Evidence at the Wrong Grain

Truth, Comprehension, and Restricted Reductio

Abstract

Some uses of truth are epistemically blind. A speaker can have excellent grounds for saying that a sentence is true while being unable to identify or understand the sentence itself. Inferential deflationists turn this fact into a solution to the Liar: truth rules transmit commitment, not evidence, so they may not be used within the evidence-preserving form of reductio that yields outright rejection. I argue that the epistemic fact and the proof-theoretic restriction do not line up. Blind uses show that unrestricted truth rules are not uniformly evidence-preserving. They do not show that ordinary, fully transparent disquotation fails to preserve evidence. Yet the proposed solution excludes transparent cases as well. Its verdict therefore depends on treating blind and transparent uses as instances of one coarse rule. Split the rule while preserving every commitment transition, and transparent disquotation re-enters reductio—and the Liar returns. Excluding it because the paradox returns makes paradox, not epistemic opacity, do the explanatory work. The same problem recurs in an inferentialist treatment of comprehension and Russell's paradox. The lesson is methodological: evidential restrictions must track inferential performances, not the accidental grain of a rulebook.

Keywords: truth; inferentialism; bilateralism; evidence; Liar paradox; Russell's paradox

1. A blind inference and a transparent one

Truth lets us endorse what we cannot conveniently repeat. I can say that everything in a long report is true without reciting it, or that what a speaker just said is true without remembering her exact words. More strikingly, I can endorse words I do not understand. A devoted but monolingual disciple might insist that everything her multilingual guru says is true, including pronouncements in a language the disciple does not speak. Truth is useful partly because it permits endorsement at this distance.

Truth is also used at no distance at all. Someone who understands *Snow is white* can move from *'Snow is white' is true* to *Snow is white*. The sentence is displayed, its quotation is canonical, and drawing the conclusion requires no discovery about which sentence is at issue. This is the humdrum use represented by the disquotational inference from $T\ulcorner A\urcorner$ to A. The disciple's endorsement and the humdrum inference exploit the same relationship between truth predication and assertion. Epistemically, they could hardly be more different.[1]

[1]The contrast between transparent attitudes toward a displayed content and non-transparent truth ascriptions involving quantification, names, or descriptions is familiar from Kalderon (1997). Shapiro (2003) makes the contrast central to the expressive role of truth.

Luca Incurvati and Julian J. Schlöder place this difference at the centre of *inferential deflationism*. Their theory separates transitions that preserve *commitment* from transitions that preserve *evidence*. A speaker who endorses her guru becomes committed to the guru's claims and their consequences even when she does not understand them. What she cannot do is use her endorsement to acquire evidence for a conclusion she is unable to formulate or assess. Truth rules can therefore preserve commitment without uniformly preserving evidence.[2]

That distinction promises a striking diagnosis of the Liar. The truth rules remain unrestricted and meaning-conferring, so the theory retains truth's power to express blind and compendious endorsement. But a form of reductio that yields outright rejection is reserved for evidence-preserving derivations. Because truth rules do not preserve evidence, they may not occur inside such a reductio. The Liar and its negation can then both be rejected in a weak sense without contradiction (Incurvati and Schlöder 2023a, 2023b, ch. 7).

The proposal is attractive because it appears to locate the fault in the Liar independently of the paradox. We already know from ordinary epistemic life that truth can carry commitments further than evidence can travel. The paradox merely brings that familiar difference into a proof. Or so the diagnosis suggests.

There is, however, a gap between the epistemic cases and the proof-theoretic restriction. The disciple shows that *some* applications of the truth rules do not transmit evidence. The solution to the Liar excludes *every* application of those rules from evidence-preserving reductio, including applications in which the sentence is displayed, its quotation is recognised, and the speaker understands both sides of the inference. Incurvati and Schlöder see no gap here. Evidence preservation, they insist, is a property of rules rather than particular applications. One counterexample is enough to show that a rule is not evidence-preserving; that some applications are impeccable is "neither here nor there" (2023a, 551).

One counterexample is indeed enough to defeat a claim of *uniform* evidence preservation. But a prior question remains: what makes a collection of applications one rule for the purpose of restricting reductio? The question cannot be dismissed as bookkeeping, because the proof theory is sensitive to the answer. We can retain an unrestricted truth rule as a commitment-preserving, meaning-conferring rule while separately certifying a transparent subclass of its applications as evidence-preserving. The certificate adds no primitive commitment transition: every certified transition was already licensed. It nevertheless changes which derivations may occur beneath reductio. In the truth case, that apparently innocuous change restores the Liar.

