

GENERALISED REVENGE

Julien Murzi and Lorenzo Rossi

Since Saul Kripke’s influential work in the ’70s, the revisionary approach to semantic paradox, the idea that semantic paradoxes must be solved by weakening classical logic, has been increasingly popular. In this paper, we present a new revenge argument to the effect that the main revisionary approaches breed new paradoxes they are unable to block.

1. Introduction

Let λ be a sentence which says of itself that it is not true. On the plausible if naïve assumption that, for every sentence φ , φ and ‘ φ is true’ are in some sense equivalent, a little reflection shows that λ is true if and only if it isn’t—a contradiction. In classical logic, this entails any sentence, that is, the reasoning makes one’s theory trivial. This is the Liar Paradox. Because the existence of sentences such as λ can be proved from basic syntactic principles, it is often thought that there are only two main ways out of the problem: one can either give up naïve principles about ‘true’ and other semantic notions, or revise classical logic. Since Saul Kripke’s influential work in the ’70s, the latter *revisionary* option has been increasingly popular.¹ Authors such as Hartry Field have forcefully argued that the truth predicate plays a key expressive role in our cognitive lives—one that requires that φ and ‘ φ is true’ be intersubstitutable. In a slogan, truth must be *naïve*.² As a result, classical logic must be restricted on pain of triviality but—revisionary theorists argue—this is not too high a cost, since classical principles are restricted where and only where they create trouble.³

Different non-classical theories of truth offer different explanations of the failure of classical principles. For instance, sentences that do not satisfy all the principles of classical logic have been characterised as ‘paradoxical’ [Kripke 1975], ‘unstable’ [Zardini 2011], ‘indeterminate’ [McGee 1991; Field 2008], ‘glutty’ [Beall 2009], both ‘tolerantly assertible and deniable’ [Cobreros et al. 2013], and so on. In turn, these notions have been thought to give rise to specific *revenge arguments*: Liar-like reasonings aimed at showing that while restricting certain classical principles allows non-classical theories to express a naïve notion of truth (and perhaps other semantic notions), notions such as absolute indeterminacy can only be expressed in those theories on pain of triviality.⁴ Revisionary theorists have responded by rejecting the coherence of revenge-breeding notions. For instance, Field writes that a unified notion of indeterminacy is ‘ultimately unintelligible’ [Field 2008: 356]; similarly, Field, Jc Beall, and Graham Priest have all rejected the coherence of the notion of *Boolean negation*.⁵ More generally, revisionary theorists typically dismiss semantic revenge arguments, on the grounds that they assume (a non-instrumental reading of) classical semantics. However, revisionary theorists either reject classical semantics outright [Ripley 2013], or interpret it instrumentally,⁶ or argue that it should be no surprise that non-classical notions cannot be expressed from within a classical framework [Beall 2007a; Field 2008].

¹ See e.g. Priest [2006a]; Field [2008]; Beall [2009]; Zardini [2011]; Cobreros et al. [2012].

² See e.g. Field [2008: ch. 13]; Beall [2009: sec. 1.1].

³ See e.g. Priest [2006a: 221]; Beall [2009: 111-12]; Zardini [2011: 518-19]; Woods [2018].

⁴ See e.g. Gauker [2006: sec. 3]; Leitgeb [2007]; Rayo and Welch [2007]; Welch [2014].

⁵ See e.g. Priest [2006b: ch. 5]; Field [2008: sec. 21.1]; Beall [2009: ch. 3]. The Boolean negation $\neg\varphi$ of φ takes value 1 whenever φ has a value *other than* 1.

⁶ See e.g. Field [2008: 356], and Beall [2009: 39 and 57].

More recently, it has been argued that revisionary approaches validating the classical structural rules cannot express notions of *naïve validity* and that this fact should be taken to favour a *substructural approach*—one that restricts some of the classical structural rules (for discussion see e.g. Shapiro [2011a], Beall and Murzi [2013], Zardini [2014], Field [2017], Murzi and Rossi [2018a]). Substructural approaches can express naïve truth, Boolean negation, and naïve validity (Zardini [2011], Ripley [2013], Nicolai and Rossi [2018]). Moreover, they have been argued to be ‘surprisingly strong’ and to approximate ‘the simplicity and symmetry of classical logic to an extent unmatched by its naïve rivals’ [Zardini 2011: 512]. Indeed, David Ripley has argued in a number of papers that his favourite nontransitive logic of paradox just *is* classical logic.⁷

But are substructural approaches revenge immune? Is there a *general* revenge problem afflicting all kinds of revisionary approaches? In this paper, we present a new proof-theoretic revenge argument to the effect that the main revisionary approaches, structural and sub-structural alike, breed new paradoxes that they are unable to block. Our argument does not rely on semantic notions and, unlike existing revenge arguments, it applies in a uniform way to any minimally strong revisionary theory.

Our argument unfolds in two main stages. We start from the observation that current revisionary theories feature sentences such as $t = t$ that satisfy all the principles of classical logic in a given theory S , and sentences such as λ that satisfy such principles in S only on pain of triviality. We call sentences of the former kind *unparadoxical-in- S* and sentences of the latter kind *paradoxical-in- S* . We argue that these notions are perfectly intelligible, even by non-classical lights, and provide a general recipe for generating *revenge paradoxes* to the effect that the main revisionary theories can only be closed under naïve principles for paradoxicality and unparadoxicality on pain of triviality.

From a revisionary perspective, the most natural way out of the problem is to treat the new paradoxes in the same way as the paradoxes of truth, that is by further weakening the logic. Since our revenge paradoxes rely on very weak logical resources, the upshot is that the revisionary approach is much more radical than it is usually thought.

To be sure, a more conservative reaction to the paradoxes of paradoxicality and unparadoxicality would be to question the intelligibility of these notions, much in the same way as notions such as absolute indeterminacy have already been questioned. However, we don’t think this would ultimately do. For one thing, the distinction between paradoxical and unparadoxical sentences in our sense is a *simple fact* about revisionary theories—one that encodes a minimal lesson to be learned from the semantic paradoxes, namely that, if truth is naïve, sentences such as λ yield absurdity if reasoned with classically, while sentences such as $t = t$ don’t (see e.g. Zardini [2011: 499]). For another, the distinction plays a crucial role in the main revisionary approaches to semantic paradox: it allows revisionary theories to ‘recapture’ classical theories such as classical mathematics, even if their logic is non-classical.

The plan of the paper is as follows. Sections 2-3 introduce the Liar Paradox and its four main revisionary ways out. Section 4 offers a precise definition of classical recapture. Section 5 presents four new revenge paradoxes, which trivialise the approaches introduced in section 4. Sections 6-7 discuss the relevance of our results and address potential objections. Section 8 concludes. The proofs of our results are given in an Appendix.

⁷ See e.g. Ripley [2012], and Ripley [2013: 146].

2. The Liar Paradox

We begin with some technical preliminaries. Let \mathcal{L}_{Tr} be a first-order language with identity whose logical vocabulary includes \neg , \wedge , \vee , \rightarrow , \forall , and \exists . In addition, \mathcal{L}_{Tr} contains a propositional *absurdity* constant \perp , a propositional *logical truth* constant \top and a predicate Tr expressing truth. Terms and formulae of \mathcal{L}_{Tr} are defined as usual. Closed formulae are called ‘sentences’. We let t and s (possibly with indices) range over closed terms of \mathcal{L}_{Tr} , and use φ , ψ , and χ (possibly with indices) as schematic variables for sentences of \mathcal{L}_{Tr} .

We require that any theory we consider satisfies two further requirements:

- (i) There is a function $[\]$ such that for every sentence φ , $[\varphi]$ is a closed term. Informally, $[\]$ can be understood as a quote-name forming device, so that $[\varphi]$ is a name of φ .
- (ii) For every open formula $\varphi(x)$ there is a term t_φ such that $[\varphi(t_\varphi/x)]$ is t_φ , where $\varphi(t_\varphi/x)$ is the result of replacing every occurrence of x with t_φ in φ .

In order to satisfy (i) and (ii), a theory has to interpret a *modicum* of arithmetic or syntax theory. For simplicity, we only consider theories in which (i) and (ii) provably hold.

A *sequent* is an expression of the form $\Gamma \vdash \varphi$, where Γ is finite *multiset* of sentences.⁸ The multiset to the left of \vdash is the *antecedent* of a sequent; the sentence on the right of \vdash is its *consequent*. We now recall the rules of classical propositional logic (henceforth, CPL).⁹ Our axiomatisation is highly redundant, in order to simplify the definition of classical recapture to be given in section 4 below.¹⁰

$$\begin{array}{c}
 \frac{}{\varphi \vdash \varphi} \text{SRef} \quad \frac{\Gamma \vdash \chi}{\Gamma, \varphi \vdash \chi} \text{SWeak} \quad \frac{\Gamma, \varphi, \varphi \vdash \chi}{\Gamma, \varphi \vdash \chi} \text{SContr} \quad \frac{\Gamma \vdash \varphi \quad \Delta, \varphi \vdash \psi}{\Gamma, \Delta \vdash \psi} \text{Cut} \\
 \\
 \frac{\Gamma \vdash \varphi \quad \Delta \vdash \psi}{\Gamma, \Delta \vdash \varphi \wedge \psi} \wedge\text{-I} \quad \frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \varphi} \wedge\text{-E}_1 \quad \frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \psi} \wedge\text{-E}_2 \\
 \\
 \frac{\Gamma \vdash \varphi}{\Gamma \vdash \varphi \vee \psi} \vee\text{-I}_1 \quad \frac{\Gamma \vdash \psi}{\Gamma \vdash \varphi \vee \psi} \vee\text{-I}_2 \quad \frac{\Gamma \vdash \varphi \vee \psi \quad \Delta_0, \varphi \vdash \chi \quad \Delta_1, \psi \vdash \chi}{\Gamma, \Delta_0, \Delta_1 \vdash \chi} \vee\text{-E} \\
 \\
 \frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \varphi \rightarrow \psi} \rightarrow\text{-I} \quad \frac{\Gamma \vdash \varphi \quad \Delta \vdash \varphi \rightarrow \psi}{\Gamma, \Delta \vdash \psi} \rightarrow\text{-E} \\
 \\
 \frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \neg(\neg\varphi \vee \neg\psi)} \text{DM}_1 \quad \frac{\Gamma \vdash \varphi \vee \psi}{\Gamma \vdash \neg(\neg\varphi \wedge \neg\psi)} \text{DM}_2 \quad \frac{\Gamma \vdash \neg(\varphi \wedge \psi)}{\Gamma \vdash \neg\varphi \vee \neg\psi} \text{DM}_3 \quad \frac{\Gamma \vdash \neg(\varphi \vee \psi)}{\Gamma \vdash \neg\varphi \wedge \neg\psi} \text{DM}_4 \\
 \\
 \frac{\Gamma \vdash \neg\neg\varphi}{\Gamma \vdash \varphi} \neg\neg\text{-I/E} \quad \frac{\Gamma, \varphi \vdash \perp}{\Gamma \vdash \neg\varphi} \neg\text{-I} \quad \frac{\Gamma \vdash \varphi \quad \Delta \vdash \neg\varphi}{\Gamma, \Delta \vdash \perp} \neg\text{-E} \quad \frac{\Gamma \vdash \perp}{\Gamma \vdash \varphi} \perp\text{-E}
 \end{array}$$

⁸ A *multiset* is just like a set, except that repetitions count. We use $\{ \}$ as brackets for sets, and $[\]$ as brackets for multisets. Thus, $\{\varphi, \psi, \psi\}$ and $\{\varphi, \psi\}$ are the same set but $[\varphi, \psi, \psi]$ and $[\varphi, \psi]$ are distinct multisets. We omit brackets from multisets in sequents, e.g. writing $\varphi, \varphi \vdash \psi$ instead of $[\varphi, \varphi] \vdash \psi$.

