Journal of Philosophical Logic (2022) 51:1583–1611
https://doi.org/10.1007/s10992-022-09682-4

The Final Cut

Elia Zardini[1,2]

Received: 9 July 2020 / Accepted: 22 August 2022 / Published online: 14 November 2022

Abstract

In a series of works, Pablo Cobreros, Paul Égré, David Ripley and Robert van Rooij have proposed a nontransitive system (call it '**K3LP**') as a basis for a solution to the semantic paradoxes. I critically consider that proposal at three levels. At the level of the background logic, I present a conception of classical logic on which **K3LP** fails to

I've often been asked why, since I do adopt a nontransitive system in order to solve the paradoxes of vagueness, I don't adopt the nontransitive system to be discussed in this paper in order to solve the semantic paradoxes instead of adopting my favoured, noncontractive system (or why, since, again, I do adopt a nontransitive system in order to solve the paradoxes of vagueness, I don't adopt that simpler nontransitive system in order to solve those paradoxes instead of adopting my favoured, more complex nontransitive system). This paper provides (part of) my answer to those questions. Earlier versions of the material in the paper were presented in 2020 at the BA Logic Group WIP Seminar (University of Buenos Aires); in 2021, at the Conference *Formal Philosophy 2021* in Moscow (Higher School of Economics) and at the 10th SLMFCE 2021 Conference (University of Salamanca); in 2022, at the Current Debates in the Philosophy of Logic Seminar (University of Padua), at the 24th Valencian Philosophy Congress (Jaume I University) and at the 13th Panhellenic Logic Symposium in Volos (University of Thessaly). I'd like to thank all these audiences for very stimulating comments and discussions. Special thanks go to Eduardo Barrio, Carlos Benito, Colin Caret, Vitalij Dolgorukov, Elena Dragalina, Filippo Ferrari, Hartry Field, Michael Glanzberg, Anil Gupta, Volker Halbach, Ulf Hlobil, Antonis Kakas, Ben Martin, José Martínez, Sergi Oms, Federico Pailos, Simone Picenni, Lucas Rosenblatt, Yannis Stephanou, Damián Szmuc, Diego Tájer, Jordi Valor and several anonymous referees. None of them should be held responsible for any remaining mistake or long footnote in the paper. I'm also grateful to the guest editors Eduardo Barrio (again!) and Paul Égré for inviting me to contribute to this special issue and for their extraordinary support and patience throughout the process. I'm additionally grateful to Paul (again!) for his extremely perceptive and fair comments on several versions of what is for him not precisely a sympathetic paper. Work on the paper was supported by the Ramón y Cajal Research Fellowship RYC-2017-22883. Additionally, support from the Basic Research Program of the National Research University Higher School of Economics is gratefully acknowledged. I also benefited from the FCT Project PTDC/FER-FIL/28442/2017 *Companion to Analytic Philosophy 2*, from the Project PID2019-105746GB-I00 of the Spanish Ministry of Science and Innovation *Linguistic Relativity and Experimental Philosophy*, from the Project 2019PIDPID-107667GB-I00 of the Spanish Ministry of Science and Innovation *Worlds and Truth Values: Challenges to Formal Semantics*, from the BBVA-Foundation Project Grants for Scientific-Research Projects 2021 *Unstable Metaphysics* and from the FCT Project 2022.03194. PTDC *New Perspectives on the Objects and Grounds of Structural Rules*.

✉ Elia Zardini
ezardini@ucm.es

Extended author information available on the last page of the article.

vindicate classical logic not only in terms of structural principles, but also in terms of operational ones. At the level of the theory of truth, I raise a cluster of philosophical difficulties for a **K3LP**-based system of naive truth, all variously related to the fact that such a system proves things that would seem already by themselves repugnant, even in the absence of transitivity. At the level of the theory of validity, I consider an extension of the **K3LP**-based system of naive validity that is supposed to certify that validity in that system does not fall short of naive validity, argue that such an extension is untenable in that its nontriviality depends on the inadmissibility of a certain irresistible instance of transitivity (whence the advertised "final cut") and conclude on this basis that the **K3LP**-based system of naive validity cannot coherently be adopted either. At all these levels, a crucial role is played by certain metaentailments and by the extra strength they afford over the corresponding entailments: on the one hand, such strength derives from considerations that would seem just as compelling in a general nontransitive framework, but, on the other hand, such strength wreaks havoc in the particular setting of **K3LP**.

Keywords Classical logic · Naive truth · Naive validity · Nontransitive logics

1 Historical Context

In a series of works (e.g. Ripley [33]; Cobreros et al. [7] respectively), David Ripley as well as Pablo Cobreros, Paul Égré, David Ripley and Robert van Rooij (henceforth, contracting on Ripley, 'CERV') have proposed a *nontransitive* system for solving the *semantic paradoxes*. The system consists of a background system (henceforth '**K3LP**'[1]) that is extended to a system of truth and validity. **K3LP** has fallen out as a particular case (sort of) from the general family of *tolerant logics* that I've introduced in Zardini [40, 41].[2] Without going into many details, this is a family of logics whose value space is a lattice where values are systematically distinguished into *designated* and not, and the latter are in turn distinguished into *tolerated* and not (with designated values' also being considered tolerated). The consequence relation is defined

[1]Following CERV's own usage since Cobreros et al. [6], in the literature, the system is more usually known as '**ST**'. I prefer and use the '**K3LP**'-label both because it makes explicit what the system is in terms of the basic and familiar logics **K3** [24] and **LP** (Asenjo [1], where the unobvious label in turn comes from the "logic of paradox" of Priest [28]) and because the '**ST**'-label brings with it a whole lot of additional theoretical baggage (on different kinds of truth, their relations and foundations) that, as far as I can tell, receives its motivation from the case of vagueness rather than from that of the semantic paradoxes (and, if we abstract away from that baggage by simply retaining the idea that the standard for the premises is higher than that for the conclusions, such an idea, far from picking out **K3LP** in particular, characterises a much more general class of logics, one that extends even beyond the class of logics mentioned in fn 2). Thanks to Paul Égré and an anonymous referee for prompting this clarification.

[2]'(sort of)' because, while I was actually developing a *wider* class of logics (to which **K3LP** belongs without '(sort of)'), 'tolerant logics' was reserved for the *narrower* family of logics within that class *where naive vagueness* (which will be introduced below in the main text and which includes the principle of *tolerance*) *is consistent* [40, p. 339], whereas it is actually not so in **K3LP** (I'll indirectly delve into the philosophical significance of this fact in Section 3 by directly delving into the philosophical significance of the fact that certain other naive theories are inconsistent in **K3LP**).

not as *preservation of designated value from the premises to the conclusions*, but as *connection between designated value in the premises and tolerated value in the conclusions*, which clearly opens up the possibility of failure of transitivity. Different logics in the family are obtained essentially by imposing different constraints on the value space or on the logical operations on it (or both). My own aim in developing tolerant logics was to provide a logic suitable for *naive vagueness*: roughly, the theory according to which *there are both positive and negative cases of a vague property, but the property is also tolerant in that cases that are similar to positive (negative) cases are positive (negative)*. I set myself that aim because I believe that, in the case of vagueness, failure of transitivity is utterly compelling (see Zardini [54] for a more recent statement and defence of this position; in fact, CERV too initially applied their system to vagueness, see e.g. Cobreros et al. [6]). When I then moved to the semantic paradoxes, I did consider (briefly) an application of tolerant logics to them, but I couldn't find a *natural reason* for thinking that transitivity fails in that context and the logics that seemed to emerge didn't have the *desired features* anyway.

Throughout the years, a new reason has become clear to me for why a nontransitive approach to the semantic paradoxes is problematic (see fn 39 for nontransitive approaches different from the **K3LP**-based one). That reason is the main topic of this paper. On the way to that reason, I'll also raise several other issues, which variously connect to the second worry I've mentioned at the end of the last paragraph (all those other issues have analogues in the case of vagueness, but I'll save you the details, some of which are given in Zardini [54], p. 180, fn 26). More in detail, I'll critically consider the proposal of **K3LP** as a basis for a solution to the semantic paradoxes at three levels. At the level of the background logic (Section 2), I present a conception of classical logic on which **K3LP** fails to vindicate classical logic not only in terms of structural principles, but also in terms of operational ones. At the level of the theory of truth (Section 3), I raise a cluster of philosophical difficulties for a **K3LP**-based system of naive truth, all variously related to the fact that such a system proves things that would seem already by themselves repugnant, even in the absence of transitivity. At the level of the theory of validity (Section 4), I consider an extension of the **K3LP**-based system of naive validity that is supposed to certify that validity in that system does not fall short of naive validity, argue that such an extension is untenable in that its nontriviality depends on the inadmissibility of a certain irresistible instance of transitivity (whence the advertised "final cut") and conclude on this basis that the **K3LP**-based system of naive validity cannot coherently be adopted either. At all these levels, a crucial role is played by certain metaentailments and by the extra strength they afford over the corresponding entailments: on the one hand, such strength derives from considerations that would seem just as compelling in a general nontransitive framework, but, on the other hand, such strength wreaks havoc in the particular setting of **K3LP**.