The central claim of this paper is that the evidential restriction is imposed at the wrong grain. Opacity shows that the unrestricted truth rules are not uniformly evidence-preserving. It does not show that no transparent restriction of those rules preserves evidence. If transparent applications

[2]Inferential deflationism belongs to a broader proof-theoretic and inferentialist tradition in which the use of an expression, rather than a prior denotation, explains its meaning; see Dummett (1991) and Brandom (1994). The present argument concerns the specifically multilateral, evidence-sensitive development of that tradition, not inferentialism in general.

are instead excluded because admitting them restores paradox, then the paradox—not the independent opacity cases—determines the evidential boundary. Inferential deflationism may still yield a coherent non-transparent theory of truth. What it no longer yields is the promised independent explanation of why the Liar derivation goes wrong.

I first reconstruct the point at which the restriction operates. I then isolate the exact lesson of the opacity cases and formulate an invariance requirement for evidentially restricted proof theories: evidential consequences should survive conservative refinements that preserve both primitive commitment transitions and the epistemic profile of their applications. Inferential deflationism violates this requirement. The violation returns the Liar, blocks compelling grounded reasoning, and reappears in Schlöder's treatment of unrestricted comprehension and Russell's paradox. The result is not an objection to multilateralism as such. On the contrary, multilateralism provides the expressive resources needed to see the problem.

2. Where the paradox is stopped

The relevant machinery is small. Incurvati and Schlöder represent several attitudes toward a sentence A. Write $+A$ for strongly asserting A, $-A$ for strongly rejecting it, and $\ominus A$ for rejecting it in a weaker sense: declining or refraining from committing oneself to A. Finally, $\bot$ marks the act of registering an incoherence. Strong assertion and strong rejection express opposed beliefs. Weak rejection need express no belief in the negation of the sentence rejected.[3]

The two rejections answer a problem pressed by Dickie (2010). One might reject *Homer wrote the Iliad* because Homer did not write it, but one might instead reject it because Homer never existed. Only the first ground also supports the corresponding negative assertion. Rejection *tout court* is therefore evidentially unspecific: evidence that warrants rejecting A need not warrant asserting $\neg A$. Marking strong rejection separately allows the negation rules to connect $-A$ with $+\neg A$ while leaving $\ominus A$ weaker (Incurvati and Schlöder, 2017).

The same distinction governs reductio. Suppose an assumption $+A$ leads to $\bot$. The derivation may show only that one ought not commit to A, not that one has evidence for $\neg A$. Ordinary Smileian reductio accordingly yields weak rejection:

> If an assumption $+A$ leads to $\bot$, discharge it and infer $\ominus A$.

To obtain strong rejection, Incurvati and Schlöder introduce a stronger rule:

> If an assumption $+A$ leads to $\bot$ by an evidence-preserving derivation, discharge it and infer $-A$.

[3] For the bilateral tradition in which assertion and rejection are equally basic, see Smiley (1996) and Rumfitt (2000). The move from bilateral to multilateral systems, and the proof-theoretic coordination of their attitudes, is developed by Incurvati and Schlöder (2017, 2022) and del Valle-Inclán and Schlöder (2023).

Call the latter rule *strong reductio*. Given the rules connecting strong rejection and negation, it plays the role of classical reductio in the assertoric fragment. Its restriction ensures that grounds hypothetically supplied for A could be transformed into grounds for incoherence and hence into grounds for $\neg A$. A route to $\bot$ that merely records incompatible commitments can support withdrawal without supporting disbelief.

This restriction relies on an agentive conception of inference associated with Prawitz. The point of an inference is to justify its conclusion: a subject who possesses grounds for the premises can use the inference to acquire grounds for the conclusion (Prawitz 2015, 71–73; 2024). Evidence preservation is therefore not simply truth preservation. Whether a subject can use an inference depends on whether she can grasp and perform it. An open derivation represents the same activity hypothetically: if grounds for its assumptions were supplied, its steps would transform those grounds into grounds for its conclusion. Strong reductio is legitimate only when the derivation under its assumption has this character.

The machinery becomes philosophically interesting when the language is extended with truth. Inferential deflationism takes the meaning of the truth predicate to be given partly by its relation to assertion. The positive truth rules are:

$$\frac{+A}{+T^{\ulcorner}A^{\urcorner}}\ (+T\text{-IN}) \qquad \frac{+T^{\ulcorner}A^{\urcorner}}{+A}\ (+T\text{-OUT}).$$

There are corresponding rules for strong rejection. Together they capture direct disquotation as well as truth's roles in indirect and compendious endorsement and opposition. Crucially, they are unrestricted at the level of commitment. If a speaker asserts that everything the guru says is true, she is committed to what the guru says even when she cannot identify or understand every claim. This is why the truth predicate cannot simply be limited to sentences already understood by a speaker. Its ability to bridge such gaps is part of its expressive point (Shapiro 2003; Incurvati and Schlöder 2023b, §7.3).