⁹ This suffices for the purposes of this paper: the results of section 5 only require propositional logical rules. For simplicity, we have opted for a single-conclusion natural deduction calculus in sequent-style in which structural rules are explicitly formulated.

¹⁰ A double line indicates that a rule can be read in both directions.

The sequents over the horizontal bar of a rule are its *premisses*; the sequent below is its *conclusion*. A rule is an *inference* if its premisses are empty, and a *meta-inference* otherwise.

In keeping with revisionary orthodoxy, we assume a *naïve* view of truth, that is, that the truth predicate satisfies the following truth rules (for convenience, we assume both positive and negative forms):

$$\frac{\Gamma \vdash \varphi}{\Gamma \vdash \text{Tr}(\ulcorner \varphi \urcorner)} \text{Tr-I} \quad \frac{\Gamma \vdash \text{Tr}(\ulcorner \varphi \urcorner)}{\Gamma \vdash \varphi} \text{Tr-E} \quad \frac{\Gamma \vdash \neg \varphi}{\Gamma \vdash \neg \text{Tr}(\ulcorner \varphi \urcorner)} \neg\text{-Tr-I} \quad \frac{\Gamma \vdash \neg \text{Tr}(\ulcorner \varphi \urcorner)}{\Gamma \vdash \neg \varphi} \neg\text{-Tr-E}$$

Other forms of naïveté include the T-Schema

$$(\text{T-Schema}) \text{Tr}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi,$$

and *transparency*, viz. the intersubstitutivity *salva veritate* of $\text{Tr}(\ulcorner \varphi \urcorner)$ and φ in all non-opaque contexts.

We are now in a position to present the Liar Paradox. Given our assumptions on \mathcal{L}_{Tr} we can prove that there is a sentence λ identical to $\neg \text{Tr}(\ulcorner \lambda \urcorner)$, so that λ says of itself that it isn't true.¹¹ We may then reason thus. We first prove $\text{Tr}(\ulcorner \lambda \urcorner) \vdash \perp$:

$$\frac{\frac{\frac{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \text{Tr}(\ulcorner \lambda \urcorner)}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \text{Tr}(\ulcorner \lambda \urcorner)} \text{SRef} \quad \frac{\frac{\frac{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \text{Tr}(\ulcorner \lambda \urcorner)}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \text{Tr}(\ulcorner \lambda \urcorner)} \text{SRef} \quad \frac{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \neg \text{Tr}(\ulcorner \lambda \urcorner)}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \lambda} \text{Definition of } \lambda}}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \neg \text{Tr}(\ulcorner \lambda \urcorner)} \neg\text{-E}}{\text{Tr}(\ulcorner \lambda \urcorner), \text{Tr}(\ulcorner \lambda \urcorner) \vdash \perp} \neg\text{-E}}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \perp} \text{SContr}}$$

Call the above derivation D_0 . We then derive $\text{Tr}(\ulcorner \lambda \urcorner)$ from D_0 :

$$\frac{\frac{\frac{\mathcal{D}_0}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \perp}}{\vdash \neg \text{Tr}(\ulcorner \lambda \urcorner)} \neg\text{-I} \quad \text{Definition of } \lambda}{\vdash \lambda} \text{Definition of } \lambda}{\vdash \text{Tr}(\ulcorner \lambda \urcorner)} \text{Tr-I}}$$

Call this derivation D_1 . D_0 and D_1 can now be combined together to yield a proof of absurdity, courtesy of Cut:

$$\frac{\frac{\mathcal{D}_1}{\vdash \text{Tr}(\ulcorner \lambda \urcorner)} \quad \frac{\mathcal{D}_0}{\text{Tr}(\ulcorner \lambda \urcorner) \vdash \perp}}{\vdash \perp} \text{Cut}$$

Given \perp -E, it yields a proof of *any sentence* φ , thus trivialising any theory in which the paradox can be derived.¹²

¹¹ More precisely, λ is the sentence $\neg \text{Tr}(t_\lambda)$, where t_λ is a closed term such that $t_\lambda = [\neg \text{Tr}(t_\lambda)]$. However, in the theories we consider λ and $\neg \text{Tr}(t_\lambda)$ are always intersubstitutable, and we will therefore stick to this simpler formulation. The same goes for the other 'self-referential' sentences to be introduced later.

3. Four Revisionary Ways Out

If naïve semantic principles such as Tr-I and Tr-E are non-negotiable, as revisionary theorists typically maintain, then one must blame the logic in order to avoid non-triviality. To be sure, such a revision is not to be taken lightly, and there is no shortage of classical treatments.¹³ But, contemporary logical wisdom has it, these alternatives are dire, the naïve semantic principles *are* non-negotiable, and there might be independent reasons for weakening classical logic in the first place.

The Liar Paradox makes use of four main logical ingredients: \neg -I, \neg -E, SContr, and Cut. Each of these rules can be, and indeed has been, questioned.¹⁴ We briefly consider the corresponding four revisionary strategies, and introduce, for each such strategy, the most representative corresponding formal theory.

3.1 Paracomplete and Paraconsistent

The most popular revisionary approaches to paradoxes such as the Liar involve revising the classical theory of negation and the conditional, according to which \neg satisfies both \neg -I and \neg -E, and \rightarrow satisfies both \rightarrow -I and \rightarrow -E. According to *paracomplete* theorists, sentences such as λ are *gappy*: they either lack a semantic value, or they have an intermediate value between truth and falsity. According to *paraconsistent* theorists, sentences such as λ are *glutty*, that is, they are both true and false. We briefly review both approaches in turn.¹⁵

Paracomplete theorists typically advocate the so-called *strong Kleene logic* K3 (Kleene [1952: 332-40]), or some extension thereof. K3 is given by the rules of classical logic minus \neg -I and \rightarrow -I. As a consequence, the Law of Excluded Middle

$$\text{(LEM)} \quad \vdash \varphi \vee \neg\varphi$$

is not unrestrictedly valid either. We call K3TT the theory resulting from adding the naïve truth rules to a sufficiently expressive theory based on the logic K3.

Dually, paraconsistent theories are typically based on the logic LP, or some extension thereof (Asenjo [1966]; Priest [1979]). LP is given by the rules of classical logic minus \neg -E and \rightarrow -E. As a result, the Law of Non-contradiction

$$\text{(LNC)} \quad \varphi \wedge \neg\varphi \vdash \perp$$

must be given up. We call LPTT the theory resulting from adding the naïve truth rules to a sufficiently expressive theory based on the logic LP.

¹² Triviality can also be established, without making use of \perp -E, via Curry's Paradox. The paradox involves a sentence γ identical to $\text{Tr}([\gamma]) \rightarrow \psi$ (where ψ is any sentence). Given SRef, SContr, Cut, \rightarrow -I, and \rightarrow -E, a Liar-like argument allows one to 'prove' ψ .

¹³ Examples of classical hierarchical treatments include Tarski [1936]; Parsons [1974]; Glanzberg [2004]; Murzi and Rossi [2018b]; for classical, non-hierarchical approaches see e.g. Feferman [1991]; Leitgeb [2005]; Halbach [2011: sec. 19.3-19.5].

¹⁴ The revisionary literature to date has almost exclusively focused on theories validating SRef. While it is possible to devise a revenge paradox for SRef-free approaches along the lines of the ones to be developed in section 5, we don't give the argument here for reasons of space. For a recent proposal involving a restriction of SRef, see Nicolai and Rossi [2018].

¹⁵ Paracomplete theories have been developed in Kripke [1975]; Field [2008, 2013]; Halbach and Horsten [2006]; Horsten [2012]. For paraconsistent logics and their application to semantic paradoxes, see e.g. Asenjo [1966]; Priest [1979, 2006a]; Goodship [1996]; Beall [2009].

3.2 Substructural Approaches: Non-contractive and Non-transitive

We now turn to approaches which restrict the *structural rules* SContr and Cut. *Non-contractive* approaches advocate a restriction of SContr. That is, according to these approaches the fact that ψ follows from $[\varphi, \varphi]$ does not entail that ψ follows from $[\varphi]$ alone.¹⁶ Elia Zardini [2011] proves syntactic consistency for a transparent theory of truth whose underlying logic is a suitable strengthening of *multiplicative affine linear logic* (henceforth, MALL). MALL's propositional fragment is CPL without SContr and with \vee -E replaced by the following weaker version:

$$\frac{\Gamma \vdash \varphi \vee \psi \quad \Delta_0, \varphi \vdash \chi_0 \quad \Delta_1, \psi \vdash \chi_1}{\Gamma, \Delta_0, \Delta_1 \vdash \chi_0 \vee \chi_1} \vee\text{-E}_w.$$

We call the propositional fragment of Zardini's theory MALLTT, for a sufficiently expressive theory based on the logic MALL with transparent truth.¹⁷

Finally, *non-transitive* approaches recommend a restriction of Cut.¹⁸ In particular, Pablo Cobreros, Paul Egré, Robert van Rooij, and David Ripley have recently put forward a non-transitive theory based on the non-transitive logic ST, which is essentially classical logic, with all its theorems and inferences, but without the rules Cut, \rightarrow -E, \vee -E, and \neg -E. The theory, labelled STTT for *strict tolerant transparent truth*, allows for a uniform treatment of the semantic and indeed soritical paradoxes. For simplicity, we consider a sufficiently expressive theory of transparent truth, which we call STTT₀, given by a sub-logic of ST with the addition of the naïve truth rules. More precisely, the logic of STTT₀ is given by the rules of CPL minus Cut, \rightarrow -E, \vee -E, and \neg -E.

4. Classical Recapture

The four families of non-classical theories we have just introduced all share a common feature: despite their non-classicality, they have *fully classical* fragments. That is, all the theories presented in section 3 limit their restrictions to classical logic to *some* sentences. This is not only a basic fact about those theories; it also allows one to apply those theories to mathematics and science more generally. As it is sometimes said, non-classical theories can *recapture* classical reasoning when needed.¹⁹ For instance, Field sees himself as being engaged in the project of finding

a generalisation of classical logic that takes the classical rules to be appropriate for dealing with 'ordinary' predicates (such as those of standard mathematics and physics) but which allows only weaker rules when dealing with certain 'extraordinary' predicates [such as 'true']. [Field 2008: 7]

¹⁶ See e.g. Shapiro [2011a]; Mares and Paoli [2014]; Zardini [2011].

¹⁷ Zardini's full theory IKT^ω, as he calls it, includes some controversial infinitary rules for the quantifiers [Zardini 2011: 508]. The revenge paradox to be developed in section 5.5 applies not only to Zardini's full theory, but also a version of his theory in which the conjunction and disjunction operators are additive, i.e. governed by context-sharing rules, and the quantifiers are governed by standard, finitary rules (for discussion, see Zardini [2011: 509-10]).

¹⁸ See for instance Weir [2005]; Ripley [2012]; Cobreros et al. [2013].

¹⁹ See e.g. Priest [2006a: 221], Field [2008], Beall [2009: 111-12], Zardini [2011].

Classical logic is restricted where, and only where, it creates trouble.²⁰ To see how non-classical theories recapture classical theories, our starting point is a particularly simple way of characterising the classical fragment of K3TT, LPTT, MALLTT, STTT₀, and their extensions. Such theories all enjoy the following informal property:

(Classicality Principles) There are finitely many classically valid principles such that a sentence satisfies such principles only if it satisfies all classical principles.

We can then say that a theory recaptures classical logic if it is closed under weaker versions of classical rules which, whenever some extra conditions are satisfied, reduce to their classical counterparts. The following definition formally captures this idea.

Definition 4.1 (Classical recapture). Let S be a non-trivial theory. Then, S enjoys a *classical recapture property* if it is \mathfrak{P} -classical recapturing, for some classically valid principle \mathfrak{P} invalid in S . The following classical recapture properties correspond to the revisionary approaches reviewed in sections 3.1-3.2.