2 The Background Logic

Without going into formal details, **K3LP** can be thought of in at least two ways. *Model-theoretically*, as the relation of logical consequence arising from the *strong*

Kleene valuation scheme by setting the standard for premises to be 1 (as in **K3**) and the standard for conclusions to be $1/2$ (as in **LP**). *Proof-theoretically*, by taking a *standard Gentzen-style sequent system for classical logic* and subtracting from its defining principles the metaentailment[3] of *transitivity*:

(TRANS) If $\Gamma_0 \vdash \Delta_0, \varphi$ and $\Gamma_1, \varphi \vdash \Delta_1$ are valid, $\Gamma_1, \Gamma_0 \vdash \Delta_0, \Delta_1$ is valid.[4]

Either presentation brings out the fact that, in a *standard* language,[5] *the consequence relation of* **K3LP** *coincides with that of classical logic* (since every **K3LP**-countermodel can be transformed into a classical countermodel and since cut is admissible in a standard Gentzen-style sequent system for classical logic), which CERV themselves like to gloss by saying that **K3LP** *is* classical logic (e.g. Cobreros et al. [7], p. 847). A related fact holds for the **K3LP**-based system considered in Section 3 (to the effect that all classically valid entailments are valid in that system, see Ripley [33], pp. 358, 360 for proofs relevant to both facts), which CERV themselves sometimes like to gloss by saying that classical logic *is preserved* by that extension of **K3LP** (e.g. Cobreros et al. [7], p. 853),[6] which, especially in view of the staggering claims made by the latter (for example, the claim, for certain φs, that $\varnothing \vdash \varphi \& \neg\varphi$ is valid, see Section 3), has variously perplexed/amused/annoyed a few other people.

In my view, the ensuing debate has been very fruitful, as it has forced many of us to think harder about what a "logic" is in general and what "classical logic" (or any other philosophically interesting logic) is in particular. By my lights, while it should be emphasised that the idea of "classical logic" (or of many other philosophically

[3] *Terminological point.* There's a fast-growing tendency in the recent literature (e.g. Barrio et al. [2]) of calling '*metainferences*' principles that have some entailments as input and an entailment as output (generalisations of this are possible), where in turn an *entailment* is something to the effect that certain (sentence-like) premises entail certain (sentence-like) conclusions (generalisations of this are also possible). While by now this is probably a lost battle, I'd like to put on the record that I believe that that is unfortunate, since it would not be unfortunate only if calling entailments 'inferences' were not unfortunate, but the latter is indeed unfortunate, since 'inference' in the philosophical literature is predominantly reserved for *the action of drawing certain conclusions from certain premises* rather than for *the fact that certain premises entail certain conclusions*. So understood, *good inferences and valid entailments can drastically come apart*: some good inferences do not correspond to valid entailments (e.g. think of the inference from 'Snow is white' to 'I believe that snow is white', see Evans [13], p. 225) and some valid entailments do not correspond to good inferences (e.g. think of the entailment from the individual sentences in a history book to their conjunction, see Makinson [26]).

[4] Throughout, unqualified $\vdash$ refers to the consequence relation of *real logical consequence* (whichever that is); $\vdash$ qualified as relative to a certain system (e.g. $\vdash_{\mathbf{K3LP}}$) refers to the consequence relation *encoded by that system* (ditto for 'valid' and its relatives).

[5] As an anonymous referee indicated to me, if we consider instead a *nonstandard* language with an additional *logical constant* † *for the 1/2-value*, then, for example, both $\varnothing \vdash_{\mathbf{K3LP}} \dagger$ and $\dagger \vdash_{\mathbf{K3LP}} \varnothing$ are valid, but, for one thing, there is no possible interpretation of † in a classical semantics that validates both those entailments.

[6] Some other times CERV take the opposite view [6, pp. 373, 384], yet some other times they remain more or less neutral (Cobreros et al. [9], p. 1075, although the only explicit reservation they enter against the gloss is one that Cobreros et al. [7], pp. 852–853 essentially discard). Thanks to an anonymous referee for urging me to pay heed to the nuances of CERV's views.

interesting logics) is *flexible* enough as to allow for a *variety of understandings each of which is in certain respects theoretically fruitful*, one very plausible conclusion to draw from that debate is that, for many a philosophically interesting logic, there is at least one theoretically fruitful understanding on which it is *not* reducible to the set of its valid entailments. Going "upwards", some logic, on at least one theoretically fruitful understanding of it, might just as crucially be characterised *by the validity of some metaentailments* (which might not be fully captured by the set of its valid entailments): for example, on at least one theoretically fruitful understanding of classical logic, (TRANS) is arguably just as essential to classical logic as e.g. the entailment of *transitivity of implication* ($\varphi \rightarrow \psi, \psi \rightarrow \chi \vdash \varphi \rightarrow \chi$). Going "downwards", some logic, on at least one theoretically fruitful understanding of it, might just as crucially be characterised *by certain properties of its logical operations* (which might not be fully captured by the set of its valid entailments): for example, on at least one theoretically fruitful understanding of classical logic, that the truth of a negation rules out the truth of what it denies is arguably just as essential to classical negation as the entailment of *explosion* ($\varphi, \neg\varphi \vdash \psi$). Naturally, these two directions can *interact*: some logic, on at least one theoretically fruitful understanding of it, might just as crucially be characterised by the validity of some metaentailments that fully capture certain properties of its logical operations (which might not be fully captured by the set of its valid entailments). I'd like to exemplify this interaction by briefly presenting what I regard as one theoretically fruitful understanding of classical logic.

On this understanding, classical logic includes a particular conception of the logical operations: *operational classicism*. In turn, operational classicism is better understood as having a richer-than-usual conception of which *structural objects* are there. It inherits from the mainstream the idea that sentences can be *combined* as *facts* (as the mainstream invariably does for the premises of an entailment) or as *alternatives* (as the mainstream invariably does for the conclusions of an entailment).[7] It however severs the exclusive connection (in the mainstream) of fact combination with premise combination and of alternative combination with conclusion combination, so that these two modes of combination can *freely interact* both in combining premises and in combining conclusions. For example, letting the comma (,) refer to fact combination and the colon (:) to alternative combination, $\varphi_0 : (\varphi_1, \varphi_2) \vdash \psi_0, (\psi_1 : \psi_2)$ is the form of a possible entailment. Operational classicism also distinguishes two senses of *nothingness*: *positive* nothingness (*nihil privativum*, as is familiar from the

[7]To clarify, the combination of φ and ψ as facts amounts to φ and ψ's *holding together as facts* (i.e. to the holding of *both* φ *and* ψ), whereas the combination of φ and ψ as alternatives amounts to φ and ψ's *holding together as alternatives* (i.e. to the holding of *either* φ *or* ψ). Notice that "alternative facts" (on the usual understanding, not on Kellyanne Conway's) just are alternatives, so that combination as *alternative facts* just is combination as *alternatives*. That however does not necessitate any qualification of the notion of combination as *facts*, since alternative facts are no more facts than, say, alleged facts are. Thanks to an anonymous referee for suggesting this clarification.

idea that a logical truth follows from "nothing")[8] and *negative* nothingness (*nihil negativum*, as is familiar from the idea that an absurdity entails "nothing"),[9] and treats both nothingnesses as *freely combinable*. For example, letting $\circ$ refer to positive nothingness and $\bullet$ to negative nothingness, $\varphi : \circ \vdash \psi, \bullet$ is the form of a possible entailment. Operational classicism finally follows the mainstream in adopting a general format where complexes built up by the previous structural objects and sentences are always further embedded by a *nonembeddable entailment*. Letting Γ and Δ range over nonempty complexes built up by the previous structural objects and sentences, every possible entailment thus has the form $\Gamma \vdash \Delta$.

On the background of this kind of structures, and restricting to the sentential level with finite complexes, operational classicism is then the idea that *logical operations are object-language correlates*[10] *of the relevant structural objects* (cf Zardini [56]

[8]That idea is usually officially represented as a logical truth's following from *the empty set* (or the empty collection with the required fineness of grain: empty multiset, empty sequence or empty whatnot). But, from a philosophical point of view, this official representation can hardly be taken seriously: on what would seem the only philosophically defensible understanding, a logical truth does not follow from the empty set in the same *real* way in which, say, $\varphi \vee \psi$ follows from φ; rather, it follows from the empty set in the same broadly *artificial* way in which, say, it follows from my nose (where the nonartificial element of the representation concerns the fact that a logical truth is guaranteed to hold independently of any premise, a fact that is however equally well represented by means of any object that does not contain any premise). A less usual official representation has it that a logical truth follows from *the logical-truth constant* t. While it does not suffer from the defect of the former representation, this other official representation misrepresents the idea that a logical truth follows *from no premises* by representing it as a logical truth's following *from a certain premise* (i.e. the logical-truth constant); such a representation is also *circular* in that the logical-truth constant is in turn typically explained as being *the (big) conjunction of all logical truths (that do not themselves involve the logical-truth constant)*. And, of course, it does not make much sense to represent the idea that a logical truth follows from "nothing" by adopting a pedantic, Carnap-style [5, pp. 229–233] reading of that construction, according to which there is no premise x such that a logical truth follows from x: such a representation overcomes the problems of the former two representations only at the expense of *manifest falsity*. It is against the background of such difficulties (for which see also Zardini [52], p. 269, fn 34) that it becomes attractive to represent the idea in question *by hypostasising nothingness as a structural object*: by appealing to a notion that would seem to find meaningful application in circumstances of this broad kind (see also fn 9), the resulting representation is not artificial; by appealing to an object that, *qua* structural, is not a premise, the resulting representation is also neither a misrepresentation nor circular; by appealing to something rather than (Carnapian) nothing, the resulting representation is also not manifestly false. While I thus think that the proposed representation captures well the idea in question, I hasten to add that, for reasons better left for another occasion (and partially presented in Zardini [52], pp. 254–259), for many logics of philosophical interest, I don't think that that idea (however represented) can in turn serve to *ground logical truth in logical consequence*. Analogous comments apply to the idea that [an absurdity entails "nothing"] to follow in the main text. Thanks to an anonymous referee for a comment that occasioned these clarifications.