Unrestricted truth rules produce the Liar. Let L be a sentence materially equivalent to $\neg T^{\ulcorner}L^{\urcorner}$. Given classical negation, truth introduction takes $+L$ to $+T^{\ulcorner}L^{\urcorner}$, which is equivalent to $+\neg L$. In the other direction, $+\neg L$ yields $+T^{\ulcorner}L^{\urcorner}$, and truth elimination then yields $+L$. Thus:

$$+L \vdash +\neg L \qquad \text{and} \qquad +\neg L \vdash +L.$$

From the assumption $+L$, the first derivation produces both $+L$ and $+\neg L$, and hence $\bot$. If strong reductio applies, one obtains $-L$, and therefore $+\neg L$. The second derivation then produces $+L$, yielding contradiction.

Incurvati and Schlöder block the final movement. The derivations from L to $\neg L$ and back remain valid because they preserve commitment. They do not count as evidence-preserving because they use truth rules. Ordinary reductio therefore permits $\ominus L$ and $\ominus\neg L$: both the Liar and its

negation ought to be weakly rejected. Strong reductio cannot be applied, so neither rejection becomes assertion of the opposite sentence. The two weak rejections are jointly consistent.

The decisive restriction is now visible: a derivation containing any application of a truth rule cannot occur beneath strong reductio. This includes the move from *'5 + 7 = 12' is true* to *5 + 7 = 12*. Incurvati and Schlöder accept the consequence. Many ordinary arguments can be rearranged so that truth is used outside the restricted subderivation, and unrestricted transparency is in any event unavailable to a consistent classical theory (2023b, §7.4). The question is not whether the paradox forces a restriction somewhere. It is whether the epistemic cases explain this restriction—a ban imposed by rule name and at maximum generality.

3. What opacity establishes

Consider first Incurvati and Schlöder's whiteboard. Thomas writes one sentence on it each day and reliably writes truths. Someone who knows Thomas's habit can have excellent evidence for *The sentence on Thomas's whiteboard is true* without knowing what the sentence is. Suppose today's sentence is *It is raining*. The subject may have no information about the weather and may not even know that *It is raining* is the relevant sentence. She can justify the truth predication by citing Thomas's reliability, but she cannot thereby formulate, much less justify, the assertion *It is raining*. Truth elimination does not transmit evidence in this setting (Incurvati and Schlöder, 2023a, 548–50).

Shapiro's disciple makes the point more vivid. The disciple endorses every utterance of a guru who speaks both arithmetic and set theory, although the disciple speaks only the arithmetical language. The guru asserts a set-theoretic sentence A, together with bridge principles connecting it to an arithmetical sentence B. A logician informs the disciple that B follows. The disciple can recognise the arithmetical conclusion and become committed to it, perhaps even believing it on the logician's testimony. Yet she cannot use her evidence for *A is true* to obtain evidence for A, because she cannot understand A or grasp what would count as evidence for it. Truth predication can therefore be meaningful and inferentially consequential in the absence of competence with the sentence called true (Shapiro, 2003).

These are genuine counterexamples. Their force should not be weakened. But notice what makes them work. In the whiteboard case, the apparent inference

$$T(d) \ \therefore \ A$$

suppresses several epistemically significant transitions. The subject must identify the referent of d as $\ulcorner A \urcorner$, substitute the canonical quotation into the truth predication, and then disquote. She cannot perform the first two transitions. In the disciple case, the quotation may be available, but the subject cannot grasp A and so cannot perform an inference whose conclusion is A. Each

subject lacks something required to turn a distant endorsement into a grasped conclusion.

Nothing similar is missing when a competent speaker moves from *'Snow is white' is true* to *Snow is white*. The sentence is present, the quotation is recognised, and the conclusion is understood. The comparison does not prove that every case of pure disquotation preserves evidence. It does establish that the explanation of failure in the opaque cases has disappeared. The natural verdict is mixed: the general rule has applications that cannot be used to acquire evidence, while transparent applications have a direct claim to do exactly that.