- S is *LEM-classical recapturing* if it is closed under the rules of CPL, where \rightarrow -I and \neg -I are replaced by the following weaker versions:

$$\frac{\Gamma, \varphi \vdash \psi}{\Gamma, \varphi \vee \neg\varphi \vdash \varphi \rightarrow \psi} \rightarrow\text{-I}_w \qquad \frac{\Gamma, \varphi \vdash \perp}{\Gamma, \varphi \vee \neg\varphi \vdash \neg\varphi} \neg\text{-I}_w$$

- S is *LNC-classical recapturing* if it is closed under the rules of CPL, where \rightarrow -E and \neg -E are replaced by the following weaker versions:²¹

$$\frac{\Gamma \vdash \varphi \quad \Delta \vdash \varphi \rightarrow \psi}{\Gamma, \Delta \vdash \psi \vee (\varphi \wedge \neg\varphi)} \rightarrow\text{-E}_w \qquad \frac{\Gamma \vdash \varphi \quad \Delta \vdash \neg\varphi}{\Gamma, \Delta \vdash \perp \vee (\varphi \wedge \neg\varphi)} \neg\text{-E}_w$$

- S is *LContr-classical recapturing*,

$$(\text{LContr}) \vdash \varphi \rightarrow (\varphi \wedge \varphi),$$

if it is closed under the rules of CPL, where SContr is replaced by the following weaker version:

$$\frac{\Gamma, \varphi, \varphi \vdash \psi}{\Gamma, \varphi, \varphi \rightarrow \varphi \wedge \varphi \vdash \psi} \text{SContr}_w$$

- S is *Cut-classical recapturing* if it is closed under the rules of CPL *minus* Cut, where \rightarrow -E, \neg -E, and \vee -E are replaced by the following weaker versions:

²⁰ Field's quote is strictly speaking misleading: current non-classical approaches to semantic paradox seek to preserve classical logic also for unproblematic uses of the truth predicate, such as 'All the theorems of Peano Arithmetic are true', 'If $[65 + 57 = 125]$ is true, then $[65 + 57 \neq 125]$ is not true', and so on.

²¹ Our characterisation of classical recapture in LP is very much in line with an account discussed in Priest [2006a: 117-18] and Beall [2011].

$$\frac{\Gamma \vdash \varphi \quad \Delta \vdash \varphi \rightarrow \psi}{\Gamma, \Delta, \varphi \vdash \psi} \rightarrow\text{-E}_{\text{WC}} \quad \frac{\Gamma \vdash \varphi \quad \Delta \vdash \neg\varphi}{\Gamma, \Delta, \varphi \vdash \perp} \neg\text{-E}_{\text{WC}}$$

$$\frac{\Gamma \vdash \varphi \vee \psi \quad \Delta_0, \varphi \vdash \chi \quad \Delta_1, \psi \vdash \chi}{\Gamma, \Delta_0, \Delta_1, \varphi, \psi \vdash \chi} \vee\text{-E}_{\text{WC}}$$

We now show that classical logic can be recaptured in the sense of Definition 4.1, in each of the non-classical approaches introduced in section 3. We do so by adding the classical recapturing rules to our target theories and by then establishing that classical logic holds for φ whenever the relevant classical principles hold for φ .

Definition 4.2 (K3TT⁺). K3TT⁺ is the result of adding $\rightarrow\text{-I}_{\text{W}}$ and $\neg\text{-I}_{\text{W}}$ to K3TT.

By definition, K3TT⁺ is LEM-classical recapturing. To see that full CPL holds for φ in K3TT⁺ given $\varphi \vee \neg\varphi$, it is sufficient to notice that whenever $\varphi \vee \neg\varphi$ is derivable in K3TT⁺, then both $\rightarrow\text{-I}$ and $\neg\text{-I}$ hold in K3TT⁺. More precisely, if $\Gamma, \varphi \vdash \perp$ is derivable together with $\varphi \vee \neg\varphi$, then we can apply $\neg\text{-I}_{\text{W}}$ and apply Cut to $\varphi \vee \neg\varphi$, thus deriving $\Gamma \vdash \neg\varphi$, that is, the conclusion of full $\neg\text{-I}$. Similarly for $\rightarrow\text{-I}$.

Definition 4.3 (LPTT⁺). LPTT⁺ is the result of adding $\rightarrow\text{-E}_{\text{W}}$ and $\neg\text{-E}_{\text{W}}$ to LPTT.

By definition, LPTT⁺ is LNC-classical recapturing. As above, to see that full CPL holds for φ in LPTT⁺ if $\varphi \wedge \neg\varphi \vdash \perp$ does, it is sufficient to notice that full $\rightarrow\text{-E}$ and $\neg\text{-E}$ hold for φ in LPTT⁺ whenever $\varphi \wedge \neg\varphi \vdash \perp$ is derivable in LPTT⁺. More precisely, if $\Gamma \vdash \varphi$ and $\Delta \vdash \varphi \rightarrow \psi$ are derivable together with $\varphi \wedge \neg\varphi \vdash \perp$, one can derive $\Gamma, \Delta \vdash \psi$ courtesy of $\rightarrow\text{-E}_{\text{W}}$ and $\vee\text{-E}$. At a glance:

$$\frac{\frac{\Gamma \vdash \varphi \quad \Delta \vdash \varphi \rightarrow \psi}{\Gamma, \Delta \vdash \psi \vee (\varphi \wedge \neg\varphi)} \rightarrow\text{-E}_{\text{W}} \quad \frac{}{\psi \vdash \psi} \text{SRef} \quad \frac{\varphi \wedge \neg\varphi \vdash \perp}{\varphi \wedge \neg\varphi \vdash \psi} \perp\text{-E}}{\Gamma, \Delta \vdash \psi} \vee\text{-E}$$

The reasoning for $\neg\text{-E}$ is analogous.

Definition 4.4 (MALLTT⁺). MALLTT⁺ is the result of adding SContr_W to MALLTT.

As above, by definition MALLTT⁺ is LContr-classical recapturing. To see that classical logic holds for φ if LContr holds for φ , we reason in two steps, keeping in mind that MALL is classical logic minus SContr and with $\vee\text{-E}$ replaced with $\vee\text{-E}_{\text{W}}$.²² First, it is immediate to see that whenever LContr holds for φ , then SContr also holds:

$$\frac{\vdash \varphi \rightarrow (\varphi \wedge \varphi) \quad \frac{\Gamma, \varphi, \varphi \vdash \psi}{\Gamma, \varphi, \varphi \rightarrow (\varphi \wedge \varphi) \vdash \psi} \text{SContr}_{\text{W}}}{\Gamma, \varphi \vdash \psi} \text{Cut}$$

Second, we show that

$$(\text{RContr}) \vdash (\varphi \vee \varphi) \rightarrow \varphi$$

²² We are adapting Theorem 3.19 of Zardini [2011] to our framework.

is derivable from SContr and that, in turn, \vee -E is derivable from RContr. The following derivation establishes the first claim:

$$\begin{array}{c}
\frac{\frac{\frac{}{\neg\varphi \vdash \neg\varphi} \text{SRef} \quad \frac{}{\neg\varphi \vdash \neg\varphi} \text{SRef}}{\neg\varphi, \neg\varphi \vdash \neg\varphi \wedge \neg\varphi} \wedge\text{-I} \quad \frac{\frac{\frac{}{\neg\varphi \wedge \neg\varphi \vdash \neg\varphi \wedge \neg\varphi} \text{SRef} \quad \frac{\frac{\frac{}{\varphi \vee \varphi \vdash \varphi \vee \varphi} \text{SRef}}{\varphi \vee \varphi \vdash \neg(\neg\varphi \wedge \neg\varphi)} \text{DM}_2}}{\neg\varphi \wedge \neg\varphi, \varphi \vee \varphi \vdash \perp} \neg\text{-E}}{\varphi \vee \varphi, \neg\varphi, \neg\varphi \vdash \perp} \text{Cut}}{\frac{\frac{\frac{}{\varphi \vee \varphi, \neg\varphi, \neg\varphi \vdash \perp} \text{SContr}}{\varphi \vee \varphi, \neg\varphi \vdash \perp} \neg\text{-I}}{\frac{\frac{}{\varphi \vee \varphi \vdash \neg\neg\varphi} \neg\text{-I}}{\varphi \vee \varphi \vdash \varphi} \neg\text{-E}} \rightarrow\text{-I}}{\vdash (\varphi \vee \varphi) \rightarrow \varphi} \rightarrow\text{-E}
\end{array}$$

The second claim is proved as follows:

$$\frac{\frac{\frac{\Gamma \vdash \chi_0 \vee \chi_1 \quad \Delta_0, \chi_0 \vdash \varphi \quad \Delta_1, \chi_1 \vdash \varphi}{\Gamma, \Delta_0, \Delta_1 \vdash \varphi \vee \varphi} \vee\text{-E}_w \quad \vdash (\varphi \vee \varphi) \rightarrow \varphi}{\Gamma, \Delta_0, \Delta_1 \vdash \varphi} \rightarrow\text{-E}$$

Definition 4.5 (STTT_0^+). STTT_0^+ is the result of adding $\rightarrow\text{-E}_{\text{WC}}$, $\neg\text{-E}_{\text{WC}}$, and $\vee\text{-E}_{\text{WC}}$ to STTT_0 .

By definition, STTT_0^+ is Cut-classical recapturing. It can be verified that, whenever Cut holds for φ , full classical logic holds for φ . To see this, consider $\rightarrow\text{-E}$ and suppose Cut holds for φ . Then, given $\vdash \varphi$, full $\rightarrow\text{-E}$ is derived as follows:

$$\frac{\frac{\frac{\Gamma \vdash \varphi \quad \Delta \vdash \varphi \rightarrow \psi}{\Gamma, \Delta, \varphi \vdash \psi} \rightarrow\text{-E}_{\text{WC}}}{\frac{\frac{}{\Gamma, \Gamma, \Delta \vdash \psi} \text{Cut}}{\Gamma, \Delta \vdash \psi} \text{(possibly multiple uses of) SContr}}{\Gamma, \Delta \vdash \psi}$$

A similar reasoning applies to $\neg\text{-E}_{\text{WC}}$ and $\vee\text{-E}_{\text{WC}}$.²³

The classical recapturing properties of the non-classical theories introduced in section 3 are at the heart of our general revenge argument, to which we now turn.

5. Revenge

Revenge arguments fall into two broad categories: *object-linguistic* and *meta-theoretic*.²⁴ Meta-theoretic revenge arguments point to the inexpressibility in a theory S of notions definable in S 's meta-theory (which is typically classical). They are standardly dismissed on

²³ We should mention at least one alternative proposal for recapturing classical logic within a non-classical theory, viz. Priest's *minimally inconsistent* LP, or *mLP* for short [Priest 2006a: 222 and ff]. *mLP* is a non-monotonic logic that behaves like classical logic in the case of arguments with consistent premises and behaves like LP in the case of arguments with inconsistent premises. As far as we know, *mLP* has not been axiomatised. However, our account of classical recapture can be extended to *mLP*, provided such a logic can be given an axiomatisation satisfying **Classicality Principles**.

²⁴ For recent discussion of revenge, both object-linguistic and meta-theoretic, see e.g. Beall [2007a]; Field [2007]; Shapiro [2011b]. For general background on revenge, see the essays in Beall [2007c] and Scharp [2013: ch. 8].

the grounds that it is no surprise that classical notions are not expressible in a non-classical theory.²⁵ Object-linguistic revenge arguments typically point to the inexpressibility in a theory S of some notion N that plays some explanatory or expressive role in S . Notions such as indeterminacy [Field 2007, 2008] and instability [Zardini 2011] are cases in point.²⁶ The revenge paradoxes to be developed in this section are of the second, object-linguistic kind. In particular, they don't rely on classical semantic notions, and they apply to theories (such as the one developed in Zardini [2011]) for which no semantics is known. Sections 5.1-5.2 motivate naïve principles for paradoxicality and unparadoxicality. Sections 5.3-5.6 introduce our revenge paradoxes.