[9]Arguably, the two nothingnesses are not confined to logical theorising—they are systematically used in all areas of thought. Positive nothingness is typically expressed by the phrase '*nothing in particular*' and its relatives, whereas negative nothingness by the phrase '*nothing*' and its relatives. For example, think of the contrast between 'I expect nothing in particular' (which expresses a state of mind where one is *open* to the realisation of any contextually relevant possibility—as when e.g. one is totally agnostic about which parties, if any, will win the elections and be able to form the government) and 'I expect nothing' (which expresses a state of mind where one is *closed* to the realisation of any contextually relevant possibility—as when e.g. one totally believes that no party will win the elections and be able to form the government).

[10]I remain neutral in this paper as to whether, deeper down than the *correlation*, it is the logical operations that *ground* the structural objects or *vice versa* (or neither). I've lifted the neutrality in Zardini [55], pp. 511–515; [57].

for a preliminary, partial analysis of operational classicism, which the current one extends and completes). In detail, the conjunction of φ with ψ correlates *φ and ψ's being combined as facts*. Consequently, $\varphi\&\psi$ is fully intersubstitutable with φ, ψ.[11] The disjunction of φ with ψ correlates with *φ and ψ's being combined as alternatives*. Consequently, $\varphi \vee \psi$ is fully intersubstitutable with $\varphi : \psi$. The negation of φ correlates, *together with φ as a combined fact, with negative nothingness*, and, *together with φ as a combined alternative, with positive nothingness*. Consequently, $\varphi, \neg\varphi$ is fully intersubstitutable with $\bullet$ and $\varphi : \neg\varphi$ is fully intersubstitutable with $\circ$.[12] The implication from φ to ψ correlates with *φ's entailing ψ*. Consequently, keeping in mind the general, entailment-oriented format adopted by operational classicism,

[11] In this context, the relevant notion of full intersubstitutability is such that Γ is fully intersubstitutable with Δ iff $[\ldots\Gamma\ldots \vdash \Theta$ is valid iff $\ldots\Delta\ldots \vdash \Theta$ is, and $\Theta \vdash \ldots\Gamma\ldots$ is valid iff $\Theta \vdash \ldots\Delta\ldots$ is$]$.

[12] Notice that, while binary logical operations like conjunction, disjunction or implication are "*composers*" of their operands and so naturally correlate *by themselves* with structural objects that, being binary, are incomplete in that they *stand only as a connection of a complex with a complex*, a unary logical operation like negation is an "*interactor*" with its operand and so naturally correlates *together with its operand* with a structural object that, being nullary, is complete in that it *already is a self-standing complex*. You might think that nullary logical operations like the logical-truth constant and the *absurdity constant* f also, and more straightforwardly, correlate with positive and negative nothingness respectively (cf the discussion in fn 8). They too arguably do so correlate, but, in a *nontransitive* framework, the issue is actually moot (you'll get a taste of it in Section 3), to the extent that it is negation that *more straightforwardly* has the function of representing together with its operand positive and negative nothingness. Moreover, and more importantly, as we've already touched on in fn 8, the logical-truth constant is typically explained as being the conjunction of all logical truths; but then it is appealing to suppose that it is a *simple* logical truth such as the law of *excluded middle* ($\varphi \vee \neg\varphi$, obviously related to $\varphi : \neg\varphi$) that *more fundamentally* correlates with positive nothingness, and that every other, *complex* logical truth (including the logical-truth constant) only so correlates *derivatively* (ditto, *mutatis mutandis*, for the absurdity constant, which is typically explained as being *the (big) disjunction of all absurdities (that do not themselves involve the absurdity constant)*). Finally, and most importantly, if negation is to be countenanced as a *primitive* logical operation and thereby as having its own correlates among structural objects, the proposal in the main text would seem by far the most *natural* one (whether or not the same structural objects also correlate with the logical-truth constant and with the absurdity constant). (If negation is not so countenanced, a natural option would be to *define it in terms of implication and the absurdity constant*. Keeping fixed the rest of operational classicism and the structural classicism to be introduced in the next paragraph, that would still result in the *consequence relation* of classical logic—though not in a *system* some of whose operational principles are problematic for **K3LP**. While that indicates that, in the relevant respect, **K3LP** is more comfortable with negation as reducible to implication and the absurdity constant rather than with negation as primitive, the understanding of classical logic I'm presenting goes instead for the latter option, partially on the grounds that, letting a *basic* logical truth (absurdity) be any logical truth (absurdity) where what occurs essentially is only one sentence (at most twice) and at most two logical operations (each at most once), it sees in the *interaction between negation and disjunction the generator (besides implication, if that is countenanced as primitive) of basic logical truths* and that it sees in the *interaction between negation and conjunction the generator (besides the interaction between negation and implication, if the latter is countenanced as primitive) of basic absurdities* (where this very last view forecloses the reduction of negation to implication and the absurdity constant).) Thanks to an anonymous referee for causing an overhaul of the material in this fn.

$\varphi \to \psi$ as a conclusion is fully intersubstitutable with $\varphi \vdash \psi$; more precisely, $\Gamma \vdash \Delta : \varphi \to \psi$ is valid iff $\Gamma, \varphi \vdash \Delta : \psi$ is.[13]

Classical logic also includes a particular conception of how structural objects behave: *structural classicism*. In detail, as for *properties* of single structural objects, fact and alternative combinations are *associative* (letting the semicolon (;) ambiguously stand either for uniform occurrences of , or for uniform occurrences of :, $(\Gamma; \Delta); \Theta = \Gamma; (\Delta; \Theta)$), *commutative* ($\Gamma; \Delta = \Delta; \Gamma$) and *idempotent* ($\Gamma; \Gamma = \Gamma$); entailment is *reflexive* ($\Gamma \vdash \Gamma$ is valid) and *transitive* (if $\Gamma \vdash \Delta$ and $\Delta \vdash \Theta$ are valid, $\Gamma \vdash \Theta$ is valid). As for *relations* between different structural objects, fact combination is *selective* over alternative combination (if $\Gamma \vdash \ldots \Delta, (\Theta : \Lambda) \ldots$ is valid, $\Gamma \vdash \ldots (\Delta, \Theta) : \Lambda \ldots$ is; if $\ldots (\Gamma, \Theta) : \Lambda \ldots \vdash \Delta$ is valid, $\ldots \Gamma, (\Theta : \Lambda) \ldots \vdash \Delta$ is), they *juxtapose* (if $\Gamma \vdash \Delta$ and $\Theta \vdash \Lambda$ are valid, $\Gamma; \Theta \vdash \Delta; \Lambda$ is) and are *monotonic* (if $\ldots \Gamma \ldots \vdash \Delta$ is valid, $\ldots \Gamma, \Theta \ldots \vdash \Delta$ is; if $\Gamma \vdash \ldots \Delta \ldots$ is valid, $\Gamma \vdash \ldots \Delta : \Theta \ldots$ is); positive nothingness is an *identity for fact combination* ($\Gamma, \circ$ is fully intersubstitutable with Γ) and negative nothingness is an *identity for alternative combination* ($\Gamma : \bullet$ is fully intersubstitutable with Γ).

If anything, the richer-than-usual conception of structural objects with fact combination and alternative combination's freely interacting both in combining premises and in combining conclusions is helpful in avoiding *an awkwardness arising in the framework of typical logics distinguishing between multiplicative and additive logical operations* (such as *noncontractive* and *nonmonotonic* logics). For the usual consequence relation of such logics can only distinguish the basic principles governing these two kinds of logical operations *by presupposing a certain amount of transitivity that is however foreign to that distinction*. For example, the relevant basic principle governing multiplicative conjunction is supposed to be the metaentailment that, if $\Gamma, \varphi, \psi \vdash \Delta$ is valid, $\Gamma, \varphi \& \psi \vdash \Delta$ is, whereas the relevant basic principle governing additive conjunction is supposed to be the metaentailment that, if $\Gamma, \varphi \vdash \Delta$ is valid, $\Gamma, \varphi \wedge \psi \vdash \Delta$ and $\Gamma, \psi \wedge \varphi \vdash \Delta$ are. That in effect presupposes that there is no failure of transitivity in the route first from $\varphi \& \psi$ to φ and ψ and then from these to Δ, and that there is no failure of transitivity in the route first from $\varphi \wedge \psi$ or $\psi \wedge \varphi$ to φ and then from this to Δ. Such a presupposition is of course understandable in the framework of those logics, since they validate the required amount of transitivity, but *the basic distinction between multiplicative and additive conjunction would not*

[13] It's true that the *sheer fact* that $\varphi \to \psi$ holds is (so much) *weaker* than the *sheer fact* that φ entails ψ (for example, 'If Argentina won the World Cup four times, they won it more times than Italy' holds as it has a false antecedent, but 'Argentina won the World Cup four times' does not entail 'Argentina won the World Cup more times than Italy' even in an extended "World-Cup logic" that contains [World-Cup]-related principles (like e.g. the entailment from 'National team x wins the ith World Cup' to 'National team x has the right to play in the $i+1$th World Cup') that could reasonably be regarded as materially valid). However, you must pay heed both to the fact that, in the full-intersubstitutability principle in question, $\varphi \to \psi$ is a *conclusion* and to the role of *side premises* and *side conclusions*. Once you do that, you'll see that the fact that $\varphi \to \psi$ is *valid under the assumptions of the truth of all side premises* Γ *and of the falsity of all side conclusions* Δ does correlate with the fact that, *under the same assumptions*, φ entails ψ (so that, in the offending example, the correlates are the fact that 'If Argentina won the World Cup four times, they won it more times than Italy' is valid under the (as a matter of fact, correct) assumption of the truth of 'Argentina did not win the World Cup four times' and the fact that, under the same assumption, 'Argentina won the World Cup four times' does entail, by explosion, 'Argentina won the World Cup more times than Italy').

seem to require even that amount of transitivity.[14] To wit, informally, the multiplicative conjunction of φ and ψ entails *both* φ *and* ψ *together*, whereas the additive conjunction of φ and ψ entails *either* φ *or* ψ—that all makes sense even in the total absence of transitivity. The operative richer-than-usual conception of structural objects helps us to capture this informal distinction *directly* (and so without presupposing any amount of transitivity): either (noncontractive-logic case) $\varphi \& \psi \vdash \varphi, \psi$ and $\psi \& \varphi \vdash \varphi, \psi$ are valid but neither $\varphi \wedge \psi \vdash \varphi, \psi$ nor $\psi \wedge \varphi \vdash \varphi, \psi$ are and only $\varphi \wedge \psi \vdash \varphi$ and $\psi \wedge \varphi \vdash \varphi$ are, or (nonmonotonic-logic case) $\varphi \wedge \psi \vdash \varphi$ and $\psi \wedge \varphi \vdash \varphi$ are valid but neither $\varphi \& \psi \vdash \varphi$ nor $\psi \& \varphi \vdash \varphi$ are and only $\varphi \& \psi \vdash \varphi, \psi$ and $\psi \& \varphi \vdash \varphi, \psi$ are.

