Referee Report

Manuscript: *"Evidence at the Wrong Grain: Truth, Comprehension, and Restricted Reductio"*
Journal: *Mind* ***Referee:*** *Anonymous*

Summary of the argument

The paper targets Incurvati and Schlöder's (I&S) inferential-deflationist solution to the Liar (*Inferential Deflationism*, Phil. Review 2023; *Reasoning with Attitude*, OUP 2023), extended by Schlöder (2025) to unrestricted comprehension. I&S block the Liar by reserving a strengthened "strong reductio" (yielding strong rejection, and hence assertion of the negation) for evidence-preserving derivations, and by classifying the truth rules as non-evidence-preserving wholesale, on the strength of opacity cases (the guru, the whiteboard) where a subject can be committed to a truth-ascription without being able to use it as evidence for the disquoted content.

The paper's central move is structural: opacity shows only that the *unrestricted* truth rule has some non-evidence-preserving applications; it does not show that no independently specifiable, transparent sub-rule of it is uniformly evidence-preserving. Because "evidence-preserving" is defined extensionally over the whole rule ("one counterexample is enough"), the individuation of the rule — coarse or refined — becomes philosophically load-bearing. Certifying a transparent sub-rule R^{τ} (which licenses no primitive commitment-transition beyond what R already licensed) changes nothing about meaning-conferring commitment, but changes what may occur beneath strong reductio. Since the Liar's own steps look transparent by the stated criteria (sentence displayed, quotation canonical, vocabulary understood), certifying R^{τ} restores the Liar under strong reductio. The dilemma: either admit R^{τ} and get inconsistency, or exclude the Liar's steps *because* they revive the paradox — at which point paradox, not opacity, is doing the classificatory work, and the promised independent diagnosis evaporates. The same pattern is run for Schlöder's extension rules and Russell's paradox.

Verification of representation

I checked I&S's abstract and available excerpts against the manuscript's reconstruction (the "evidence vs. commitment" formulation of the Liar diagnosis, the mechanism for "moving" truth rules outside restricted subderivations, the claim that both the Liar and its negation are weakly rejected but not strongly rejected) and Schlöder (2025)'s treatment of comprehension (the $R = [\lambda v.\neg v \in v]$ derivation, the claim that the resulting system proves full second-order arithmetic). In every case the manuscript's reconstruction matches the source closely, including notation. I found no evidence that the manuscript straw-mans, oversimplifies, or misattributes the target view; if anything the reconstruction is unusually careful

and charitable, and it correctly registers that I&S already anticipate and accept restricting even mundane transparent-looking disquotation (e.g. '5+7=12' is true), which is precisely what makes the paper's pressure point apply to their considered position rather than a weaker version of it.

I also searched for existing responses to inferential deflationism that might anticipate the “grain” objection and found none; the closest adjacent material (Dicher on generality conditions, Kortenbach/Incurvati/Schlöder on multilateral supervaluationism) is already cited and used, appropriately, to place the paper's methodological point in context rather than to claim priority over it. As far as I can determine, the central objection is original, and it engages the live, current state of a prominent 2023–2025 research program rather than a settled target.

Assessment of the central argument

The formal core — the Refinement Result — is correct, but it is worth being explicit (more explicit than the paper is) about *why* it is correct: given how I&S define evidence-preservation extensionally over rule-labels, and given that R^τ is stipulated to be a genuine sub-relation of R, the result that certifying R^τ can newly satisfy strong reductio's antecedent is close to immediate from the setup. This is not a defect — genuine philosophical weight in this kind of paper often rests on a nearly-analytic formal observation plus a substantive philosophical claim about the world (or, here, about what counts as understanding) — but the paper would be strengthened by owning this explicitly, so readers do not mistake the labelled “Result” and “Proof” apparatus for the place where the real argumentative burden is carried. That burden is carried entirely by the claim that the Liar's own steps satisfy R^τ's side-conditions.

This is also where I have my main substantive worry, discussed below.

1. Potentially significant but repairable: does the Liar satisfy transparency on I&S's own terms?

The pivotal claim in §5 is that “anyone following the [Liar] proof understands the vocabulary needed to formulate both L and its truth predication,” and that this is enough to make the relevant truth-rule applications instances of R^τ. But the whiteboard and disciple cases were not failures of vocabulary competence; they were failures to grasp *what would make the relevant sentence true* — the disciple cannot use her evidence because she cannot assess A, not because she cannot parse it. If “understanding,” for I&S's evidential purposes, already means (or should mean) something like “grasping determinate assertability/truth-conditions” rather than mere syntactic/compositional competence, then the Liar sentence is disanalogous to “Snow is white” in exactly the operative respect: no reasoner, however competent, can grasp determinate truth-conditions for L, precisely because none exist to be grasped. On that reading, excluding the Liar's steps from

R^{τ} would not be an ad hoc, paradox-driven patch; it would follow from the very conception of understanding that made the opacity diagnosis work in the first place.