Incurvati and Schlöder respond that one counterexample is enough. They define evidence preservation, like truth preservation, over a whole rule: a rule preserves evidence just when every application does. On this definition the unrestricted truth rule plainly fails. Let $I(R)$ be the class of applications of a rule R, and let $EP(i)$ mean that application i can transmit grounds in the required way. The whiteboard and disciple establish:

$$\neg\forall i \in I(R)\, EP(i).$$

That is sufficient for R not to be uniformly evidence-preserving. It does not establish:

$$\forall i \in I(R)\, \neg EP(i),$$

More importantly, it does not establish that $I(R)$ contains no proper and independently specifiable subclass all of whose applications preserve evidence. The gap is proof-theoretically consequential. Strong reductio is sensitive to which *rule* a proof step instantiates. Whether a transparent subclass appears as a rule in its own right can therefore change what the calculus permits.

4. Evidence at the wrong grain

The problem does not depend on truth. Imagine a proof system with two generic rules, one evidence-preserving and one not. Strong reductio may use the first but not the second. Now replace them with a single rule whose side condition is the disjunction of their respective side conditions. The merged rule licenses exactly the same premise-conclusion transitions. Yet because it has counterexamples to evidence preservation, none of its applications may now occur beneath strong reductio. A presentational merger has removed strong conclusions; splitting the rule restores them. Nothing about any subject's evidence has changed.

The example exposes a general problem for any proof theory that classifies derivations by the generic rules they contain. Three levels must be distinguished. An *application* is a transition actually made in a derivation. A *generic inference* groups applications of a certain form. A *presentation* groups generic inferences into the displayed rules of a calculus. Prawitz (2024) distinguishes

generic inferences from their applications, since the generic inference specifies, among other things, which assumptions may be discharged while an application records what is actually discharged. But generic inferences themselves can be presented more or less coarsely. Predicating validity of generic inferences does not settle which collection of applications constitutes the relevant generic inference.

Treat a rule R extensionally as a set $I(R)$ of licensed applications. A *refinement* of R is a family $R_1, \ldots, R_n$ such that

$$I(R) = I(R_1) \cup \cdots \cup I(R_n).$$

Refinement changes the classification of applications, not the local transitions licensed. In the case at hand, let R^τ comprise transparent applications and R^ω opaque ones. The distinction need not be syntactic. R^τ may carry epistemic side conditions: the subject understands the conclusion, recognises the relevant quotation or extension term, and can perform the transition as a means of acquiring evidence. The target account already appeals to these facts when explaining why opaque applications fail.[4]

Refinement need not replace the original rule. Inferential deflationists may retain R as the unrestricted, meaning-conferring rule that preserves commitment while recognising R^τ as an evidentially certified route. This addition is conservative at the level of primitive commitment transitions. Every transition licensed by R^τ was already licensed by R, so any derivation using R^τ can be transformed into a derivation of the same conclusion by relabelling that occurrence as R. The inferential role used to explain the expression's meaning remains unchanged.[5]

The key result is correspondingly simple.

> **Refinement Result.** If R^τ licenses only applications already licensed by R, adding R^τ as a separately certified rule introduces no primitive commitment transition. It can nevertheless introduce new strong-rejection consequences whenever strong reductio is sensitive to the evidential status of rule labels.

For the first claim, replace each occurrence of R^τ with the corresponding occurrence of R. For the second, take a derivation of $\bot$ from $+A$ whose only prohibited step is a transparent application of R. In the coarse presentation, the label R prevents strong reductio and yields at most $\ominus A$. In the

[4] Proof-theoretic semantics contains a related dispute about whether defining rules must satisfy a condition of maximal generality. Dicher (2017) argues that such a generality condition is neither mandatory nor well supported. The present argument needs no general verdict on that debate; it requires only that epistemically significant restrictions of a meaning-conferring rule remain legitimate candidates for evidence preservation.

[5] This local notion of conservativity should not be confused with every stronger notion bearing that name. In proof-theoretic semantics, conservativity often requires that adding a connective or its rules generate no new consequences in the old vocabulary; see Belnap (1962) and Dummett (1991, 218–20). The present refinement satisfies what the argument needs: it adds no primitive commitment transition and changes no derivation until a higher-order rule inspects the label attached to a step.

refined presentation, label that same transition R^{τ}. The commitment derivation is unchanged, but strong reductio now yields $-A$. If the system connects $-A$ with $+\neg A$, assertoric consequences change as well. Two presentations can therefore agree on every meaning-conferring transition while disagreeing over what may be strongly rejected.

That instability motivates a constraint on evidentially restricted proof systems:

> **Evidential-Grain Requirement.** When a proof restriction is justified by whether an inferential transition transmits evidence, its consequences should be invariant under refinements that preserve both primitive commitment transitions and the epistemic profile of their applications.

Without this requirement, evidential consequences depend on editorial decisions about the rule sheet. An evidence-preserving transition can be made unavailable beneath reductio by merging it with unsafe transitions and restored by separating it from them. A restriction that purports to track the transmission of grounds cannot be manipulable in this way.