It's easy to see that, restricting (as in the mainstream) to entailments where premises are combined as facts and conclusions as alternatives, operational classicism *plus* structural classicism validates exactly the classically valid entailments.[15] It's also easy to see that, lifting that restriction, operational classicism *plus* structural classicism validates exactly the entailments that, on any serious understanding, should be considered as "classically valid" within the richer-than-usual conception of structural objects presented in the third paragraph of this section.[16] I take it that the analysis of classical logic as the sum of operational classicism and structural classicism is one theoretically fruitful understanding of classical logic.[17] On such analysis, unsurprisingly given the remark in the second paragraph of this section concerning

[14]A similar presupposition of transitivity is actually present in *all* the operational principles of a *standard sequent calculus*. But, as far as I can tell, that presupposition has *absolutely no philosophical raison d'être* and is instead *only made with the purpose of dispensing with transitivity as a separate principle* (aka *the admissibility of cut*) because that is in turn *advantageous in many proof-theoretic investigations*. In many of those cases, it's indeed a child's play to see what the transitivity-free basic principle should be: for example, the relevant basic principle governing additive conjunction mentioned above (and usually known as *simplification*) can be reformulated to the effect that $\varphi \wedge \psi \vdash \varphi$ and $\psi \wedge \varphi \vdash \varphi$ are valid. My point in this paragraph is that, without employing something like the operative richer-than-usual conception of structural objects, I don't see how the relevant basic principle governing multiplicative conjunction mentioned above in the main text (or, for that matter, the corresponding one governing multiplicative disjunction) could similarly be reformulated as to be made transitivity-free (while also remaining free from other extraneous presuppositions). Thanks to an anonymous referee for discussion of these issues.

[15]*Idea of proof.* This can be proven e.g. by showing that operational classicism *plus* structural classicism entails all the principles of a standard sequent-calculus presentation of classical logic.

[16]*Idea of proof.* This can be proven by first showing that operational classicism *plus* structural classicism validates $\Gamma \vdash \Delta$ iff it validates $\gamma \vdash \delta$ (where γ and δ are the sentences resulting from Γ and Δ by replacing in those complexes fact combination with conjunction, alternative combination with disjunction, positive nothingness with, say, $\varphi \vee \neg\varphi$ and negative nothingness with, say, $\varphi \& \neg\varphi$), and then appealing to the previous result (fn 15) to infer (and give a precise sense to the idea) that operational classicism *plus* structural classicism validates exactly the entailments that, on any serious understanding, should be considered as "classically valid" within the operative richer-than-usual conception of structural objects.

[17]*Both* components are essential for yielding classical logic: there are nonclassical (indeed substructural) logics that vindicate operational classicism (e.g. the logics mentioned in the next paragraph) and there are nonclassical logics that vindicate structural classicism (e.g. an extended multiple-conclusion version of intuitionistic logic *à la* Dragalin [11]). (Indeed, depending on the exact definition of substructurality (cf Zardini [55], pp. 507–516; [57]), there might even be substructural logics that vindicate structural classicism (e.g. the trivial logic where, for every Γ and Δ, it is the case that $\Gamma \vdash \Delta$ is valid can easily vindicate structural classicism and might count as substructural since [it lacks the property of classical logic that, for some Γ and Δ, it is not the case that $\Gamma \vdash \Delta$ is valid and that property does not concern any particular object-language expression]).) Thanks to an anonymous referee for urging this clarification.

(TRANS) and classical logic, transitivity is a defining principle of classical logic, and so, on the analysis, **K3LP** falls short of classical logic. But the analysis does more than simply provide an instance of the point made generally by the remark in the second paragraph of this section concerning (TRANS) and classical logic—for, on the analysis, **K3LP** *falls short of classical logic not only at the structural level, but even at the operational one*. To wit, we'll see in Section 3 that **K3LP** can be extended to a system where, for certain φs, $\varnothing \vdash \varphi\&\neg\varphi$ is valid, and so where, by monotonicity, for every ψ and χ, $\psi \vdash \chi : \varphi\&\neg\varphi$ is valid. According to operational classicism, by the properties of conjunction, that is equivalent with $\psi \vdash \chi : (\varphi, \neg\varphi)$ (assuming for the rest of this discussion that **K3LP** is [extended to the richer-than-usual conception of structural objects presented in the third paragraph of this section] along the lines mentioned in fn 16), which, by the properties of negation, is in turn equivalent with $\psi \vdash \chi : \bullet$. Given that ψ and χ are arbitrary, that is already bad enough; it's even worse once we consider that $\bullet$ as it occurs in operational classicism's characterisation of negation refers to negative nothingness, and so forces the implication from $\psi \vdash \chi : \bullet$ to $\psi \vdash \chi$, thereby making the system officially trivial. Therefore, **K3LP** is incompatible with the combination of the property of conjunction and the property of negation appealed to above (more with the property of negation rather than with the property of conjunction, I submit),[18] and so, on the analysis of classical logic as the sum of operational classicism and structural classicism, falls short of classical logic not only at the structural level, but even at the operational one. On such an analysis, thus falling short on both counts, **K3LP** can actually reasonably be taken to be a *deeply nonclassical* system.

While operational classicism would seem to constitute a *fairly appealing principled, orderly conception of kindred correlations enjoyed by logical operations*, structural classicism would seem to constitute a *rather unappealing unprincipled, disorderly list of sundry properties and relations enjoyed by structural objects*,[19] and so, in this respect, deviations from structural classicism would seem less problematic than deviations from operational classicism, to the extent that a system deviating only in the former way can reasonably be taken to be a *superficially nonclassical* system. It is then interesting to note that, contrary to the deeply nonclassical **K3LP**, there are systems for solving the semantic paradoxes and the paradoxes of vagueness that manage to get rid of structural classicism while upholding operational classicism, and are therefore superficially nonclassical. I myself have developed a broadly *noncontractive* system for solving the semantic paradoxes, which (when extended to the richer-than-usual conception of structural objects presented in the third paragraph of this section along the lines mentioned in fn 16) vindicates operational classicism and

[18]Notice that the fact that the principles *are admissible in* **K3LP** is neither here nor there, for sheer admissibility in a logic does not entail that a principle *characterises* (nor even that it *is compatible with*) the logic: for example, the metaentailment that, if $\varnothing \vdash \exists xy\, x \neq y$ is valid, $\varphi \vee \neg\varphi \vdash \varnothing$ is valid is admissible in classical logic, but it is in a very natural sense incompatible with it, as it is incompatible with classical logic's sound application to many correct target theories. Notice also that, starting from the fact that, in the same system, for certain φs, $\varphi \vee \neg\varphi \vdash \varnothing$ is valid, an analogous argument is available that focuses on disjunction and (the *exhaustivity* of) negation rather than on conjunction and (the *exclusivity* of) negation.

[19]See however Zardini [53] for an attempt at bringing in some order into that mess by imposing on it a certain architecture.

only abandons the *idempotency* component of structural classicism [42, 46, 49, 51, 53]; moreover, in my works mentioned in Section 1 (with the further developments of Zardini [45, 49, 54]), I've developed a broadly *nontransitive* system for solving the paradoxes of vagueness, which (when extended to the operative richer-than-usual conception of structural objects along the lines mentioned in fn 16) vindicates operational classicism and "only" abandons the *transitivity*, *selection* and *juxtaposition* components of structural classicism (which interestingly indicates that the noncontractive system has more of a right to be simply labelled as 'noncontractive' than the "nontransitive" system has to be simply labelled as 'nontransitive').

3 The Theory of Truth

The distinctive nonclassical streak of **K3LP** emerges when *further constraints* on admissible models are imposed (model-theoretically) or when *further principles* are added (proof-theoretically). In particular, *model-theoretically*, we can impose constraints on admissible models so as to have a *truth predicate* T such that φ *and* $T(\ulcorner\varphi\urcorner)$ *always have the same value* (where $\ulcorner\varphi\urcorner$ is a singular term referring to φ). *Proof-theoretically*, we can add rules that guarantee that φ and $T(\ulcorner\varphi\urcorner)$ *are proof-theoretically indiscriminable*. Either way yields a system (henceforth '$\mathbf{K3LP^T}$'[20]) where φ and $T(\ulcorner\varphi\urcorner)$ are *fully intersubstitutable*, which some may take as capturing the essence of *naive truth* (we'll see). Letting λ be identical with $\neg T(\ulcorner\lambda\urcorner)$, the *Liar paradox* still shows that both $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda$ and $\lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ are valid.[21] However, while (TRANS), though not included as a defining principle of **K3LP**, was nevertheless admissible in it, it is no longer admissible in $\mathbf{K3LP^T}$, so that $\varnothing \vdash_{\mathbf{K3LP^T}} \varnothing$ is not valid. That is essentially supposed to be $\mathbf{K3LP^T}$'s solution to the Liar paradox.