5.1 Paradoxicality and Unparadoxicality

General approaches to revenge are discussed in Beall [2007c], Priest [2007], Shapiro [2011b], and Scharp [2013, sec. 4.3]. For instance, Graham Priest argues that

[t]here is, in fact, a uniform method for constructing the revenge paradox—or extended paradox, as it is called sometimes. All semantic accounts have a bunch of Good Guys (the true, the stably true, the ultimately true, or whatever). These are the ones that we target when we assert. Then there's the Rest. The extended liar is a sentence, produced by some diagonalising construction, which says of itself that it's in the Rest. The diagonal construction . . . may then play havoc. This shows, incidentally, that the extended paradox is not really a different paradox. The pristine liar is the result of the construction when the theoretical framework is the standard one (all sentences are true or false, not both, and not neither) . . . 'Extended paradoxes' are simply the results of applying the construction in different theoretical frameworks. [Priest 2007: 226]

We are sympathetic to Priest's claim that revenge paradoxes are structurally similar to the run of the mill semantic paradoxes. However, his revenge recipe only describes extremely general features of revenge arguments and cannot be used to actually generate in a uniform way revenge paradoxes for a wide range of theories. Our aim in what follows is to provide a general revenge strategy for constructing revenge paradoxes for several non-classical theories satisfying **Classicality principles**, including some of the non-classical theories defended by Priest.

Our starting point is the distinction, present in each of the theories presented in section 3, between sentences that satisfy all the principles of classical logic and sentences that do so on pain of triviality. More precisely, let S be a $\mathfrak{F}_0, \dots, \mathfrak{F}_n$ -classical recapturing, non-trivial theory. We then say that a sentence φ is *paradoxical-in- S* if and only if \perp follows in S from the assumption that φ satisfies $\mathfrak{F}_0, \dots, \mathfrak{F}_n$; and that a sentence φ is *unparadoxical-in- S* if and only if it satisfies $\mathfrak{F}_0, \dots, \mathfrak{F}_n$ in S .²⁷ Paradoxicality and unparadoxicality so understood are intelligible notions at the core of the revisionary approach to semantic paradox. In

²⁵ See e.g. Field [2008: sec. 21.1] and Beall [2009: sec. 3.4].

²⁶ Field's theory involves a hierarchy of ever stronger notions of indeterminacy, whereby sentences such as λ can be only said to be indeterminate $_{\alpha}$, for some level α in the hierarchy [Field 2007].

²⁷ Our notion of paradoxicality-in- S is closely related to Jc Beall's notion of a *trivialiser* for S [Beall 2015].

keeping with the revisionist's treatment of truth, we treat them as *object-linguistic predicates*, Par and Un.²⁸

5.2 The Expressive Role of Par and Un

Revisionary theorists typically give the semantics of a language \mathcal{L} in \mathcal{L} . Here are two representative quotes:

If the formal language is to provide an adequate explication of the informal language that we use, it must contain its own metalanguage. [Reinhardt 1986: 227-9]

[M]y claim will be that there are languages that are sufficiently powerful to serve as their own meta-languages. [Field 2008: 18]

They further argue that the notion of truth for \mathcal{L} to be captured in \mathcal{L} must be *naïve*.²⁹ For instance, the following reasoning is taken to motivate the unrestricted rule Tr-E:

Agreement. All the theorems of Peano Arithmetic are true. φ is a theorem of Peano Arithmetic. Therefore, φ is true. Therefore, φ .

A parallel reasoning is taken to establish Tr-I:

Disagreement. Everything Lois says is not true. Lois says φ . Therefore, φ is not true. Therefore, $\neg\varphi$.

Similar considerations can be put forward for paradoxicality and unparadoxicality. For reasons of space, we only give one example, in the context of a paracomplete theory of naïve truth S . Consider the following case:

The logic student. Lois is a logic student who is learning how to reason in S . She (mistakenly) assumes $\lambda \vee \neg\lambda$. As a result, she carries out the Liar reasoning in S and derives \perp . She concludes that assuming that λ satisfies LEM trivialises S . As she puts it, λ is paradoxical, that is, Lois asserts Par($[\lambda]$).

In the above example, Lois adopts the following principle: that if S derives the sequent $\varphi \vee \neg\varphi \vdash \perp$, then it also derives the sequent $\vdash \text{Par}([\varphi])$. We call this principle Par-introduction, or Par-I for short. The principle immediately rules out the possibility of interpreting paradoxicality by means of a conditional. That is, given Par-I, Par(x) cannot be interpreted as (slightly abusing notation) $\text{Tr}(x \vee \neg x) \rightarrow \perp$, since in a paracomplete setting \rightarrow -I is not unrestrictedly valid and, as a result, $\text{Tr}([\lambda] \vee \neg[\lambda]) \rightarrow \perp$ cannot in general be inferred from a derivation of \perp in S from $\lambda \vee \neg\lambda$.

It might be tempting to interpret Par as derivability in S instead. After all, if S interprets a *modicum* of arithmetic, if there is derivation in S of \perp from $\lambda \vee \neg\lambda$, then S derives $\text{Der}_S([\lambda \vee \neg\lambda], [\perp])$, where Der_S is a standard, arithmetically definable derivability predicate for S . However, this can't be either, as shown by the following scenario:

²⁸ Since the chosen theory S will always be clear from context, we simply write Par(x) for paradoxicality and Un(x) for unparadoxicality.

²⁹ For a general case for naïve truth principles, see Field [2008: 209 and ff.] and Beall [2009: sec. 1.1].

Misguided reasoning. Clark reasons in S and assumes that everything Lois says is paradoxical. Lois asserts that φ . As a result, Clark infers that φ is paradoxical. However, Clark also proves that φ satisfies LEM, and hence all the principles of classical logic. From his claim that φ is paradoxical, that is such that $\varphi \vee \neg\varphi$ entails \perp , and his proof of $\varphi \vee \neg\varphi$, Clark concludes \perp .

The above scenario requires the following elimination rule: from $\text{Par}([\varphi])$ and ‘ φ satisfies LEM’, one may infer \perp . We call this principle Par-elimination, or Par-E for short. Just like Par-I rules out interpreting paradoxicality by means of a conditional, Par-E rules out interpreting such a notion as derivability-in- S . This is essentially a consequence of Löb’s Theorem, as we will see more fully in section 6.1 below.

We conclude that paradoxicality-in- S must be expressed via a single, primitive predicate Par, obeying Par-I and Par-E. Similar arguments apply to unparadoxicality-in- S , and to paraconsistent, non-contractive, and non-transitive theories.

The notions of paradoxicality and unparadoxicality now give rise to a *revenge argument*, to the effect that any theory extending the theories presented in section 3 expresses such notions only if it is trivial. In particular, consistent theories cannot express the notion of paradoxicality, while inconsistent theories cannot express the dual notion of unparadoxicality. We consider theories formulated in the language $\mathcal{L}_{\text{Tr}}^+$ obtained by adding Par and Un to \mathcal{L}_{Tr} . We extend to $\mathcal{L}_{\text{Tr}}^+$ and the theories formulated in it all the conventions and requirements stated in section 2 for languages and theories.

5.3 Paracomplete Revenge

We focus on K3TT^+ -based theories as our representative, catch-all paracomplete theories. Since paracomplete theories reject LEM for ‘paradoxical’ sentences and since K3TT^+ is LEM-classical recapturing, the rules for Par are as follows:³⁰

$$\frac{\Gamma, \varphi \vee \neg\varphi \vdash \perp}{\Gamma \vdash \text{Par}(\ulcorner \varphi \urcorner)} \text{LEM-Par-I} \qquad \frac{\Gamma \vdash \text{Par}(\ulcorner \varphi \urcorner) \quad \Delta \vdash \varphi \vee \neg\varphi}{\Gamma, \Delta \vdash \perp} \text{LEM-Par-E}$$

Definition 5.1 (K3TTP). K3TTP is the theory resulting from closing K3TT^+ under LEM-Par-I and LEM-Par-E.

Proposition 5.2. K3TTP is trivial, and so is the closure under LEM-Par-I and LEM-Par-I of any theory extending K3TT^+ .

It follows from Proposition 5.2 that (among others) the theories developed in Field [2002, 2008, 2013] and Yablo [2003] cannot express the notion ‘ φ yields absurdity if $\varphi \vee \neg\varphi$ holds’, on pain of triviality.

5.4 Paraconsistent Revenge

³⁰ With the exception of non-contractive theories, all the rules for Par and Un presented in this paper can be given a context-sharing formulation without affecting our revenge paradoxes.

Consider now paraconsistent approaches. In keeping with our account of classical recapture, we focus on LPTT⁺-based theories. We show that no extension of LPTT⁺ can express the notion of *unparadoxicality* introduced in section 5.1. Keeping in mind that LPTT⁺ is LNC-classical recapturing, a sentence φ is unparadoxical in LPTT⁺ if LNC holds for φ , that is, if LPTT⁺ proves $\varphi \wedge \neg\varphi \vdash \perp$. Conversely, if φ is unparadoxical in LPTT⁺, then LNC holds for φ , that is, if LPTT⁺ proves $\varphi \wedge \neg\varphi$ from Γ, Δ , then it also proves \perp from the same multi-set of assumptions. More formally:

$$\frac{\Gamma, \varphi \wedge \neg\varphi \vdash \perp}{\Gamma \vdash \text{Un}(\ulcorner \varphi \urcorner)} \text{LNC-Un-I} \quad \frac{\Gamma \vdash \text{Un}(\ulcorner \varphi \urcorner) \quad \Delta \vdash \varphi \wedge \neg\varphi}{\Gamma, \Delta \vdash \perp} \text{LNC-Un-E.}$$

Definition 5.3 (LPTTU). LPTTU is the theory resulting from closing LPTT⁺ under LNC-Un-I and LNC-Un-E.

Proposition 5.4. LPTTU is trivial, and so is the closure under LNC-Un-I and LNC-Un-E of any theory extending LPTT⁺.

It follows from Proposition 5.4 that the theories developed in Priest [2006b] and Beall [2009, 2011] cannot express the notion ‘ φ behaves classically if $\varphi \wedge \neg\varphi \vdash \perp$ holds’, on pain of triviality.

5.5 Non-contractive Revenge

Now to contraction-free approaches. Because of its prominence, we focus on Zardini’s non-contractive theory, but our result generalises. We begin by recalling classical recapture in a contraction-free setting. As we have seen in section 4 (Definition 4.4 and subsequent remarks), full SContr and \vee -E, and hence full classical logic, hold for φ in MALLTT⁺ whenever MALLTT⁺ derives $\varphi \rightarrow (\varphi \wedge \varphi)$. Keeping in mind that, according to SContr-free wisdom, SContr is the culprit of the semantic paradoxes, the paradoxicality predicate can now be interpreted as follows: if absurdity is derivable from the assumption that φ satisfies $\varphi \rightarrow (\varphi \wedge \varphi)$, then φ is paradoxical. Conversely, if φ is paradoxical and φ satisfies $\varphi \rightarrow (\varphi \wedge \varphi)$, then \perp is derivable.