Truth preservation provides no escape. A schematic rule is invalid if it has a counterexample, even when some instances preserve truth. But logicians freely recognise valid restrictions of invalid schemata. The invalidity of an unrestricted rule does not make every application of every valid subrule unusable in a proof. What is distinctive here is that a verdict about the coarse schema prohibits individual applications inside a higher-order rule. That prohibition requires more than a definition of uniform validity. It requires a reason to regard the coarse schema as the governing unit of evidential permission.

The agentive character of evidence makes the point sharper. Truth preservation abstracts from what any subject understands. Evidence preservation, on the target view, does not. The counterexamples succeed because one subject cannot identify a sentence and another cannot understand it. Those differences cannot matter when generating counterexamples and then become irrelevant when transparent applications are classified. A rule-wide universal tells us whether a schema is safe without qualification. It does not tell us which inferential performances can transmit grounds under a supposition.

Nor can meaning fix the grain. The unrestricted truth rules explain both direct and compendious endorsement. Restricting their *meaning-conferring range* to sentences already understood would sacrifice part of truth's function. The proposed refinement does no such thing. The unrestricted commitment rule remains and continues to explain distant endorsement. The transparent rule is an evidential certificate, not a substitute meaning rule. Inferential expressivism already distinguishes transitions that determine commitment from transitions that transmit evidence. The second classification need not reproduce the first one's grain.

The programme already exploits this flexibility. Incurvati and Schlöder rearrange many ordinary arguments so that a truth rule is applied in the global context rather than inside a restricted

subderivation (2023b, §7.4). Schlöder later introduces an *Absoluteness* metarule that recovers evidence-preserving inferences among extension claims even though the extension rules themselves are globally classified as non-preserving (2025, 7–8). Neither device establishes that transparent disquotation or comprehension preserves evidence. Both show that epistemically acceptable routes can be certified without abandoning unrestricted meaning-conferring rules.

The remaining question is substantive: do transparent subrules preserve evidence? Opacity does not answer it. The paradoxes do.

5. Disquotation under supposition

Consider transparent truth elimination:

$$\frac{+T^{\ulcorner}A^{\urcorner}}{+A}\ (+T\text{-OUT}^{\tau}),$$

where the subject understands A, recognises its canonical quotation, and can perform the displayed inference. There is a corresponding transparent truth-introduction rule. These side conditions are not chosen to protect a preferred derivation. They negate the features that make the whiteboard and disciple cases work. No denoting expression must be resolved, no unknown sentence recovered, no conclusion brought within a competence the subject lacks.

On the Prawitzian conception invoked by inferential deflationism, these rules preserve evidence. Suppose a subject has testimonial evidence for *'Snow is white' is true*, understands *Snow is white*, and recognises the quotation. She can cite the testimony in support of *Snow is white*. It would be perverse to regard it as supporting only the metalinguistic predication and not the content transparently predicated true. Conversely, someone with evidence that snow is white can use truth introduction to justify *'Snow is white' is true*. Neither direction requires the discovery of a hidden identity or the acquisition of a competence the subject lacks.

The certificate is not unrestricted transparency. Evidence can be acquired through an opaque description of a sentence the subject could understand but has not identified. The transparent rule covers the performance in which the content and its canonical truth predication are both before a competent subject. If that performance fails to transmit grounds, the explanation must differ from the one supplied by the guru and whiteboard.

The Liar uses exactly this kind of disquotation. The sentence L is displayed, its canonical quotation is explicit, and anyone following the proof understands the vocabulary needed to formulate both L and its truth predication. The derivation from $+L$ to $+\neg L$ employs transparent truth introduction; the derivation from $+\neg L$ to $+L$ employs transparent truth elimination. Certify those applications and strong reductio no longer blocks the paradox:

$$\begin{array}{c}[+L]\\ \vdots\,(+T\text{-IN}^{\tau})\\ +\neg L\\ \bot\end{array} \quad \Rightarrow \quad -L \quad \Rightarrow \quad +\neg L \quad \Rightarrow \quad +L \quad \Rightarrow \quad \bot.$$

The coarse presentation contains only unrestricted truth rules, classified as non-preserving because of opaque applications. The refined presentation retains them and adds transparent certificates. Both presentations license the same primitive, meaning-conferring commitment transitions. But the refined presentation admits the Liar steps beneath strong reductio and is inconsistent. The solution is therefore not invariant under a refinement motivated by its own explanation of why evidence transmission sometimes fails.