I think that such a solution raises several philosophical issues that have not yet been sufficiently highlighted. To start with, the fact that $\varnothing \vdash_{\mathbf{K3LP^T}} \varnothing$ is not valid would seem to *be beside the point*—have you ever heard a compelling informal presentation of the Liar paradox where *the troubling conclusion* is supposed to be that *the empty set entails the empty set* (?!), or that, by *monotonicity*,[22] *everything*

[20]Following CERV's own usage since Cobreros et al. [7], in the literature, the system is more usually known as '**STTT**'. The choice between the '$\mathbf{K3LP^T}$'-label and the '**STTT**'-label is obviously dependent on that between the '**K3LP**'-label and the '**ST**'-label, which I've briefly discussed in fn 1. Thanks to an anonymous referee for prompting this clarification.

[21]*Reason.* By reflexivity, $\lambda \vdash_{\mathbf{K3LP^T}} \lambda$ (i.e. $\lambda \vdash_{\mathbf{K3LP^T}} \neg T\ulcorner\lambda\urcorner$) is valid, and so, by a principle of negation related to the law of *noncontradiction* ($\neg(\varphi\&\neg\varphi)$), $\lambda, T\ulcorner\lambda\urcorner \vdash_{\mathbf{K3LP^T}} \varnothing$ is valid, and hence, by full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$, $\lambda, \lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ is valid, and thus, by the metaentailment of contraction (Section 4), $\lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ (i.e. $\neg T(\ulcorner\lambda\urcorner) \vdash_{\mathbf{K3LP^T}} \varnothing$) is valid, and so, by a principle of negation related to the law of excluded middle, $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\lambda\urcorner)$ is valid, and hence, by full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$, $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda$ is valid.

[22]Throughout, I take monotonicity for granted (as $\mathbf{K3LP^T}$ does). Da Ré [10] develops a solution to the semantic paradoxes based on the denial of that metaentailment: all the points made in this section about $\mathbf{K3LP^T}$ but the very last one apply just as well to his nonmonotonic system, since, by trying to block paradox typically even later than $\mathbf{K3LP^T}$ does, it shares with it the relevant problematic features. Thanks to Damián Szmuc and two anonymous referees for discussion of the role of monotonicity in the semantic paradoxes.

entails everything? Well, maybe you have indeed heard (a couple of times) something like the latter. But, to appreciate how beside the point that too is, suppose that the operative language is quantifier-free and that its only atomic sentence is $T(\ulcorner\lambda\urcorner)$. The Liar paradox can still be run in such an impoverished language, but, in it, the worst thing that e.g. something can entail is presumably something along the lines of $T(\ulcorner\lambda\urcorner)\&\neg T(\ulcorner\lambda\urcorner)$. That, as every other sentence in such an impoverished language, is a logical truth of $\mathbf{K3LP^T}$, and so, by $\mathbf{K3LP^T}$'s lights, in such an impoverished language, that everything entails everything only implies the platitude that everything entails a logical truth. Therefore, by $\mathbf{K3LP^T}$'s lights, in such an impoverished language, its own solution to the paradox as described in the last paragraph (which jumps off the paradoxical train only at the time of not accepting that $\varnothing \vdash \varnothing$ is valid) does not apply (since one should now accept that everything entails everything), so that, if the essence of $\mathbf{K3LP^T}$'s solution to the Liar paradox is really the one described in the last paragraph, by $\mathbf{K3LP^T}$'s lights, in such an impoverished language, there is essentially no paradox that needs a solution. But there is—have you ever heard a compelling informal presentation of the Liar paradox where it all hinges on the fact that the language can express propositions like e.g. the proposition that the Earth is at the centre of the universe? We hit paradox before and without hitting any such proposition.

Some elaboration on what happened in the last paragraph is now in order. To wit, the point does not eventually hinge on the idea that a paradox must have an *apparently false* conclusion, for it needn't (López de Sa and Zardini [25], p. 246; Oms and Zardini [27], p. 8, fn 14; Zardini [55], pp. 494–506). What the point does eventually hinge on is the idea that a paradox must have a conclusion *apparently unsupported by the apparently valid argument*. But, by $\mathbf{K3LP^T}$'s lights, in such an impoverished language, that, say, $\varnothing \vdash T(\ulcorner\lambda\urcorner)\&\neg T(\ulcorner\lambda\urcorner)$ is valid is *not* such a conclusion: every such entailment derivable by monotonicity from $\varnothing \vdash \varnothing$ *is already derivable without the detour through* $\varnothing \vdash \varnothing$, and so is apparently supported *if* $\mathbf{K3LP^T}$ is apparently supported. You can of course accept the point and say that the paradoxical conclusion that needs a solution concerns rather the validity of certain entailments that are derivable *at earlier stages of the paradoxical reasoning*, such as, in the first instance, the conclusion that $\varnothing \vdash T(\ulcorner\lambda\urcorner)$ is valid and $\varnothing \vdash \neg T(\ulcorner\lambda\urcorner)$ is valid—after all, contradictory sentences are always apparently not jointly true. But then the real solution to the paradox is presumably that *contradictory sentences are sometimes actually jointly true*, not that *transitivity fails*. That's both old news for everyone (see Priest [29], who defends an **LP**-based dialetheic approach to the semantic paradoxes that tackles straight on the fact that the paradoxical conclusion that needs a solution is that $\varnothing \vdash T(\ulcorner\lambda\urcorner)$ is valid and $\varnothing \vdash \neg T(\ulcorner\lambda\urcorner)$ is valid) and bad news for lovers of classical negation. I'm happy to grant that it'd be slightly better news *if* the main reason for rejecting contradictory sentences were that, by transitivity, they entail everything. But it isn't. Ask Stephen Read: he'll tell you that, although contradictory sentences don't entail everything, we should reject them because *they cannot be jointly true* (e.g. Read [31], pp. 136–148). I'd add that they cannot be jointly true because one (prominent) understanding of negation (vindicated e.g. by operational classicism) is such that the truth of a negation *rules out* the truth of what it denies (I'll give a fuller statement of such understanding in fn 32 under the label 'complementative negation'). If you want to accept contradictory sentences, you mainly need to face this kind of challenge (as Priest [29, 30] admirably does), where the issue of transitivity is neither here nor there.

In making these points, I should emphasise that, throughout, I (understandably) presume that, even in a $\mathbf{K3LP^T}$-framework, logical truths are... true (in the natural, strong sense, roughly, that, if one accepts that φ is a logical truth, one should accept that φ is true, see Section 4 for a way of explaining, precisifying and generalising such a sense), so that, by naive truth, if one accepts that φ is a logical truth, one should accept φ. When push towards dialetheism (or other nasty consequences)[23] comes to shove, there's actually some temptation by defenders of $\mathbf{K3LP^T}$ to (sort of) resist that, by postulating an *acceptance-like mental state* (call it 'acceptance*') that, in a $\mathbf{K3LP^T}$-framework, it is not the case that one should bear towards logical truths as such (Ripley [34], who actually mostly focuses on acceptance* rather than acceptance), so that, at least with respect to acceptance*, by naive truth, the presumption that logical truths are true would fail. Setting aside the *dubiousness* of such a split in acceptance-like mental states (so far unnoticed in psychology),[24] its *irrelevance* (any type of acceptance of dialetheism is one too many) and its *selfdefeatingness* (in what sense is $\mathbf{K3LP^T}$ the right logic if we rightly flout it when it comes to accepting* things?), the move does little to avoid the shove towards dialetheism. For, correlated to acceptance*, there is a *rejection-like mental state* (call it 'rejection*') that, in a $\mathbf{K3LP^T}$-framework, it is not the case that one should bear towards absurdities as such, so that it is not the case one should reject* that contradictions are sometimes actually true. Dialetheism is an offer that $\mathbf{K3LP^T}$ can't reject*.[25] Be that as it may with acceptance* and rejection* (and rejection), by naive truth, a defender of $\mathbf{K3LP^T}$ must accept that logical truths are true in order to be able to *accept the target logical and truth-theoretic laws* (such as e.g. the law of excluded middle[26] and the principle of correlation to be introduced in the second next paragraph) and thereby *have a view*

[23] In fact, the nasty consequences are as nasty as they can be, since, barring quibbles about $a = a$ and its relatives, by courtesy of λ and its relatives every unrestricted logical and truth-theoretic law has instances whose negations are logical truths in $\mathbf{K3LP^T}$. In the rest of this paragraph, dialetheism thus serves merely as a placeholder for any such ghastly statement.

[24] Incurvati and Schlöder [23] very plausibly defend the existence of a speech act whose effect is to present a proposition *as a possibility* (and, somewhat infelicitously, call it 'weak assertion'). However, it's clear that the existence of such a speech act provides absolutely no evidence in favour of the existence of acceptance*. Firstly, not every kind of *speech act* corresponds to a kind of *mental state* (think promises!). Secondly, "weak assertion" is not even something in the ballpark for being *closed under conjunction*, while acceptance* is supposed to be. Thirdly, *logical truths are certainly a possibility*, and so "weak assertion" should be borne towards them. Thanks to Filippo Ferrari and Diego Tájer for drawing my attention to "weak assertion".

[25] Not that the other rejection-like mental state fares better: in a $\mathbf{K3LP^T}$-framework, one should bear it towards absurdities as such, and so one should bear it towards virtually every unrestricted logical and truth-theoretic law, since again, barring quibbles about $a = a$ and its relatives, by courtesy of λ and its relatives every unrestricted logical and truth-theoretic law has instances that are absurd in $\mathbf{K3LP^T}$.