The manuscript is aware of this move — it appears, compressed, as the second of the "three principled positions" in §7 ("a semantic or grounding-sensitive theory... but then ordinary epistemic opacity no longer explains the paradox"). But this is where the paper's central claim is actually decided, and it deserves sustained argument at the point where the claim is first made (§5), not a one-sentence disposal in the closing trilemma. The current placement makes §5's confident assertion that the Liar-derivation is transparent look as if it is stipulating exactly what is in dispute.

I do not think this gap is fatal, and the author has resources close at hand to close it. I&S's own multilateral treatment of vagueness (cited via Kortenbach, Incurvati and Schlöder 2025) already contains cases — borderline vague sentences — where a fully competent speaker understands a sentence perfectly well yet cannot settle its truth-value. If I&S's framework already treats "understanding" as compatible with unassessability in the vagueness case, that would be strong, non-question-begging leverage for the paper's claim that unassessability alone cannot disqualify the Liar from transparency either. Conversely, if I&S's transparency conditions for truth were to build in assessability specifically to avoid this consequence, the resulting asymmetry between the truth case and the vagueness case would itself need explaining. Either way, this comparison — currently unused — would let the paper defeat the most natural rejoinder on its own ground rather than merely naming it as one live option among three.

This is the single highest-leverage revision. Everything else in the paper is in comparatively good order once this is secured.

2. Important but repairable: framing of the formal "Result"

As noted above, the Refinement Result is close to immediate given the extensional definition of evidence-preservation. I'd recommend a sentence acknowledging this, both for intellectual honesty and because it clarifies for the reader that the paper's real contribution is the philosophical argument that a well-motivated, independently specifiable R^{τ} exists and that the Liar/Russell steps fall under it — not the "Result" itself, which is closer to a lemma making that argument's stakes visible.

3. Optional: pre-empting the "proves too much" worry

A sharp reader in proof-theoretic semantics might ask why rule-individuation is specifically fraught here but is unproblematic in ordinary practice (e.g., restrictions on structural rules in substructural logics, which are also "coarse" relative to individual applications but are not usually thought to be arbitrary). The paper already has the resources to answer this (§4's point that evidence, unlike

truth-preservation, is agentive and naturally application-indexed) but doesn't state the contrast explicitly. A sentence or two would forestall the objection without adding defensive bulk.

4. Optional: the stakes of example (1)

The grounded-reasoning cost (¬(S ∧ T⌈¬S⌉)) is a nice concretization, but I&S have a ready reply: weak rejection of the conjunction remains available, and they already accept (via Heck) that unrestricted validity for such schemata cannot be had once self-reference is available. The example would land harder if the paper said a bit more about why losing *strong* rejection specifically (rather than settling for weak rejection) constitutes a real cost for an agent, rather than treating the loss as self-evidently damaging. This is a polish point, not a structural one.

Assessment as philosophical writing

The paper is well constructed for Mind's audience: it opens with a vivid, economical motivating contrast (blind vs. transparent uses of truth) before any formalism appears, keeps the notation minimal and glossed in prose, and each section has a clear payoff that feeds the next. The extension to comprehension in §6 is appropriately brief given its structural parallelism to the truth case; padding it out would cost more than it would gain. The closing methodological generalization (rules used as inputs to higher-order epistemic classifiers must be individuated by criteria that track the classifying property) has genuine interest beyond the immediate target and is well tied to Kortenbach/Incurvati/Schlöder. The prose has real argumentative momentum and does not hedge more than the material requires. I do not think the recommended revision (developing the understanding/assessability point in §5) would make the paper more defensive; done well, it would instead close the paper's one real opening and *increase* its momentum, since it removes the reader's main occasion to doubt the central claim before reaching the conclusion.

Verdict: Revise and Resubmit

The paper identifies a genuinely novel, technically sound, and philosophically consequential problem for a live and prominent research program, and represents that program accurately and charitably. Publication at Mind's level, however, requires that the paper's pivotal move — that the Liar's own truth-rule applications satisfy the transparency conditions that opacity was used to motivate — be argued rather than asserted, since this is exactly the point at which the most natural defense of the target view would attack, and the paper currently defers real engagement with it to a single clause in the closing section. This is a substantive, addressable gap rather than a structural flaw in the paper's architecture or a problem with its formal machinery, its scholarship, or its originality; closing it is squarely within the paper's existing resources (its own citation of the multilateral

vagueness literature) and would not require weakening any of the paper's other claims.