Incurvati and Schlöder explicitly consider restricting truth rules to pure disquotation and reject the proposal. The semantic paradoxes, they say, show that truth rules fail to preserve evidence even in pure applications (2023b, §7.3). The response is coherent. The Liar proves that transparent rules cannot combine with strong reductio and the classical laws governing negation. Something must give.

But the response abandons the promised explanation. The disciple and whiteboard show that distant endorsement does not uniformly transmit evidence. They do not reveal a defect in the pure Liar steps. Those steps are classified as non-preserving because allowing them beneath reductio produces paradox. At that point the Liar, rather than opacity, fixes the evidential boundary. The paradox may support this package just as it supports familiar paracomplete, paraconsistent, or non-transitive packages.[6] What inferential deflationism loses is an independent reason to select its diagnosis.

This is not a complaint about how the view is advertised. The argument from opacity was supposed to explain *why* reductio is misapplied in the Liar. Without it, "the derivation does not preserve evidence" redescribes the need to block the derivation somewhere. A genuine diagnosis needs a non-paradoxical criterion that identifies the pure Liar steps as defective while distinguishing them from ordinary transparent disquotation. Classification by rule name supplies a formally sufficient boundary. The opacity cases give no reason to think it is the epistemically correct one.

The price of the coarse boundary appears in grounded reasoning. Let S be an ordinary, fully understood sentence. Consider:

$$\neg\big(S \wedge T^{\ulcorner}\neg S^{\urcorner}\big). \tag{1}$$

[6]For representative alternatives, see Kripke (1975); Field (2008) on gap-based and paracomplete approaches; Priest (1979) and Beall (2009) on glut-based approaches; and Ripley (2012); Cobreros et al. (2013) on transparent truth with non-transitive consequence. The present claim is not that these theories are preferable. It is that opacity alone does not select the inferential-deflationist restriction over them.

This is one of the weak truth-theoretic principles highlighted by Heck (2012). Its ordinary instances are compelling even without unrestricted transparency. It cannot be both that snow is white and that *'Snow is not white' is true*. The natural proof assumes the conjunction, extracts S and the truth predication of $\neg S$, disquotes the latter, and registers incoherence. Strong reductio then supports the negation of the conjunction.

In the inferential-deflationist calculus, this proof contains a truth rule beneath strong reductio. It can justify weakly rejecting the conjunction, but not strongly rejecting it or asserting (1). The loss is not confined to an exotic self-referential sentence. Set S to $5 + 7 = 12$. A competent reasoner can transform any putative grounds for

$$5 + 7 = 12 \text{ and } \text{'}5 + 7 \neq 12\text{' is true}$$

into grounds for incoherence using conjunction elimination and canonical disquotation. The opacity cases supply no reason why this hypothetical transformation fails.

The procedure for moving truth rules outside restricted contexts does not recover this proof. That procedure works when the desired output of a truth rule does not depend on assumptions to be discharged: one assumes the output inside the restricted context and applies the truth rule only after returning to the global context. In the proof of (1), however, the truth predication is extracted from the conjunction assumed for reductio. Its disquoted output $\neg S$ depends on the very assumption to be discharged. Moving disquotation outside leaves no categorical truth premise to which it can apply.

Heck shows that principles like (1) cannot be retained without limit in a sufficiently expressive self-referential setting. His result is a reason not to demand unrestricted validity for the schema. It is no reason to regard the arithmetic instance as epistemically defective. On the contrary, it sharpens the dilemma. If grounded instances are withheld because the schema becomes paradoxical under self-reference, the restriction tracks the global paradox-generating potential of a principle, not whether this subject can use this inference to acquire evidence. That boundary may define a coherent theory of truth. It is not the agentive boundary described by the opacity cases.

The truth case leaves three possibilities. Admit transparent evidential subrules, and the Liar returns. Exclude them because the Liar returns, and paradox rather than opacity does the classificatory work. Or supply a further, independently motivated explanation of why competent canonical disquotation fails to transmit evidence under a supposition. The existing diagnosis takes the second route while presenting itself as if the opacity cases had established the third.

6. The same fault line in comprehension

The grain problem is not peculiar to truth or quotation. Schlöder (2025) extends the same strategy to unrestricted comprehension. Write εP for the extension of a concept P. The characteristic

extension rules move between an object's falling under a concept and its belonging to that concept's extension:

$$\frac{+Px}{+x \in \varepsilon P}\ (+\varepsilon\text{-IN}) \qquad \frac{+x \in \varepsilon P}{+Px}\ (+\varepsilon\text{-OUT}).$$

There are corresponding rules for strong rejection. Schlöder argues that the rules preserve commitment but not evidence. Excluding them from strong reductio blocks Russell's paradox while leaving them unrestricted at the commitment level. The resulting calculus is consistent and proves full second-order arithmetic. The ambition is substantial: recover classical arithmetic in a logic with unrestricted comprehension while giving Russell's paradox the same evidential diagnosis as the Liar.