This informal reasoning can be formalised thus. Let $[\varphi]^n$ be the multiset consisting of n occurrences of φ . Moreover, let us assume that Γ in LC-Par-I does not contain instances of $\varphi \rightarrow (\varphi \wedge \varphi)$, and let $m \geq 1$. Then, paradoxicality in a non-contractive setting is characterised by the following rules:

$$\frac{\Gamma, [\varphi \rightarrow (\varphi \wedge \varphi)]^m \vdash \perp}{\Gamma \vdash \text{Par}(\ulcorner \varphi \urcorner)} \text{LC-Par-I} \quad \frac{\Gamma \vdash \text{Par}(\ulcorner \varphi \urcorner) \quad \Delta \vdash \varphi \rightarrow (\varphi \wedge \varphi)}{\Gamma, \Delta^{n+1} \vdash \perp} \text{LC-Par-E}$$

where n is the highest number of occurrences of $\varphi \rightarrow (\varphi \wedge \varphi)$ occurring on the left-hand side of the sequents in the subderivation of $\Gamma \vdash \text{Par}(\ulcorner \varphi \urcorner)$ if Γ is non-empty, and 0 otherwise. Intuitively, the I-rule tells us that, if contracting m times on φ yields absurdity (where m contractions on φ are represented by $[\varphi \rightarrow (\varphi \wedge \varphi)]^m$), then φ is paradoxical. Conversely, the E-rule says that, if φ is paradoxical, then the assumption that φ can be contracted on (at least as many times as it is needed to declare it paradoxical) yields absurdity.

Definition 5.5 (MALLTTP). MALLTTP is the theory resulting from closing MALLTTP⁺ under LC-Par-I and LC-Par-E.

Proposition 5.6. *MALLTTP is trivial, and so is the closure under LC-Par-I and LC-Par-E of any theory extending MALLTTP⁺.*

It follows from Proposition 5.6 that (among others) the theory developed in Zardini [2011] cannot express the notion ‘ φ yields absurdity if $[\varphi \rightarrow (\varphi \wedge \varphi)]_m$ holds’, on pain of triviality.

It might be objected the non-contractive theorist who rejects contraction *in all its forms* has a reason to reject contracting on sentences of the form $\varphi \rightarrow (\varphi \wedge \varphi)$, and hence to reject LC-Par-I, which allows one to discharge multiple occurrences of $\varphi \rightarrow (\varphi \wedge \varphi)$. The resulting conception of paradoxicality would be problematic, however. It would commit the non-contractive theorist to distinguishing between different numbers of applications of SContr in a derivation, which would sit poorly with her diagnosis of what goes wrong in paradoxical derivations. According to non-contractive wisdom, *indiscriminate uses* of SContr must be rejected *in general*. That is, non-contractive theorists disallow the following generalised version of SContr:

$$\frac{\Gamma, [\varphi]^j \vdash \Delta}{\Gamma, [\varphi]^i \vdash \Delta} \text{SContr}^* \quad (\text{where } j > i),$$

according to which, if Δ follows from Γ and i occurrences of φ , then Δ follows from Γ and at least one occurrence of φ . The idea that if SContr* applied to φ leads to \perp then φ is non-contractable is at the heart of the non-contractive approach to semantic paradox: one must disallow *whatever number* of applications of SContr to φ lead to \perp in a paradoxical derivation. This is captured by our rule LC-Par-I, but cannot be expressed by the non-contractive theorist who expresses paradoxicality by means of denumerably many paradoxicality predicates.

5.6 Non-transitive Revenge

We finally turn to non-transitive approaches. We focus on the theory STTT₀⁺ but, again, our results generalise. To begin with, we notice that, in STTT₀⁺, full classical logic holds for φ whenever Cut does (see Definition 4.5 and subsequent remarks). This in turn justifies the following characterisation of unparadoxicality. On the one hand, if φ is ‘cuttable on’, that is, if the conclusion of an instance of Cut applied to φ is derivable from its premises, then φ is unparadoxical. On the other, if φ is unparadoxical and the premises of an instance of Cut applied to φ are derivable, so is their conclusion.

Our revenge argument against STTT₀⁺ shows that such a theory cannot express unparadoxicality so understood. It makes use of *higher-order* rules: rules that allow one to discharge *entire sequents*, as well as sentences.³¹ We are now in a position to formulate the rules governing the unparadoxicality predicate:

³¹ To our knowledge, higher-order rules were first introduced by Peter Schroeder-Heister [1984: 1284-5], who pointed out that if sentence-assumptions are technically *temporary axioms*, nothing should prevent one from also making use of *rule-assumptions*, understood as *temporary rules* (in our setting, such temporary rules are sequent-assumptions).

$$\begin{array}{c}
\boxed{\Gamma \vdash \varphi}^n \quad \boxed{\Delta, \varphi \vdash \psi}^n \\
\vdots \quad \vdots \\
\vdots \\
\frac{\Gamma, \Delta \vdash \psi}{\Gamma, \Delta \vdash \text{Un}(\ulcorner \varphi \urcorner)} \text{Cut-Un-I, } n \qquad \frac{\Gamma \vdash \text{Un}(\ulcorner \varphi \urcorner) \quad \Delta_0 \vdash \varphi \quad \Delta_1, \varphi \vdash \psi}{\Gamma, \Delta_0, \Delta_1 \vdash \psi} \text{Cut-Un-E,}
\end{array}$$

where the box left of the discharge line in Cut-Un-I signals that the rule-assumptions $\Gamma \vdash \varphi$ and $\Delta, \varphi \vdash \psi$ may not be discharged vacuously.³² Again, the rules are justified by the account of classical recapture given in section 4 (see especially Definition 4.5). Cut-Un-I says that, if φ is ‘cuttable on’, then it is unparadoxical. Conversely, Cut-Un-E tells us that, if φ is unparadoxical (given Γ), and hence ‘cuttable’, and both $\Delta_0 \vdash \varphi$ and $\Delta_1, \varphi \vdash \psi$ are provable, then $\Gamma, \Delta_0, \Delta_1 \vdash \psi$ follows.

Definition 5.7 (STTTU₀). STTTU₀ is the theory resulting from closing STTT₀⁺ under Cut-Un-I and Cut-Un-E.

Proposition 5.8. STTTU₀ is trivial, and so is the closure under Cut-Un-I and Cut-Un-E of any theory extending STTT₀⁺.

It follows from Proposition 5.8 that (among others) the theories developed in Ripley [2012]; Cobreros et al. [2013] cannot express the notion ‘ φ behaves classically given a derivation of $\vdash \psi$ from $\vdash \varphi$ and $\varphi \vdash \psi$ ’, on pain of triviality.

We notice that the derivation of \perp in the proof of Proposition 5.8 (see Appendix) is not normal, since it involves a use of Cut-Un-E immediately after a use of Cut-Un-I.³³ This suggests that, unlike STTT₀⁺, Neil Tennant’s *Core Logic*, a logic in which all proofs are normal proofs, may support the rules for naïve truth *together with* Cut-Un-I and Cut-Un-E, in keeping with Tennant’s conjecture that the semantic paradoxes all involve derivations that cannot be brought into normal form [Tennant 1982, 2015].³⁴ Does it follow that our revenge argument doesn’t apply to the non-transitive approach defended in Tennant [2015]? Tennant [2015: 593] advocates a positive answer. However, we do not share Tennant’s optimism. In the proof of Proposition 5.8, we give *normal* proofs of $\vdash \text{Un}(\ulcorner \zeta \urcorner)$, $\vdash \zeta$, and $\zeta \vdash \perp$, where ζ is $\neg \text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)$. That is, Tennant’s theory proves both that φ is ‘cuttable’ and the premises of a cut on φ . Yet, *one cannot cut on φ* in such a theory. This means that $\text{Un}(\ulcorner \zeta \urcorner)$, viz. that φ is ‘cuttable’, no longer has its intended meaning in Tennant’s framework. The

³² See Tennant [2012: 4-5].

³³ The derivation is also non-normal in the sense of Tennant [2012], since the major premise of Cut – Un – E in the last step of the derivation has non-trivial proof work above it.

³⁴ Core Logic is a non-transitive logic all of whose proofs are in normal form. The logic comes in a constructive and in a non-constructive variety; our remarks apply to both. As Tennant [2015] observes, the derivations of semantic paradoxes such as the Liar are all invalid in Core Logic supplemented by suitable rules for naïve truth. More precisely, just like in STTT₀⁺, the final use of Cut in the paradoxical derivations (or of the unrestricted \neg -E and \rightarrow -E) is disallowed, on the grounds that it would bring in non-normality. Although Tennant doesn’t prove consistency for a theory of naïve truth whose underlying logic is Core Logic, it can be shown that the consistency proofs available for STTT [Cobreros et al. 2013; Ripley 2012] immediately extend to the theory given by closing Core Logic under the rules for naïve truth (this follows from the fact that Core Logic is a relevant subsystem of the logic ST). This in turn answers the technical question raised in Tennant [2015: sec. 4.3.4] whether one can prove the non-triviality of the approach he recommends.

framework is not trivial, but non-triviality is only restored at the price of expressive incompleteness.

6. What our Results Show

The paradoxes of sections 5.3-5.6 make use of logical rules that are valid in the theories they trivialise. Insofar as the theories introduced in section 3 are representative of the revisionary approach to semantic paradox, it follows that revisionary treatments of the Liar Paradox and of other run of the mill paradoxes don't apply to the paradoxes of sections 5.3-5.6. Yet the notions of paradoxicality and unparadoxicality codify a minimal lesson to be drawn from the semantic paradoxes: that given the naïve truth rules, sentences such as λ satisfy all the classical rules only on pain of triviality, whereas sentences such as $t = t$ unproblematically satisfy those rules. The results of sections 5.3-5.6 show that the expression of such a truism is precluded to most non-classical theorists, on pain of adopting an extremely weak, and possibly unworkable, logic. For instance, it is a consequence of the proof of Proposition 5.2 that a paracomplete logic of paradox cannot contain all of SRef, SContr, \rightarrow -E, and \vee -I. Likewise, it follows from the proof of Proposition 5.8 that a non-transitive logic of paradox cannot contain SRef, SContr, the rules for conjunction, and a very weak form of negation elimination. And so on. Our revenge strategy is perfectly general. Although the paradoxes of sections 5.3-5.6 make use of theory-specific notions of paradoxicality and unparadoxicality, it can be shown that the naïve rules for Par and Un are all instances of a more general template.³⁵

Solomon Feferman [1984: 95] once wrote, referring to theories of truth based on the logic K3, that 'nothing like sustained ordinary reasoning can be carried out' in them. While his remark may apply to weak logics such as K3 and LP, it may be thought to be unfair as a criticism of the stronger non-classical theories developed since 1984, such as the structural ones given in Field [2002, 2008, 2017] and Beall [2009], and the substructural ones given in Zardini [2011] and Cobreros et al. [2012]. Even classical theorists concede that, *pace* Feferman, some such theories are surprisingly strong. Vann McGee [2010], for instance, reports to have been 'astonished' by the 'combination of transparency and logical strength' exhibited by Field's paracomplete theory.

The results of sections 5.3-5.6 vindicate the spirit of Feferman's remark. Just like classical logic, and many other strong logics, are known to be incompatible with naïve truth, our results show that a wide range of reasonably strong non-classical logics are incompatible with naïve paradoxicality and unparadoxicality. And, as we argued in sections 5.1-5.2, just like there are strong reasons for wanting truth to be naïve, and hence to adopt one of the non-classical logics introduced in section 4 (or some extension thereof), there are parallel reasons for wanting paradoxical and unparadoxicality to be *also* naïve, and hence to adopt a weaker still non-classical logic—one in which the arguments of sections 5.3-5.6 no longer go through. By the revisionary theorist's own lights, strong non-classical theories such as Field's are ultimately incompatible with the project of giving the semantics of a language \mathcal{L} in \mathcal{L} .

In what follows, we briefly explore the relation between our results and Löb's Theorem (section 6.1). We argue that the naïve principles for paradoxicality and unparadoxicality can be seen to be compatible with classical limitative results such as Löb's Theorem just in the same way as a naïve notion of truth can be seen to be compatible with classical limitative results such as Tarski's Theorem. We then point to a parallelism between our arguments and a recent revenge argument for classical theories (section 6.2).