[26] On this point, at least with respect to acceptance*, any sense in which "**K3LP** is classical logic" comes across as particularly idle: what's the point of accepting* classically valid *entailments* if one doesn't even accept* classically valid *laws*?

that is at least a minimally fit contender in the debate on the semantic paradoxes (see also Zardini [51], pp. 319–320, fn 43 on these issues).[27]

Moreover, the fact that $\varnothing \vdash_{\mathbf{K3LP^T}} \varnothing$ is not valid would also seem to *come too late*. For one thing, $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda$ and $\lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ are still both valid: λ is a *logical truth* but also an *absurdity*. In other words, *something must be the case while at the same time it cannot be the case*. But such a circumstance itself is something *that cannot be the case* (indeed, that's exactly the paradox!). Therefore, $\mathbf{K3LP^T}$ *cannot be correct* (indeed, it just reinstates the paradox!).[28]

For another thing, an absolutely central feature of naive truth is the principle of *correlation* between *reality* and *truth*, which, subject to qualifications that needn't detain us here (but that did detain Zardini [48]), in its *convergence-oriented* version can be formulated as $\varphi \leftrightarrow T(\ulcorner\varphi\urcorner)$. The Liar paradox is an argument against (the convergence-oriented version of) that compelling principle, allegedly showing that it is absurd (or, if you really prefer, allegedly showing that it entails that geocentrism is true). But $\lambda \leftrightarrow T(\ulcorner\lambda\urcorner) \vdash_{\mathbf{K3LP^T}} \varnothing$ is indeed valid. Therefore, $\mathbf{K3LP^T}$ *condemns naive truth to the Liar paradox just as the grimmest classical logician does*: for $\mathbf{K3LP^T}$ just as well, correlation is absurd and entails that geocentrism is true.[29] And the second conjunct entails that, keeping fixed the normative principle mentioned in

[27] Notice that a defender of $\mathbf{K3LP^T}$ cannot obviate to the last point by making the move of saying that the target law is accepted *on other grounds*. For then (setting aside the issue that, in such a case, the background logic and theory of truth as they stand would be woefully incomplete and that the overall theory would only be adequately represented by a stronger system that also incorporates those "other grounds"), by a normative principle that is just as plausible in a nontransitive framework to the effect that, *if one accepts a sentence not because it is the conclusion of an entailment of type X that one knows to be valid and all of whose premises one accepts, one should accept what one knows are the logical consequences of that sentence* (where type X is the type of entailment that features in the relevant failures of (TRANS), see Zardini [50], pp. 255–265), since one accepts that the target unrestricted law holds (presumably, on the move in question, not because it is the conclusion of an entailment of type X) and knows that it entails f, one should accept f (of course, the normative principle would require a lot of fine-tuning in order to deal with a host of well known issues arising from its present formulation, but I take it that none of those issues is relevant for the particular applications we're envisaging). Notice also that all the points made in this paragraph concerning *acceptance* of *logical truths* have analogues concerning *rejection* of *absurdities*. Thanks to Ulf Hlobil for propounding to me the move discussed in this fn, to Colin Caret and Ben Martin for then pressing me on the normative principle and, wrapping up a rather convoluted paragraph, to Federico Pailos for spurring me to consider the acceptance* twist in the first place.

[28] The point has been made harder to appreciate by the typically ⊢-*centric* setting adopted in these discussions, which makes it look as though the position in question could be made coherent by denying (TRANS). However, if we abandon that setting (cf fn 8) and, instead of considering the claim that $\varnothing \vdash \lambda$ and $\lambda \vdash \varnothing$ *are both valid*, consider the consequent claim that λ *both holds and is absurd* (for, generally, the claim that $\varnothing \vdash \varphi$ is valid licences the claim that φ holds and the claim that $\varphi \vdash \varnothing$ is valid licences the claim that φ is absurd), it becomes easier to appreciate that the position in question remains incoherent (or, if you wish, Zen-sounding) never mind whether (TRANS) is denied.

[29] In a natural sense, it is not even possible to "add" correlation to $\mathbf{K3LP^T}$ while preserving nontriviality. For, in a natural sense (adding *as an axiom licensing entailments*), "adding" φ to a system $\mathbf{S}$ results in a system $\mathbf{S}^\varphi$ such that, if $\Gamma, \varphi \vdash_{\mathbf{S}} \Delta$ is valid, $\Gamma \vdash_{\mathbf{S}^\varphi} \Delta$ is valid. In that natural sense, for $\mathbf{K3LP^T}$ just as well, it is not even possible to "add" correlation to it while preserving nontriviality. What is possible of course is to "add" correlation to $\mathbf{K3LP^T}$ in the less natural sense (adding *as a theorem*) in which "adding" φ to a system $\mathbf{S}$ results in a system $\mathbf{S}^\varphi$ such that $\varnothing \vdash_{\mathbf{S}^\varphi} \varphi$ is valid (indeed, in that sense, correlation has already been "added" in $\mathbf{K3LP^T}$ itself!). We'll take up the double-edged significance of this latter fact below in the main text.

fn 27, $\mathbf{K3LP^T}$ *cannot be adopted by anyone who accepts correlation not because it is the conclusion of an entailment of type X*, and so *a fortiori by anyone who*—as many theorists of truth since Plato (*Cratylus*, 385b, though, like Tarski [36], pp. 342–343, you might be more familiar with Aristotle's *Metaphysica*, Γ, 7, 1011b27) are wont to do—*accepts correlation as a basic principle* (by fn 25, analogous points apply for virtually every other unrestricted logical and truth-theoretic law). It's true that $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda \leftrightarrow T(\ulcorner\lambda\urcorner)$ is also valid, and so, contrary to the classical logician, $\mathbf{K3LP^T}$ also maintains that correlation holds. But that only makes things worse (if possible), for it means that $\mathbf{K3LP^T}$ maintains that *something holds that entails that geocentrism is true*—in a very natural sense, that *there is something (the correlation between reality and truth) that makes geocentrism true*.[30] With all due respect for what is in some *technical* respects an interesting system, those are dumbfounding statements, which constitute abundant reason for conclusively rejecting $\mathbf{K3LP^T}$ in its envisaged *philosophical* applications—one can hardly hope for a more conclusive refutation of a theory of truth than a proof that it is committed to a truthmaker for geocentrism! Notice that it's no use to say that, being somewhat "soft", the entailment does not really guarantee, given correlation, the truth of geocentrism, for then, by the same token, one should conclude that, being somewhat "soft", the entailment $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda \leftrightarrow T(\ulcorner\lambda\urcorner)$ does not really guarantee, given anything, that correlation holds.

Similarly, in its *nondivergence-oriented* version, subject to qualifications that needn't detain us here (the same ones as those alluded to in the last paragraph), correlation can be formulated as $\neg(\varphi \& \neg T(\ulcorner\varphi\urcorner))$ and $\neg(T(\ulcorner\varphi\urcorner) \& \neg\varphi)$. Of course, those are also absurd etc. in $\mathbf{K3LP^T}$ (and in virtually every *nondialetheic* approach to the semantic paradoxes except mine, see Heck [20]; Zardini [43] for the details of the argument), but I'm not going to belabour that. Rather, notice that they are supposed to express the informal ideas that *it cannot be the case that what a sentence says is the case without the sentence's being true* and that *it cannot be the case that a sentence is true without what it says being the case*. But $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda \& \neg T(\ulcorner\lambda\urcorner)$ and $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\lambda\urcorner) \& \neg\lambda$ are indeed valid. Therefore, according to $\mathbf{K3LP^T}$, it is indeed the case that what a sentence says is the case without the sentence's being true and it is indeed the case that a sentence is true without what it says being the case (this is really an argument against virtually every *dialetheic* approach, whether the background logic be **K3LP**, or **LP**, or what have you). It is tempting to reply that $\varphi \& \neg\psi$ is *not strong enough* in $\mathbf{K3LP^T}$ to mean that φ holds without ψ's holding (more because of the inadequacy of negation rather than because of any inadequacy of conjunction, I submit). But, if this is not meant by that, it is not meant by anything in $\mathbf{K3LP^T}$. And, if we try to introduce conjunction-like and negation-like logical operations $\&^*$ and $\neg^*$ strong enough so that $\varphi \&^* \neg^* \psi$ does mean that φ holds

[30] As has been observed *ad nauseam* (e.g. Restall [32]), the entailment from x's existence to φ's truth is not always a good reason for thinking that, in a very natural sense, x makes φ true. But it is not a good reason in those cases where φ's truth is entailed by everything (in spite of Restall [32], p. 333's mystical feeling to the effect that "there is something quite touching in the view that every particle in the universe (and everything else besides!) is a witness to all necessary truths"), and the truth of geocentrism is far from being so.

without ψ's holding, they'll still have to obey the compelling "*with-or-without-you*" law $(\mathrm{t}\&^*\varphi) \vee (\mathrm{t}\&^*\neg^*\varphi)$ (either the logical-truth constant holds together with φ's holding or without it) and the compelling "*lonely-together*" metaentailment from $\varnothing \vdash \varphi$ and $\varnothing \vdash \neg^*\psi$ to $\varnothing \vdash \varphi\&^*\neg^*\psi$ (if it is a logical truth that φ holds and a logical truth that $\neg^*\psi$ holds, it is a logical truth that φ holds without ψ's holding). The above argument will then kick in (for a λ^* identical with $\neg^*T(\ulcorner\lambda^*\urcorner)$).[31] What's the point of trying to make the notion of naive truth expressible at the cost of making the notion of *withoutness*—which crucially enters in the formulation of the nondivergence-oriented version of correlation—*inexpressible*? Don't leave naive truth without 'without'![32,33]