The independent case for non-preservation again exploits opacity. A person may know that a cube belongs to the extension of *Platonic solid* without understanding the concept expressed by the predicate. Perhaps she knows only that the relevant extension comprises the tetrahedron, cube, octahedron, dodecahedron, and icosahedron. She can justify the extensional claim from the list. She cannot use that evidence to justify the intensional claim that the cube is a Platonic solid, understood as a regular convex polyhedron in three-dimensional space, because she does not know which properties constitute the concept. Nevertheless, the extensional assertion commits her to consequences of the intensional claim once those consequences are explained (Schlöder, 2025, 6–7).

Schlöder notes that natural language hides an important difference. *The cube is a Platonic solid* may use *Platonic solid* as shorthand for an extension or may express an intension. The two uses have different inferential roles: only the latter immediately supports the conclusion that the cube is convex. The motivating failure thus occurs when a subject moves from an extensional understanding to an intensional content she does not grasp. It shows that unrestricted extension rules are not uniform evidence transformers. It does not show that conversion fails for a subject who understands P, recognises εP as its extension, and has both sides transparently before her.

Here too, a transparent certificate adds no commitment transition. Every inference from Px to membership in the extension of P, or back, is already licensed by the unrestricted rule. The certificate matters only beneath strong reductio. Once it is admitted, Russell's paradox returns. Let

$$R = [\lambda v.\neg v \in v].$$

Then:

$$+\varepsilon R \in \varepsilon R \ \vdash\ +R(\varepsilon R) \ \vdash\ +\neg \varepsilon R \in \varepsilon R,$$

while in the other direction:

$$+\neg\varepsilon R \in \varepsilon R \ \vdash \ +R(\varepsilon R) \ \vdash \ +\varepsilon R \in \varepsilon R.$$

The concept R is displayed and constructed from familiar logical vocabulary. No disciple has failed to grasp which concept a guru expressed. If the extension conversions are certified as transparent, strong reductio turns the first derivation into rejection of self-membership and the second back into its assertion. The calculus becomes inconsistent for the same reason as in the truth case.

Schlöder's Absoluteness principle makes the issue especially clear. Suppose a relation between concepts can be grasped without understanding the concepts themselves—for example, that a conjunctive concept entails each conjunct. Absoluteness carries an evidence-preserving inference between concept claims over to an evidence-preserving inference between extension claims. A disciple may therefore infer from membership in the extension of a conjunctive concept to membership in the extension of a conjunct while remaining ignorant of the concepts' contents. Schlöder introduces the metarule because the blanket restriction on extension rules would otherwise withhold too much ordinary reasoning (2025, 7–8).

Absoluteness does not certify the direct movement from extension to intension used in Russell's paradox, so it does not create inconsistency. It confirms the structural point. The proof theory can distinguish epistemically acceptable routes involving extension vocabulary from unrestricted rules that determine commitment. Why should that selectivity stop before competent, transparent conversion? The extensional list shows that conversion can fail when competence is missing. It supplies no answer for Russell's displayed concept. Once again, the paradox itself supplies the pressure to treat every direct conversion as non-preserving.

Truth and comprehension instantiate one pattern. A meaning-conferring bridge rule supports opaque and transparent uses. Opacity establishes that the bridge is not a universally safe evidence transformer. The proof theory then bans every application of the bridge inside a higher-order rule. But the ban is not invariant under refinement, and a transparent refinement restores the paradox. An evidential diagnosis of this form needs a principled reason to classify at the coarse, paradox-blocking grain. Neither opacity argument provides one.

7. The burden of an evidential cure

The distinction between commitment and evidence survives the argument. The guru and whiteboard show that speakers can undertake inferential commitments they are not yet in a position to appreciate or use as evidence. Truth rules explain those commitments without being uniformly evidence-preserving. Extension talk can likewise carry commitments beyond a speaker's present grasp of a concept. These are important achievements of inferential expressivism.

What fails is the movement from those achievements to the restriction on strong reductio. The independent cases justify a negative verdict about unrestricted rules. They do not identify the proper evidential grain. At the level of primitive, meaning-conferring transitions, coarse and refined presentations agree. Once evidence-sensitive discharge is added, they diverge. The cure needs a reason why the meaning-conferring rule, rather than a transparent subrule, is the unit whose failures determine every permissible occurrence under reductio.