³⁵ For reasons of space, we leave out the precise formulation of the template.

6.1 Paradoxicality and Derivability

It could be argued that the results of sections 5.3-5.6 are hardly surprising, on the grounds that the eliminations rules for Par and Un are unacceptable in the lights of Löb's Theorem. More precisely, let S be a theory satisfying the Hilbert-Bernays conditions for a predicate Prov_S expressing standard provability-in- S .³⁶ It is a consequence of Löb's Theorem that, if S proves every instance of $\text{Prov}_S([\varphi]) \rightarrow \varphi$, then it also proves any sentence φ . Consider the paradoxicality predicate Par. Its rules can be rewritten using a two-place derivability predicate $\text{Der}_S(x, y)$ expressing that y is derivable from x in S . For instance, the LEM-Par rules can be seen as instances of the following general rules:

$$\frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \text{Der}_S([\varphi], [\psi])} \text{Der}_S\text{-I} \qquad \frac{\Gamma \vdash \text{Der}_S([\varphi], [\psi]) \quad \Delta \vdash \varphi}{\Gamma, \Delta \vdash \psi} \text{Der}_S\text{-E}$$

However, $\text{Der}_S\text{-E}$ entails

$$\text{Der}_S([\top], [\psi]) \rightarrow \psi,$$

which is equivalent to $\text{Prov}_S([\psi]) \rightarrow \psi$, from which ψ is derivable in S via Löb's Theorem. It is now natural to object that the rules for Par and Un employed in the results of sections 5.3-5.6 are but special cases of naïve rules for provability-in- S that are *already known* to be unacceptable because of Löb's Theorem.³⁷

The foregoing reasoning requires that paradoxicality-in- S be interpreted as standard derivability-in- S . More precisely, it assumes that the paradoxicality and unparadoxicality predicates be interpreted by means of an arithmetically definable derivability predicate Der_S satisfying versions of the Hilbert-Bernays derivability conditions. On such a construal, the introduction rules for Par and Un are arithmetically derivable, while the elimination rules only hold on pain of triviality. However, we have argued in 5.2 that Par is *not* to be interpreted via a standard derivability predicate: scenarios such as **Misguided reasoning** rule out this possible interpretation.

If it is insisted that paradoxicality and unparadoxicality are to be interpreted via a standard derivability predicate, and therefore fail to obey their elimination rules because of Löb's Theorem, then a parallel argument can be given that *truth* is to be interpreted via some arithmetically definable predicate, and therefore fails to obey the naïve truth rules, because of Tarski's Theorem. For instance, it might be pointed out that sufficiently strong theories validate all instances of the T-Schema restricted to Σ_n -sentences of the base language. More precisely, they validate all instances of the following schema:

$$\varphi \leftrightarrow \text{Tr}_{\Sigma_n}([\varphi]),$$

for φ a Σ_n -sentence of the base language (for any given n) and Tr_{Σ_n} a predicate definable in the base theory. To be sure, restricting the T-Schema to Σ_n -sentences is inadequate for the purpose of giving the semantics of a language \mathcal{L} in \mathcal{L} , and for fulfilling the expressive role of 'true'.³⁸ However, it might be argued, the same holds for any restricted notion of

³⁶ Standard provability predicates satisfy the Hilbert-Bernays 'derivability conditions', i.e. predicate analogues of the rules of necessitation and of the K and 4 axioms of modal logic.

³⁷ For more discussion see Field [2017] and Murzi and Rossi [2018a].

³⁸ Suppose one tried to express her acceptance of all of PA's theorems by asserting that all the theorems of PA are true. If 'true' is modelled by a Σ_n -truth predicate, one would have thereby only expressed acceptance of the Σ_n -theorems of PA.

paradoxicality or unparadoxicality. For instance, if Par-I is restricted, some sentences that behave non-classically in S cannot be said to be paradoxical and, if Par-E is restricted, one cannot infer from the claim that φ is paradoxical that φ satisfies S 's classical recapturing principles only on pain of triviality. Just like it is possible to validate all instances of the T-Schema in spite of Tarski's Theorem, it is also possible, and consistent with Löb's Theorem, to non-trivially have all instances of the naïve rules for paradoxicality and unparadoxicality, provided one adopts a *very weak* non-classical logic.

6.2 Non-classical and Classical Revenge

The revenge paradoxes for non-classical theories given in sections 5.3-5.6 are closely related to a general revenge argument for classical theories recently introduced by Andrew Bacon [2015]. Bacon's starting point is analogous to ours: while non-classical theories of truth restrict the application of classical logic to some sentences, classical theories of truth restrict the application of *naïve truth-theoretic principles* to some sentences.³⁹ In order to express such a distinction, Bacon introduces a 'healthiness' predicate H satisfying the following scheme:

$$(SRT) H([\varphi]) \rightarrow (Tr([\varphi]) \leftrightarrow \varphi).$$

That is, Bacon characterises the healthy sentences as those that satisfy naïve truth-theoretic principles such as the T-Schema. Bacon then shows that every (sufficiently expressive) classical theory of truth that includes all instances of SRT proves sentences that it also proves to be unhealthy, that is, proves $\varphi \wedge \neg H([\varphi])$ for some φ . Under the assumption (which Bacon does not endorse) that H satisfies the following necessitation rule

$$\frac{\vdash \varphi}{\vdash H(\ulcorner \varphi \urcorner)} \text{ (H-NEC),}$$

Bacon's argument shows the resulting theories to be trivial.

The parallel between Bacon's argument and ours is easy to see. On one hand, Bacon's argument shows that *classical* theories cannot be closed under natural principles governing a healthiness predicate true of all and only the sentences that satisfy the naïve truth rules, keeping classical logic fixed (and similarly for unhealthiness). On the other hand, our results show that *non-classical* theories cannot be closed under natural principles governing a unparadoxicality predicate, true of all and only the sentences that satisfy all the principles of classical logic, keeping the naïve truth rules fixed (and similarly for paradoxicality).

7 Objections and Replies

Revisionary theorists might object to the paradoxes of sections 5.3-5.6 on the grounds that our naïve principles for paradoxicality and unparadoxicality trade on a deep misunderstanding of their views. More specifically, they might argue that our revenge arguments try to force revisionary theories to express notions whose intelligibility they have long argued against.⁴⁰ For instance, Field writes:

³⁹ For instance, classical theories do not derive the following instance of the T-Schema: $Tr([\lambda]) \leftrightarrow \lambda$.

⁴⁰ See e.g. Priest [2006b: ch. 5], Field [2008: sec. 21.1], Beall [2007b, 2009: ch. 3].

there is no negation that obeys [both of \neg -I and \neg -E] without restriction: if there were, it would be impossible to have a [naïve] truth predicate. [Field 2008: 309]

He further suggests that there is no coherent notion satisfying both of \rightarrow -I and \rightarrow -E. Similarly, it might be argued that while the non-classical theories of section 3 cannot express paradoxicality or unparadoxicality, this is not a problem, since there is no coherent notion to be expressed beyond the ones already expressible in such theories. For instance, the non-contractive theorist might insist that LC-Par-I validates some illicit, and ultimately unacceptable, uses of contraction. Likewise, the non-transitive theorist might insist that ‘cuttable’ is to be interpreted by means of a conditional: *if* one can assert the premises of a cut on φ , *then* one may assert the conclusion of such a cut. She might then point out that to assume that one can in general infer the consequent of this conditional from the premises is just to assume the unrestricted validity of \rightarrow -E, which non-transitive theorists reject (since it makes Cut admissible).

This kind of reply is perfectly coherent, as far as it goes. But how far does it go? As we observed in section 5.5, LC-Par-I expresses the basic non-contractivist diagnosis of the paradoxes, namely that *contraction in general* is at the root of those paradoxes. Whether S is trivialised by one, two, or m uses of contraction, these are all manifestations of the *same* problem. Similarly, if one’s logic doesn’t allow interpreting ‘cuttable’ in such a way that one *can* cut on a cuttable sentence, then this is a serious expressive limitation of the logic. The English expression ‘cuttable’ still means cuttable, and any adequate solution to the paradoxes should respect this basic fact about English. (Imagine the surprise of our logic student, Lois, if she were to learn that, even if φ is cuttable and one can assert the premises of a cut on φ , one may still not be allowed to derive the conclusion of such a cut.) Similarly for the other cases: rejecting the rules for Par and Un restores non-triviality only at the price of serious expressive limitations.

We can think of two main possible reactions at this point. First, upon deriving \perp in S from a classical recapturing principle, non-classical theorists might concede that φ is paradoxical, and insist that it is just a limitation about S that it cannot non-trivially prove as much—a limitation one must learn to live with. However, while classical theorists may be sympathetic to this suggestion, it does not sit well with the project of giving the semantics for a language \mathcal{L} in \mathcal{L} .

Second, one might offer instead a hierarchical treatment of the notions of paradoxicality and unparadoxicality, much in the same way as, in order to semantically characterise intuitively defective sentences such as λ , Field [2007, 2008: ch. 22-23] defines a hierarchy of determinacy operators.⁴¹ For instance, the thought would be that although the sentence ρ identical to $\text{Tr}([\rho]) \rightarrow \text{Par}([\rho])$ figuring in the proof of Proposition 5.2 cannot be said to be paradoxical in the sense expressed by Par, it can still non-trivially be said to be paradoxical in a stronger sense expressed by a new predicate Par_1 . And so on (and similarly for Un). It might then be insisted that, as Field puts it in a related context, this ‘would not nearly have the devastating impact on our reasoning a stratification of truth predicates would have’ [Field 2014: 22].

Field’s strategy has been criticised in a number of places (see e.g. Priest [2007]; Rayo and Welch [2007]; Welch [2008, 2014]). Here we limit ourselves to noticing, first, that paradoxicality and unparadoxicality appear to be just as central as truth in the revisionary theorist’s cognitive life. That λ entails absurdity if reasoned with classically, and is therefore paradoxical in our sense, is a minimal but key revisionary lesson of the Liar Paradox. Second,

⁴¹ The strategy can be generalised to other non-classical approaches (see Field [2008: ch. 27]).

the arguments that are usually put forward against non-hierarchical accounts of truth equally apply to paradoxicality and unparadoxicality. For instance, if φ 's paradoxicality-in- S can only be asserted by means of a stronger paradoxicality predicate 'paradoxicality₁-in- S ', it might be argued following Kripke [1975: 695-6] that there is no way to interpret a discourse in which two speakers attribute paradoxicality-in- S to everything they say.

8 Concluding remarks

Non-classical approaches to paradox are attractive for two main reasons: they allow one to retain extremely intuitive naïve semantic principles; and they often allow one to do so using non-classical logics that can be both natural and strong.⁴² This is a tempting, if ultimately radical, thought. Existing revisionary approaches cannot express one of the basic lessons of the semantic paradoxes, namely that certain sentences trivialise one's theory if reasoned with classically, while others don't—facts that are built into the classical recapturing properties enjoyed by each of the representative theories discussed in the course of this paper. As a result, revisionary theorists must resort to logics that are significantly weaker than the four families of logic introduced in section 4. This is especially problematic for revisionary theorists who place special emphasis in their theories' ability to recapture classical theories and restrict classical logic exactly when it creates paradox-driven trouble. The original Liar Paradox, and other run of the mill paradoxes, can be blocked by weakening classical logic. But, in view of the paradoxes of paradoxicality and unparadoxicality, the Liar Paradox inevitably reemerges in new theory-relative clothes to exact its revenge.⁴³

REFERENCES

- Asenjo, F. G. 1966. A calculus of antinomies, *Notre Dame Journal of Formal Logic* 16: 103–5.
- Bacon, A. 2015. Can the classical logician avoid the revenge paradoxes?, *Philosophical Review* 124/3: 299–352.
- Beall, J. 2007a. Prolegomenon to future revenge, in *Revenge of the Liar*, ed. J. Beall, Oxford: Oxford University Press: 1–30.
- Beall, J. 2007b. Truth and paradox: a philosophical sketch, in *Philosophy of Logic*, ed. D. Jacquette, Oxford: Elsevier: 325–410.
- Beall, J., ed., 2007c. *Revenge of the Liar*, Oxford: Oxford University Press.