At this point, let me undertake what might strike you as a bit of a digressive (3-paragraph long) rant but which will end up uncovering (in the third next paragraph) another philosophical shortcoming of $\mathbf{K3LP^T}$. Given how badly $\mathbf{K3LP^T}$ treats correlation, you might wonder how almost every well established researcher in the area could seriously consider it among the candidate systems of naive truth. I'm pretty sure that's because $\mathbf{K3LP^T}$ validates the *full intersubstitutability* of φ with $T(\ulcorner\varphi\urcorner)$ (henceforth simply 'full intersubstitutability'). But that principle is much stronger than naive truth and in fact variously inconsistent with it. For one thing, if there is *no fact of the matter* whether what φ says is the case, φ does not correlate with the facts (for there are no relevant facts to begin with), and so presumably does not correlate with reality (for presumably how reality is boils down to what facts there are), and hence, by naive truth, is not true, but, *pace* full intersubstitutability, $\neg\varphi$ doesn't follow (for example, we may assume with Field [14] that there is no fact of the matter

[31] *Reason.* Assuming throughout that t is an identity also for $\&^*$, by the with-or-without-you law $\varnothing \vdash \lambda^* \vee \neg^*\lambda^*$ is valid, and so, by full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$, $\varnothing \vdash \lambda^*\vee\neg^*T(\ulcorner\lambda^*\urcorner)$ (i.e. $\varnothing \vdash \lambda^*\vee\lambda^*$) is valid, and hence, by *idempotency* of $\vee$ (φ is fully intersubstitutable with $\varphi \vee \varphi$), $\varnothing \vdash \lambda^*$ (i.e. $\varnothing \vdash \neg^*T(\ulcorner\lambda^*\urcorner)$) is valid. Therefore, by the lonely-together metaentailment, $\varnothing \vdash \lambda^*\&^*\neg^*T(\ulcorner\lambda^*\urcorner)$ is valid. Similarly, by the with-or-without-you law, $\varnothing \vdash \neg^*\lambda^*\vee\neg^*\neg^*\lambda^*$ is valid, and so, by full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$, $\varnothing \vdash \neg^*\lambda^*\vee\neg^*\neg^*T(\ulcorner\lambda^*\urcorner)$ (i.e. $\varnothing \vdash \neg^*\lambda^*\vee\neg^*\lambda^*$) is valid, and hence, by idempotency of $\vee$, $\varnothing \vdash \neg^*\lambda^*$ is valid. Since, by full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$, the first part of this argument also yields that $\varnothing \vdash T(\ulcorner\lambda^*\urcorner)$ is valid, by the lonely-together metaentailment $\varnothing \vdash T(\ulcorner\lambda^*\urcorner)\&^*\neg^*\lambda^*$ is valid.

[32] In essence, this argument from 'without' is (what I regard as) a particularly strong version of the argument [against dialetheic approaches to the semantic paradoxes] based on *complementative negation* (i.e. *exhaustive* and *exclusive* negation): the negation that holds iff what it denies *somehow or other* (exhaustion) *fails* (exclusion) to hold.

[33] Perhaps, it might be interesting to investigate an application to the semantic paradoxes of a *variation* of **K3LP** where, roughly, *the standard for the premises is allowed to be 1/2 if the premises can get that value but not value 1* (something along the lines of the system **PrT** discussed by Cobreros et al. [8] in application to the paradoxes of vagueness). Such an application would improve on some, but not all, the problems discussed in this section (by—perhaps disappointingly—wearing its dialetheism more on its sleeve, so that e.g. $\lambda\&\neg\lambda \vdash \varnothing$ would no longer be valid—thereby raising the question of why not go for the real dialetheist thing of Priest [29]). Without engaging in an extended discussion of a system whose fine details could still be tinkered with, two of its main aspects would already seem to make it unfit for the purposes of the discussion in this paper. Firstly, the system is *nonmonotonic* in such a way as to encode a *nondeductive* consequence relation, whereas what we're looking for is a system that encodes a *deductive* consequence relation. Secondly, by assigning $1/2$ to the relevant paradoxical sentences, the system *does not validate* the principle of manifestation (right to left, see Section 4), whereas what we're looking for is a system that *does validate* manifestation.

whether Newtonian mass is relativistic mass, so that ‘Newtonian mass is relativistic mass’ does not correlate with the facts, and so presumably does not correlate with reality, and hence, by naive truth, is not true, but, *pace* full intersubstitutability, ‘Newtonian mass is not relativistic mass’ doesn’t follow). For another thing, another feature of naive truth is that, assuming φ, it is the case that [φ is true *because* what φ says is the case] (for example, ‘Snow is white’ is true because snow is white). A sentence is true because the reality it talks about is in a certain way. But, *pace* full intersubstitutability, for the vast majority of φs (Zardini [53], pp. 176–177, fn 47 provides one kind of exception), it is not the case that [φ is true because φ is true], nor, if what φ says is that P, is it the case that [P because P], let alone is it the case that [P because φ is true] (for example, it is not the case that [‘Snow is white’ is true because ‘Snow is white’ is true], nor is it the case that [snow is white because snow is white], let alone is it the case that [snow is white because ‘Snow is white’ is true]).[34]

In this connection, you might pause and quite reasonably ask what I mean by ‘naive truth’ and its relatives. Quite generally, by ‘naively F’, I mean something like ‘*really* F’, in turn in the sense (“[lower-order]-wise”) of *what our idea of being F compels us to recognise as having the property of being F or as not having the property of being F* and (“[higher-order]-wise”) of *what properties our idea of being F compels us to recognise as being had by the property of being F or as not being had by the property of being F*—either *categorically* or *hypothetically* under a certain supposition—*sometimes in spite of the fact that such recognition is not licenced by our standard formal theories of being F* (typically because those theories are inconsistent with it). Downstream, I take it that this gloss justifies what I’ve claimed so far about naive truth, including its inconsistency with full intersubstitutability (‘Newtonian mass is relativistic mass’ is not *really* true; ‘Snow is white’ is *really* true because snow is white). Upstream, I take it that this gloss is the one implicitly at work in paradigmatic discussions of “naive truth” and, taking historical priority, in paradigmatic discussions of “naive sets”.[35,36]

Given this, one might wonder how virtually every well established researcher in the area could seriously consider full intersubstitutability a desideratum of a system of naive truth. I’m pretty sure that’s because *deflationistic* propaganda has been

---

[34]To stress, the problem is not that, given full intersubstitutability, we don’t get the wonderful principle that [φ is true because what φ says is the case] (for we might get that, see Horwich [22], pp. 104–105). The problem is rather that, if we get that (as we should), given full intersubstitutability we also get the dreadful principles that [φ is true because φ is true], that, if what φ says is that P, it is the case that [P because P] and that [P because φ is true]. If you’re now tempted to restrict full intersubstitutability to rule out ‘because’-contexts, you’re ready for fn 37.

[35]Unsurprisingly, then, naive membership would seem inconsistent with full intersubstitutability of ‘x is F’ with ‘x belongs to the set of Fs’ (Newtonian mass does not *really* belong to the set of objects identical with relativistic mass; snow *really* belongs to the set of white objects because snow is white).

[36]Having noted all this, I don’t mean to rule out that there are other interesting notions of truth that it may also be reasonable to label as ‘naive’: natural-language meanings, even in a philosophical context, are typically flexible enough, and ‘naive’ is no exception (for example, one could reasonably label fully intersubstitutable truth as ‘naive’ on the grounds that to treat $T(\ulcorner\varphi\urcorner)$ as tantamount to φ is indeed in many respects a very simple and ingenuous way of proceeding). However, I do take it that the notion of truth I’m focusing on is prominent for its utter compellingness. I think it is the only one that is so hefty as to have a chance to tilt the balance in its favour when paradox weighs it against classical logic.

so efficient in the last decades. But, in the only well defined sense I'm aware of, deflationism about truth is the wildly implausible doctrine that *truth merely has a universal expressive function*,[37] driven by the wildly implausible research programme of *explaining content without recourse to the notion of truth* (in favour of which, however, there are deep and challenging philosophical arguments, see Field [15]). Contrary to *naive* truth, *deflationary* truth—and every kind of truth that, possibly among its other functions (and despite the arguments of Zardini [48]), is also supposed to have a universal expressive function—does indeed require full intersubstitutability [16, pp. 205–210].[38] It is surprising how heavily, in the recent past, the

[37] Going back to the point made in the second last paragraph concerning the *dependence of truth on reality*, recall that deflationists typically restrict full intersubstitutability to *nonattitudinal* contexts (after all, a nihilist about truth still believes that snow is white but does no longer believe that 'Snow is white' is true). Give them an inch, and they'll take a mile: they might now be tempted to reply to that point by claiming that full intersubstitutability should more generally be restricted to something like *nonhyperintensional* contexts. However, that claim is untenable *precisely by deflationistic lights*: by those lights, while it may to some extent be plausible that there is no need for an expressive device in attitudinal contexts, it is totally clear that there is every need for an expressive device in many hyperintensional contexts, including dependence contexts: for example, one might want to consider the thought that Fermat's last theorem holds because 0 is not a successor, successorship is injective, 0 is an identity for addition..., which, by deflationistic lights, one can do by entertaining the thought that Fermat's last theorem holds because every axiom of first-order **PA** is true. As a consequence of its being systematically inconsistent with the dependence of truth on reality, deflationism—and, more generally, full intersubstitutability—cannot account for very simple facts about truth such as e.g. the fact that 'Snow is white or grass is blue' is true because snow is white, facts that would seem in the same ballpark as those deflationists typically boast of being able to account for: what's the point of accounting for the fact that 'Snow is white or grass is blue' is true *if* snow is white in such a way (i.e. by appealing to full intersubstitutability) that makes it impossible to account for the very related fact that 'Snow is white or grass is blue' is true *because* snow is white? By contrast, given a couple of plausible assumptions, the antideflationistic principle of dependence of truth on reality stated in the second last paragraph can account for that fact: since 'Snow is white or grass is blue' is true, by the dependence of truth on reality 'Snow is white or grass is blue' is true because snow is white or grass is blue; given the plausible assumption that, if [snow is white or grass is blue] and grass is not blue, [snow is white or grass is blue] because snow is white (and since the assumption's antecedent holds), [snow is white or grass is blue] because snow is white; given the plausible assumption that the relevant instance of transitivity of dependence holds, 'Snow is white or grass is blue' is true because snow is white.