There are three principled positions. An application-sensitive theory can let evidential annotations track the subject's actual competence and route through a derivation; it must then explain why the transparent Liar and Russell steps fail without appealing only to their consequences. A semantic or grounding-sensitive theory can distinguish ordinary disquotation from paradoxical dependence, but then ordinary epistemic opacity no longer explains the paradox. Or the coarse syntactic ban can be retained as a formal policy. The resulting system is coherent, but its evidence-sensitive fragment is presentation-dependent and withholds strong grounded conclusions such as ordinary instances of (1). Each position is intelligible. None is the advertised inference from blind endorsement to a rule-wide evidential prohibition.

The methodological lesson extends beyond these cases. Kortenbach, Incurvati, and Schlöder (2025) show that multilateral supervaluationist systems can be classical at the levels of theorems and ordinary inferences while remaining weaker at meta- and higher inferential levels. The present problem arises at exactly such a level. Two systems agree on every local commitment transition while differing over which derivations may be embedded under reductio. Once epistemic properties of rules regulate embedding, equivalence at the base level is not enough. The evidential classification must itself be stable under harmless changes of presentation.

This is what the multilateral machinery establishes that mainstream semantic and epistemic resources do not establish alone. Epistemology can distinguish testimony whose content one understands from testimony one endorses blindly. Semantic theories can restrict transparency. But neither, by itself, displays the difference between a derivation that warrants only weak rejection and one that warrants strong rejection because of the rule labels beneath reductio. Inferential expressivism makes that difference explicit. It thereby makes its own burden explicit as well: if evidence is why a derivation is barred, the restriction must track evidence rather than the accidental grain of the rulebook.

A blind application shows that an unrestricted rule is too coarse to serve as a universal evidence transformer. It does not show that evidence should be tracked at that same coarse grain. Until inferential deflationism gives an independent reason to exclude transparent applications, opacity cannot bear the weight of its cure.

References

Jc Beall. *Spandrels of Truth*. Oxford University Press, Oxford, 2009.

Nuel D. Belnap. Tonk, plonk and plink. *Analysis*, 22(6):130–134, 1962.

Robert B. Brandom. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press, Cambridge, MA, 1994.

Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Reaching transparent truth. *Mind*, 122(488):841–866, 2013.

Pedro del Valle-Inclán and Julian J. Schlöder. Coordination and harmony in bilateral logic. *Mind*, 132(525):192–207, 2023.

Bogdan Dicher. On a generality condition in proof-theoretic semantics. *Theoria*, 83(4):394–418, 2017.

Imogen Dickie. Negation, anti-realism, and the denial defence. *Philosophical Studies*, 150:161–185, 2010.

Michael Dummett. *The Logical Basis of Metaphysics*. Harvard University Press, Cambridge, MA, 1991.

Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.

Richard Kimberly Heck. A liar paradox. *Thought*, 1:36–40, 2012.

Luca Incurvati and Julian J. Schlöder. Weak rejection. *Australasian Journal of Philosophy*, 95: 741–760, 2017.

Luca Incurvati and Julian J. Schlöder. Epistemic multilateral logic. *Review of Symbolic Logic*, 15: 505–536, 2022.

Luca Incurvati and Julian J. Schlöder. Inferential deflationism. *Philosophical Review*, 132:529–578, 2023a.

Luca Incurvati and Julian J. Schlöder. *Reasoning with Attitude: Foundations and Applications of Inferential Expressivism*. Oxford University Press, New York, 2023b.

Mark Eli Kalderon. The transparency of truth. *Mind*, 106(423):475–497, 1997.

Bas Kortenbach, Luca Incurvati, and Julian J. Schlöder. Multilateral supervaluationism and classicality. *Journal of Philosophical Logic*, 54:247–290, 2025.

Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975.

Dag Prawitz. Explaining deductive inference. In Heinrich Wansing, editor, *Dag Prawitz on Proofs and Meaning*, pages 65–100. Springer, Dordrecht, 2015.

Dag Prawitz. The aim and validity of inference and argument. *Theoria*, 90:515–527, 2024.

Graham Priest. The logic of paradox. *Journal of Philosophical Logic*, 8:219–241, 1979.

David Ripley. Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, 5(2):354–378, 2012.

Ian Rumfitt. "Yes" and "No". *Mind*, 109(436):781–823, 2000.

Julian J. Schlöder. Basic rules of arithmetic. *Australasian Journal of Philosophy*, 2025. doi: 10.1080/00048402.2025.2521739. Advance online publication.

Stewart Shapiro. The guru, the logician, and the deflationist: Truth and logical consequence. *Noûs*, 37:113–132, 2003.

Timothy Smiley. Rejection. *Analysis*, 56(1):1–9, 1996.