⁴² Leon Horsten [2009, 2012] argues for the naturalness of a certain non-classical theory of truth. Field [2008]; Zardini [2011]; Ripley [2012]; Cobreros et al. [2013] all emphasise the logical strength of their respective approaches.

⁴³ Early versions of some of this material were presented at conferences and colloquia in Amsterdam, Barcelona, Bochum, Groningen, Oxford, Rome, Florence, Montreal, and Pamplona. We thank these audiences for valuable comments. We also wish to thank Jc Beall, Andrea Cantini, Paul Egré, Hartry Field, Rosanna Keefe, Francesco Paoli, Lavinia Picollo, Dave Ripley, Lucas Rosenblatt, Lionel Shapiro, Neil Tennant, and Elia Zardini for helpful discussions over the years. Special thanks to Dave Ripley for invaluable comments on previous drafts of this paper that have led to substantial improvements. We are finally grateful to the ÖAD and the FWF (grant number P29716-G24) for generous financial support during the time this paper was written.

- Beall, J. 2009. *Spandrels of Truth*, Oxford: Oxford University Press.
- Beall, J. 2011. Multiple-conclusion LP and default classicality, *Review of Symbolic Logic* 4/1: 326–36.
- Beall, J. 2015. Trivializing sentences and the promise of semantic completeness, *Analysis* 75/4: 573–84.
- Beall, J. and Murzi, J. 2013. Two flavors of Curry’s Paradox, *The Journal of Philosophy* 110/3: 143–65.
- Cobreros, P., Egré, P., Ripley, D. and van Rooij, R. 2012. Tolerant, classical, strict, *Journal of Philosophical Logic* 41/2: 347–85.
- Cobreros, P., Egré, P., Ripley, D. and van Rooij, R. 2013. Reaching transparent truth, *Mind* 122/488: 841–66.
- Feferman, S. 1984. Toward useful type-free theories I, *Journal of Symbolic Logic* 49: 75–111.
- Feferman, S. 1991. Reflecting on incompleteness, *The Journal of Symbolic Logic* 56/1: 1–49.
- Field, H. 2002. Saving the truth schema from paradox, *Journal of Philosophical Logic* 31/1: 1–27.
- Field, H. 2007. Solving the paradoxes, escaping revenge, in *Revenge of the Liar*, ed. J. Beall, Oxford: Oxford University Press: 53–144.
- Field, H. 2008. *Saving Truth from Paradox*, Oxford: Oxford University Press.
- Field, H. 2013. Naive truth and restricted quantification: Saving truth a whole lot better, *Review of Symbolic Logic* 7/1: 147–91.
- Field, H. 2017. Disarming a paradox of validity, *Notre Dame Journal of Formal Logic* 58/1: 1–19.
- Gauker, C. 2006. Against stepping back: A critique of contextualist approaches to the semantic paradoxes, *Journal of Philosophical Logic* 35/4: 393–422.
- Glanzberg, M. 2004. A contextual-hierarchical approach to truth and the liar paradox, *Journal of Philosophical Logic* 33: 27–88.
- Goodship, L. 1996. On dialetheism, *Australasian Journal of Philosophy* 74: 153–61.
- Halbach, V. 2011. *Axiomatic Theories of Truth*, Cambridge: Cambridge University Press.
- Halbach, V. and Horsten, L. 2006. Axiomatizing Kripke’s theory of truth, *Journal of Symbolic Logic* 71: 677–712.

- Horsten, L. 2009. Levity, *Mind* 118/471: 555–81.
- Horsten, L. 2012. *The Tarskian Turn. Deflationism and axiomatic truth*, Cambridge (Mass.): MIT Press.
- Kleene, S. C. 1952. *Introduction to Metamathematics*, North-Holland.
- Kripke, S. 1975. Outline of a theory of truth, *Journal of Philosophy* 72: 690–716.
- Leitgeb, H. 2005. What truth depends on, *Journal of Philosophical Logic* 34/2: 155–92.
- Leitgeb, H. 2007. On the metatheory of Field’s ‘Solving the paradoxes, escaping revenge’, in *Revenge of the Liar*, ed. J. Beall, Oxford: Oxford University Press: 159–83.
- Mares, E. and Paoli, F. 2014. Logical consequence and the paradoxes, *Journal of Philosophical Logic* 43: 439–69.
- McGee, V. 1991. *Truth, Vagueness, and Paradox*, Indianapolis: Hackett Publishing Company.
- McGee, V. 2010. Field’s logic of truth, *Philosophical Studies* 147/3: 421–32.
- Murzi, J. and Rossi, L. 2018a. Naïve validity. Forthcoming in *Synthese*.
- Murzi, J. and Rossi, L. 2018b. Reflection principles and the liar in context. *Philosophers’ Imprint*. 18/15: 1-18
- Nicolai, C. and Rossi, L.: 2018, Principles for object-linguistic consequence: from logical to irreflexive, *Journal of Philosophical Logic* 47/3: 549-577.
- Parsons, C. 1974. The Liar Paradox, *Journal of Philosophical Logic* 3/4: 381–412.
- Priest, G. 1979. The logic of paradox, *Journal of Philosophical Logic* 8: 219–41.
- Priest, G. 2006a. *Doubt Truth to be a Liar*, Oxford: Oxford University Press.
- Priest, G. 2006b. *In Contradiction*, Oxford: Oxford University Press. Expanded edition (first published 1987 Kluwer-Dordrecht).
- Priest, G. 2007. Revenge, Field, and ZF, in *Revenge of the Liar*, ed. J. Beall, Oxford: Oxford University Press: 225–33.
- Rayo, A. and Welch, P. D. 2007. Field on revenge, in *Revenge of the Liar*, ed. J. Beall, Oxford: Oxford University Press: 234–49.
- Reinhardt, W. N. 1986. Some remarks on extending and interpreting theories with a partial predicate for truth, *Journal of Philosophical Logic* 15: 219–51.
- Ripley, D. 2012. Conservatively extending classical logic with transparent truth, *Review of Symbolic Logic* 5/2: 354–78.

- Ripley, D. 2013. Paradoxes and failures of cut, *Australasian Journal of Philosophy* 91/1: 139–64.
- Scharp, K. 2013. *Replacing Truth*, Oxford: Oxford University Press.
- Schroeder-Heister, P. 1984. A natural extension of natural deduction, *Journal of Symbolic Logic* 49: 1284–99.
- Shapiro, L. 2011a. Deflating logical consequence, *The Philosophical Quarterly* 61: 320–42.
- Shapiro, L. 2011b. Expressibility and the Liar’s revenge, *Australasian Journal of Philosophy* 89/2: 297–314.
- Tarski, A. 1936. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica Commentarii Societatis Philosophicae Polonorum* 1: 261-405, translated as “The Concept of Truth in Formalized Languages” in Tarski (1956), pp. 152-278.
- Tarski, A. 1956. *Logic, Semantics, Metamathematics*, Oxford: Oxford University Press.
- Tennant, N. 1982. Proof and paradox, *Dialectica* 36: 265–96.
- Tennant, N. 2012. Cut for Core Logic, *Review of Symbolic Logic* 5/3: 450–79.
- Tennant, N. 2015. A new unified account of truth and paradox, *Mind* 124/494: 571–605.
- Weir, A. 2005. Naïve truth and sophisticated logic, in *Deflationism and Paradox*, eds. J. Beall and B. Armour-Garb, Oxford: Oxford University Press: 218–49.
- Welch, P. 2008. Ultimate truth *vis à vis* stable truth, *Review of Symbolic Logic* 1: 126–42.
- Welch, P. 2014. Some observations on truth hierarchies, *The Review of Symbolic Logic* 7/1: 1–30.
- Woods, J. 2018. Logical partisanhood. Forthcoming in *Philosophical Studies*.
- Yablo, S. 2003. New grounds for naïve truth theory, in *Liars and Heaps. New Essays on Paradox*, ed. J. Beall, PLACE: Oxford University Press: 312–30.
- Zardini, E. 2011. Truth without contra(di)ction, *Review of Symbolic Logic* 4: 498–535.
- Zardini, E. 2014. Naïve truth and naïve logical properties, *Review of Symbolic Logic* 7/2: 351– 84.

Appendix

We provide proofs of Propositions 5.2, 5.4, 5.6, and 5.8.

Proposition 5.2. *K3TTP is trivial, and so is the closure under LEM-Par-I and LEM-Par-E of any theory extending K3TT⁺.*

Proof. We make use of the following K3TT⁺-valid form of \rightarrow -I (given \rightarrow 's materiality, this is in effect a restricted form of \vee -I):

$$\frac{\vdash \psi}{\vdash \varphi \rightarrow \psi} \rightarrow\text{-I}$$

We now reason thus, in K3TTP. Let ρ be identical to $\text{Tr}([\rho]) \rightarrow \text{Par}([\rho])$. We first prove $\rho \vee \neg\rho \vdash \text{Tr}([\rho]) \vee \neg\text{Tr}([\rho])$:

$$\frac{\frac{\frac{\overline{\rho \vdash \rho} \text{SRef}}{\rho \vdash \text{Tr}(\ulcorner \rho \urcorner)} \text{Tr-I}}{\rho \vdash \text{Tr}(\ulcorner \rho \urcorner) \vee \neg\text{Tr}(\ulcorner \rho \urcorner)} \vee\text{-I}_1 \quad \frac{\frac{\overline{\neg\rho \vdash \neg\rho} \text{SRef}}{\neg\rho \vdash \neg\text{Tr}(\ulcorner \rho \urcorner)} \neg\text{-Tr-I}}{\neg\rho \vdash \text{Tr}(\ulcorner \rho \urcorner) \vee \neg\text{Tr}(\ulcorner \rho \urcorner)} \vee\text{-I}_2}{\rho \vee \neg\rho \vdash \text{Tr}(\ulcorner \rho \urcorner) \vee \neg\text{Tr}(\ulcorner \rho \urcorner)} \vee\text{-E} \text{SRef}$$

Call this derivation D_0 . In our next step, we prove $\rho \vee \neg\rho \vdash \rho$:

$$\frac{\frac{\frac{\frac{\overline{\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Tr}(\ulcorner \rho \urcorner)} \text{SRef}}{\text{Tr}(\ulcorner \rho \urcorner) \vdash \rho} \text{Tr-E}}{\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)} \text{Def. of } \rho}{\text{Tr}(\ulcorner \rho \urcorner), \text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SContr}}{\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Par}(\ulcorner \rho \urcorner)} \rightarrow\text{-E}}{\frac{\text{Tr}(\ulcorner \rho \urcorner) \vee \neg\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)}{\text{Tr}(\ulcorner \rho \urcorner) \vee \neg\text{Tr}(\ulcorner \rho \urcorner) \vdash \rho} \rightarrow\text{-Iw Def. of } \rho} \text{Cut}}{\rho \vee \neg\rho \vdash \text{Tr}(\ulcorner \rho \urcorner) \vee \neg\text{Tr}(\ulcorner \rho \urcorner)} \text{D}_0}{\rho \vee \neg\rho \vdash \rho} \text{D}_1$$

