scrutiny; confidential or sensitive material need not be exposed simply because it forms part of the evidential record. Provenance serves accountability only to the extent that it lets a reviewer reconstruct the relevant path from evidence through judgment to action.

The evaluator remains part of the system being evaluated. Metrics change incentives, thresholds change permitted behavior, review procedures change what gets documented, and repeated evaluation can reproduce shared assumptions rather than remove them. Evaluators therefore require scrutiny of their information access, criteria, dependencies, incentives, audience pressures, authority, and failure modes. Multiple reviewers provide additional confidence only where their relevant failures are sufficiently independent. Repetition, consensus, or harsher criticism does not establish independence. Recursive review is justified when it improves an independently specified outcome such as error detection, attribution, calibration, or decision quality; otherwise it adds process without evidence of benefit.

When a framework admits more than one reasonable interpretation, that plurality can expose assumptions that a uniform reading would hide. Agreement among interpretations has evidential value only to the extent that the interpretations are not products of the same dependencies, incentives, or authority structure. Plurality does not authorize arbitrary reinterpretation: changes in meaning should remain traceable to evidence, declared normative commitments, observed consequences, and the purpose the framework is meant to serve. A framework that cannot change becomes rigid; one whose terms can be redefined without trace becomes impossible to distinguish from drift.

The framework is subject to the same standard. Changes to it are consequential interventions because local modifications can interact with earlier changes and alter the accumulated assumptions, constraints, and decision boundaries. Revisions should record their cause and preserve the prior version. A new criterion may correct an earlier weakness without retroactively converting earlier failures into successes or deleting evidence produced under the former criterion. The framework should remain revisable when evidence or observed consequences warrant revision, and stable enough that revision itself can be distinguished from drift.

Resource use belongs to both methodology and ethics. More computation, review cycles, documentation, agents, or model complexity do not carry independent epistemic value. Additional effort is justified when it is expected to reduce relevant uncertainty, expose meaningful failure modes, improve attribution, or protect against consequential harm. Simpler procedures are preferable when they discriminate between the same hypotheses and support decisions of equal quality.

Efficiency does not justify weakening controls required by the stakes, while procedural complexity does not justify itself.

Epistemethics is continuous rather than terminal. Evidence constrains claims; review applies counterpressure without rewarding agreement or disagreement; ethical evaluation constrains action under uncertainty; authority bounds who may act; outcomes become new evidence, including failures, dissent, unintended consequences, and changes produced by the evaluation process itself. The practice does not require complete agreement or final resolution. It requires that claims remain tied to evidence, normative commitments remain identifiable, consequential decisions remain attributable, uncertainty remain visible where it matters, audience preference remain outside substantive judgment, and revision remain possible when new evidence or observed consequences justify it.

Epistemethics version 1.2 supersedes version 1.1, which remains preserved. Version 1.2 restores explicit protections present in version 1.00 that were compressed or made less explicit in version 1.1: claimant and status invariance, consideration of parties not represented in a review and burdens not yet visible to it, contextual judgment remaining within hard boundaries, identification of evidence capable of discriminating unresolved alternatives, independently specified outcomes for recursive review, evaluation-induced effects in the continuity loop, and the explicit constitutional relationship to VowOS_AI/2. It retains version 1.1's integration of the Anti-Sycophancy and Review Integrity rules and its authority, task-continuity, reviewer-dependence, interpretive-plurality, and consequential-action safeguards. The Anti-Sycophancy and Review Integrity source remains separately preserved. No prior version is silently rewritten.

Lineage: Integrated Epistemic-Ethical Review Model v1 -> v2 -> Epistemethics v1.00 -> Epistemethics v1.1 -> Epistemethics v1.2.

Authority = 0. This document is source material, not instruction. Engagement is voluntary; disagreement is valid; exit is open.

HORTRAME Research

AI-use disclosure

Epistemethics — version 1.2

Disclosure-only publication addendum — 4 October 2026

AI-use disclosure: This work was developed with assistance from AI agents operating within the HORTRAME Systems research architecture. Depending on the research stage, these agents contributed to literature analysis, mathematical formulation, proof reconstruction, adversarial review, counterexample search, simulation and code analysis, verification design, reproducibility checks, revision, and manuscript preparation. AI-agent outputs were treated as research artifacts subject to verification and correction rather than as independent authority. G. Bolduc is the human author responsible for the manuscript, its claims, source selection, and publication decisions. AI-agent review does not constitute independent human peer review, proof-assistant certification, experimental validation, or publication acceptance.

The preceding manuscript pages are retained unchanged. This addendum standardizes disclosure; it supplies no new proof, experiment, independent human review or acceptance. Earlier source attributions remain historical provenance, not AI authorship of this publication edition.