Call this derivation D_1 . We use it to show that ρ is paradoxical:

$$\frac{\frac{\frac{\frac{\text{D}_1}{\rho \vee \neg\rho \vdash \rho}}{\rho \vee \neg\rho \vdash \text{Tr}(\ulcorner \rho \urcorner)} \text{Tr-I}}{\rho \vee \neg\rho, \rho \vee \neg\rho \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SContr}}{\rho \vee \neg\rho \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SContr}}{\frac{\frac{\frac{\text{D}_1}{\rho \vee \neg\rho \vdash \rho}}{\rho \vee \neg\rho \vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)} \text{Definition of } \rho}}{\rho \vee \neg\rho, \rho \vee \neg\rho \vdash \perp} \text{SContr}}{\frac{\rho \vee \neg\rho \vdash \perp}{\vdash \text{Par}(\ulcorner \rho \urcorner)} \text{LEM-Par-I}} \text{LEM-Par-E}$$

We now have a proof of $\text{Par}([\rho])$ —call it D_2 . This in turn yields absurdity, as the following derivation shows:

$$\frac{\frac{\frac{\text{D}_2}{\vdash \text{Par}(\ulcorner \rho \urcorner)}}{\vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)} \rightarrow\text{-I}}{\vdash \rho} \text{Definition of } \rho}{\frac{\text{D}_2}{\vdash \text{Par}(\ulcorner \rho \urcorner)} \quad \frac{\vdash \rho}{\vdash \rho \vee \neg\rho} \vee\text{-I}_1} \text{LEM-Par-E} \vdash \perp$$

Proposition 5.4. *LPTTU is trivial, and so is the closure under LNC-Un-I and LNC-un-E of any theory extending LPTT⁺.*

Proof. Let ζ be a sentence identical to $\neg\text{Tr}([\zeta]) \wedge \text{Un}([\zeta])$. We reason in LPTTU. We begin by proving $\zeta \wedge \neg\zeta \vdash \perp$:

$$\frac{\frac{\frac{\overline{\zeta \wedge \neg\zeta \vdash \zeta \wedge \neg\zeta}}{\zeta \wedge \neg\zeta \vdash \zeta} \text{SRef}}{\zeta \wedge \neg\zeta \vdash \zeta} \wedge\text{-E}_1 \quad \frac{\overline{\zeta \wedge \neg\zeta \vdash \neg\text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)}}{\zeta \wedge \neg\zeta \vdash \neg\text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)} \text{Def. of } \zeta}{\zeta \wedge \neg\zeta \vdash \text{Un}(\ulcorner \zeta \urcorner)} \wedge\text{-E}_2 \quad \frac{\overline{\zeta \wedge \neg\zeta \vdash \zeta \wedge \neg\zeta}}{\zeta \wedge \neg\zeta \vdash \zeta \wedge \neg\zeta} \text{SRef}}{\zeta \wedge \neg\zeta, \zeta \wedge \neg\zeta \vdash \perp} \text{LNC-Un-E}}{\zeta \wedge \neg\zeta \vdash \perp} \text{SContr}$$

Call this derivation D_0 . We can use it to prove $\neg\zeta$:

$$\frac{\frac{\frac{\overline{\zeta \vdash \zeta}}{\zeta \vdash \zeta} \text{SRef}}{\zeta \vdash \neg\text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)} \text{Def. of } \zeta}{\zeta \vdash \neg\text{Tr}(\ulcorner \zeta \urcorner)} \wedge\text{-E}_1 \quad \frac{\overline{\zeta \vdash \zeta} \text{SRef}}{\zeta \vdash \zeta} \text{SRef}}{\zeta \vdash \neg\zeta} \neg\text{-Tr-E}}{\zeta, \zeta \vdash \perp \vee (\zeta \wedge \neg\zeta)} \neg\text{-E}_w \quad \frac{\overline{\zeta \vdash \zeta} \text{SRef}}{\zeta \vdash \zeta} \text{SRef}}{\zeta \vdash \perp \vee (\zeta \wedge \neg\zeta)} \text{SContr} \quad \frac{\overline{\perp \vdash \perp} \text{SRef}}{\perp \vdash \perp} \text{SRef} \quad \frac{D_0}{\zeta \wedge \neg\zeta \vdash \perp} \text{V-E}}{\zeta \vdash \perp} \neg\text{-I}}{\vdash \neg\zeta} \neg\text{-I}$$

Call the above derivation D_1 . Together with D_0 , it yields a proof of triviality.

$$\frac{\frac{D_0}{\zeta \wedge \neg\zeta \vdash \perp} \text{LNC-Un-I} \quad \frac{\frac{D_1}{\vdash \neg\zeta} \quad \frac{\frac{\frac{D_1}{\vdash \neg\zeta}}{\vdash \neg\text{Tr}(\ulcorner \zeta \urcorner)} \neg\text{-Tr-I} \quad \frac{D_0}{\zeta \wedge \neg\zeta \vdash \perp} \text{LNC-Un-I}}{\vdash \neg\text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)} \wedge\text{-I}}{\vdash \zeta} \text{Def. of } \zeta} \quad \frac{D_1}{\vdash \neg\zeta} \wedge\text{-I}}{\vdash \zeta \wedge \neg\zeta} \text{LNC-Un-E}}{\vdash \perp}$$

Proposition 5.6. *MALLTTP is trivial, and so is the closure under LC-Par-I and LC-Par-E of any theory extending MALLTTP⁺.*

Proof. Let ρ be the sentence $\text{Tr}([\rho]) \rightarrow \text{Par}([\rho])$. We reason in MALLTTP. We first prove ρ , on the assumption that ρ satisfies $\rho \rightarrow (\rho \wedge \rho)$:

$$\frac{\frac{\overline{\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Tr}(\ulcorner \rho \urcorner)}}{\text{Tr}(\ulcorner \rho \urcorner) \vdash \rho} \text{SRef}}{\text{Tr}(\ulcorner \rho \urcorner) \vdash \rho} \text{Tr-E} \quad \frac{\frac{\overline{\rho \vdash \rho} \text{SRef}}{\rho \vdash \text{Tr}(\ulcorner \rho \urcorner)} \text{Tr-I} \quad \frac{\overline{\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Tr}(\ulcorner \rho \urcorner)}}{\rho, \text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SRef}}{\rho, \text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{Cut}}{\rho, \rho \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{Cut}}{\rho, \rho \rightarrow (\rho \wedge \rho) \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SContr}_w \quad \frac{\overline{\rho \vdash \rho} \text{SRef}}{\rho \vdash \rho} \text{SRef}}{\rho \vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)} \text{Def. of } \rho}}{\rho \rightarrow (\rho \wedge \rho), \text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{Cut}}{\rho \rightarrow (\rho \wedge \rho) \vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)} \rightarrow\text{-I}}{\rho \rightarrow (\rho \wedge \rho) \vdash \rho} \text{Def. of } \rho$$

Call this derivation D_0 . We use it to prove that ρ is paradoxical from two occurrences of $\rho \rightarrow (\rho \wedge \rho)$:

$$\frac{\frac{\mathcal{D}_0}{\rho \rightarrow (\rho \wedge \rho) \vdash \rho} \text{Def. of } \rho \quad \frac{\mathcal{D}_0}{\rho \rightarrow (\rho \wedge \rho) \vdash \rho} \text{Tr-I}}{\frac{\rho \rightarrow (\rho \wedge \rho) \vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)}{\rho \rightarrow (\rho \wedge \rho)^2 \vdash \text{Par}(\ulcorner \rho \urcorner)} \rightarrow\text{E}}$$

Call the above derivation D_1 . We now use it to prove that ρ is paradoxical:

$$\frac{\frac{\mathcal{D}_1}{[\rho \rightarrow (\rho \wedge \rho)]^2 \vdash \text{Par}(\ulcorner \rho \urcorner)} \quad \frac{\rho \rightarrow (\rho \wedge \rho) \vdash \rho \rightarrow (\rho \wedge \rho)}{\text{LC-Par-E}} \text{SRef}}{\frac{[\rho \rightarrow (\rho \wedge \rho)]^5 \vdash \perp}{\vdash \text{Par}(\ulcorner \rho \urcorner)} \text{LC-Par-I}}$$

Call the above derivation D_2 . It yields a proof of \perp , and hence of the triviality of MALLTTP:

$$\frac{\frac{\mathcal{D}_2}{\vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SWeak} \quad \frac{\mathcal{D}_2}{\text{Tr}(\ulcorner \rho \urcorner) \vdash \text{Par}(\ulcorner \rho \urcorner)} \text{SWeak}}{\frac{\vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)}{\vdash \rho} \rightarrow\text{I} \quad \frac{\vdash \text{Tr}(\ulcorner \rho \urcorner) \rightarrow \text{Par}(\ulcorner \rho \urcorner)}{\vdash \rho} \text{Def. of } \rho} \wedge\text{-I}}{\frac{\vdash \rho \wedge \rho}{\rho \vdash \rho \wedge \rho} \text{SWeak} \quad \frac{\rho \vdash \rho \wedge \rho}{\vdash \rho \rightarrow (\rho \wedge \rho)} \rightarrow\text{I}} \text{LC-Par-E}}{\frac{\vdash \text{Par}(\ulcorner \rho \urcorner)}{\vdash \perp} \text{LC-Par-E}}$$

Proposition 5.8. STTTU_0 is trivial, and so is the closure under Cut-Un-I and Cut-Un-I of any theory extending STTT_0^+ .

Proof. We reason much in the same way as in the paraconsistent case. We reason in STTTU_0 . We let ζ be identical to $\neg \text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)$ and prove $\zeta \vdash \perp$:

$$\frac{\frac{\frac{\frac{\frac{\zeta \vdash \zeta}{\zeta \vdash \neg \text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)} \text{Def. of } \zeta} \wedge\text{-E}_1}{\zeta \vdash \neg \text{Tr}(\ulcorner \zeta \urcorner)} \neg\text{-Tr-E}}{\zeta \vdash \neg \zeta} \neg\text{-E}_{\text{WC}} \quad \frac{\zeta \vdash \zeta}{\zeta \vdash \zeta} \text{SRef}}{\frac{\zeta, \zeta, \zeta \vdash \perp}{\zeta, \zeta \vdash \perp} \text{SContr} \quad \frac{\zeta, \zeta \vdash \perp}{\zeta \vdash \perp} \text{SContr}}$$

Call this derivation D_0 . We can now assume $\vdash \zeta$, use the conclusion of D_0 (namely $\zeta \vdash \perp$) to derive $\vdash \perp$, and finally discharge our assumptions and categorically conclude $\vdash \text{Un}(\ulcorner \zeta \urcorner)$ via Cut-Un-I:

$$\frac{\frac{\frac{\frac{\overline{\vdash \zeta}^1}{\vdash \neg \text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)} \text{Def. of } \zeta} \wedge\text{-E}_2}{\vdash \text{Un}(\ulcorner \zeta \urcorner)} \quad \frac{\overline{\vdash \zeta}^1}{\vdash \zeta} \quad \frac{\mathcal{D}_0}{\zeta \vdash \perp}}{\frac{\vdash \perp}{\vdash \text{Un}(\ulcorner \zeta \urcorner)} \text{Cut-Un-I, 1}} \text{Cut-Un-E}$$

Call this derivation D_1 . D_0 and D_1 can finally be combined to yield a proof of \perp :

$$\frac{
 \frac{
 \frac{
 \mathcal{D}_0
 }{
 \zeta \vdash \perp
 }
 }{
 \vdash \neg \zeta
 }
 \neg\text{-I}
 }{
 \vdash \neg \text{Tr}(\ulcorner \zeta \urcorner)
 }
 \neg\text{-Tr-I}
 \quad
 \frac{
 \mathcal{D}_1
 }{
 \vdash \text{Un}(\ulcorner \zeta \urcorner)
 }
 }{
 \vdash \neg \text{Tr}(\ulcorner \zeta \urcorner) \wedge \text{Un}(\ulcorner \zeta \urcorner)
 }
 \wedge\text{-I}
 }{
 \vdash \zeta
 }
 \text{Def. of } \zeta
 }{
 \vdash \perp
 }
 \text{Cut-Un-E}$$