[38] Gupta and Standefer [19] interestingly argue that what is only required is a weaker version of *uniform intersubstitutability*, according to which, roughly, in a sentence, all occurrences of $T(\ulcorner\varphi_0\urcorner)$, $T(\ulcorner\varphi_1\urcorner)$, $T(\ulcorner\varphi_2\urcorner)\ldots$ *en bloc* are intersubstitutable with the corresponding occurrences of φ_0, φ_1, $\varphi_2\ldots$ *en bloc*. While I think that Gupta and Standefer are right that, in many cases, the expressive function of truth only requires uniform intersubstitutability, I also think that they overlook how crucial full intersubstitutability is for the *universal* expressive function of truth. To wit, to take a basic case of the form 'If 'φ' is true, ψ', as Gupta and Standefer themselves recognise, uniform intersubstitutability does not deliver its equivalence with 'If φ, then ψ' *if ψ itself contains some truth attribution* (a case that must be countenanced not only because, in general, a *universal* expressive function is being considered, but also because, in particular, there are well known reasons—i.e. *ignorance* and *finitude*—for why the relevant truth attribution might *not be feasibly eliminable*). Gupta and Standefer propose 'If 'φ' is true, 'ψ' is true' as a replacement candidate for uniform intersubstitutability to deliver its equivalence with 'If φ, then ψ', but that too will not deliver all the target equivalences *if φ in turn contains truth attributions used to express the target hypotheses* (for example, it might be the case that φ uses the truth predicate to express the thought that some theory in a certain class Th_0 holds, where some theory in Th_0 in turn uses the truth predicate to express the thought that some theory in a certain further class Th_1 holds, where some theory in Th_1 yet in turn uses the truth predicate to express the thought that some theory in a certain yet further class Th_2 holds...). Thanks to Hartry Field and Anil Gupta for probing exchanges on these issues.

agenda of nonclassical theories of truth, with its focus on full intersubstitutability, has in effect been influenced by such extreme philosophical ideas.

That said, I do think that there is a more roundabout reason (whose presentation goes beyond the scope of this paper) for having full intersubstitutability as a desideratum of a system of naive truth, a reason that has to do with the *selfreferential* (for want of a better term, see Yablo [39], p. 340) *paradoxes of higher-order quantification* (see Zardini [56] for the details). A less roundabout reason (whose presentation does fall within the scope of this paper) is that *a lot of* applications of full intersubstitutability remain compelling and we can stipulate that the language does not contain any of the expressive resources that make it fail, so that *all* applications of full intersubstitutability become compelling.[39] Notice however that those applications are compelling *in virtue of their being grounded in (typically, the convergence-oriented version of) correlation* (for example, ask yourself why you believe that 'Snow is white or grass is green' entails ' 'Snow is white' is true or grass is green'). *The holding of full intersubstitutability in a system of naive truth should therefore be grounded in the holding of correlation in the system*. But this requirement is resoundingly *not* met by $\mathbf{K3LP^T}$ (as evidenced by the way in which I've introduced such an extension of **K3LP** in the first paragraph of this section): in general, in $\mathbf{K3LP^T}$, there's precious little that is constrained about the behaviour of φ and ψ simply by the fact that $\varnothing \vdash_{\mathbf{K3LP^T}} \varphi \leftrightarrow \psi$ is valid.

Finally, the fact that $\varnothing \vdash \varnothing$ should be valid in $\mathbf{K3LP^T}$ would seem anyway *ultimately inescapable*. Let me explain. If $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda$ is valid, λ is a logical truth (I'm focusing on λ just for concreteness: the argument I'm developing works with many other logical truths of $\mathbf{K3LP^T}$, see fn 25). Since, assuming its typical explanation, the logical-truth constant is the conjunction of all logical truths, it has λ among its conjuncts. Therefore, since even a defender of $\mathbf{K3LP^T}$ accepts that *a conjunction entails everything any of its conjuncts entails* (for it is indeed the case that, if $\Gamma, \varphi \vdash_{\mathbf{K3LP^T}} \Delta$ is valid, $\Gamma, \varphi \& \psi \vdash_{\mathbf{K3LP^T}} \Delta$ and $\Gamma, \psi \& \varphi \vdash_{\mathbf{K3LP^T}} \Delta$ are), and since $\lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ is valid, $\mathsf{t} \vdash \varnothing$ should be valid in $\mathbf{K3LP^T}$. As in the case of $\mathbf{K3LP^T}$'s declaring correlation absurd (sixth paragraph of this section), that's bad enough: it's a sort of ultrairrationalist position according to which *the laws of logic (i.e. the logical truths) cannot be the case*. Even worse, we reach total triviality by adding the compelling metaentailment of t-*deletion*/t-*insertion* that $\Gamma, \mathsf{t} \vdash \Delta$ is valid iff $\Gamma \vdash \Delta$ is: left to right (and crucially for the present argument), *entailment in virtue of the laws of logic*

[39] Starting from Weir [37], Alan Weir has also developed a nontransitive approach to the semantic paradoxes (see Weir [38] for a recent version), but his resulting system does not validate full intersubstitutability and is thus subject to the two problems just mentioned in the main text (as well as to those developed in Zardini [56]). In fact, given full intersubstitutability and a classical behaviour of negation and implication, every nontransitive (but contractive and reflexive) system is going to share the relevant features of **K3LP** this paper focuses on: for example, since, by reflexivity, $\lambda \vdash \lambda$ is valid, by the classical behaviour of negation both $\varnothing \vdash \lambda, \neg\lambda$ and $\lambda, \neg\lambda \vdash \varnothing$ are valid, and so, by full intersubstitutability, both $\varnothing \vdash \lambda, \lambda$ and $\lambda, \lambda \vdash \varnothing$ are valid, and hence, by contraction, both $\varnothing \vdash \lambda$ and $\lambda \vdash \varnothing$ are valid. Therefore, the discussion in the paper is not merely a criticism of the specific **K3LP**-based approach—it extends to every nontransitive (but contractive and reflexive) approach that validates full intersubstitutability and a classical behaviour of negation and implication.

is entailment[40] and, right to left (but not crucially for the present argument), *entailment is entailment in virtue of the laws of logic*.[41] Therefore, as promised, $\varnothing \vdash \varnothing$ should be valid in $\mathbf{K3LP^T}$.[42]

[40]To appreciate just how compelling the idea is that, *working by logical laws, entailment does not need them as premises*, consider a similar case: if Γ *physically* determines Δ in virtue of the laws of *physics*, Γ physically determines Δ *tout court*. Just so, if Γ *logically* determines (i.e. entails) Δ in virtue of the laws of *logic*, Γ entails Δ *tout court*. Interestingly, one might have thought that $\mathbf{K3LP^T}$ vindicates the idea *at least in the case of the entailment from the antecedent to the consequent of an implicational logical law*, in that, if $\varnothing \vdash_{\mathbf{K3LP^T}} \varphi \to \psi$ is valid, not only $\varphi, \varphi \to \psi \vdash_{\mathbf{K3LP^T}} \psi$ is but also $\varphi \vdash_{\mathbf{K3LP^T}} \psi$ is. However, in that case, the idea that *entailment does not need logical laws as premises* really requires that, if $\varnothing \vdash_{\mathbf{K3LP^T}} \varphi \to \psi$ is valid, not only $\varphi, \varphi \to \psi \vdash_{\mathbf{K3LP^T}} \psi$ and $\varphi \vdash_{\mathbf{K3LP^T}} \psi$ be so but also, if $\varnothing \vdash_{\mathbf{K3LP^T}} \varphi$ too is valid, $\varnothing \vdash_{\mathbf{K3LP^T}} \psi$ be so—a requirement that cannot be satisfied by $\mathbf{K3LP^T}$ (as can be gleamed from fn 46). (Should one fall back, in that case, on the weaker idea that *entailment does not need the implicational logical law as a premise*, if full intersubstitutability is understood strongly so that the content of $T(\ulcorner\varphi\urcorner)$ is tantamount to that of φ (as befits deflationism) that idea really requires that, if $\varnothing \vdash_{\mathbf{K3LP^T}} \varphi \to \psi$ is valid, not only $\varphi, \varphi \to \psi \vdash_{\mathbf{K3LP^T}} \psi$ and $\varphi \vdash_{\mathbf{K3LP^T}} \psi$ be so but also, if $\varphi = T(\ulcorner\varphi \to \psi\urcorner)$, it be the case that $\varnothing \vdash_{\mathbf{K3LP^T}} \psi$ is valid—also a requirement that cannot be satisfied by $\mathbf{K3LP^T}$ (as can also be gleamed from fn 46).) Thanks to Lucas Rosenblatt for a question that originated the developments in this fn.

[41]A similar point can be made without passing through the logical-truth constant, by simply focusing on a sentence such as λ that is both a logical truth and an absurdity (setting aside now the problem noted in the fifth paragraph of this section) and noting that, *qua* logical truth, it does not obey its own version of deletion (in the main text, I've preferred to pass through the logical-truth constant because doing so shows the incompatibility in $\mathbf{K3LP^T}$ of two standard and attractive theses about the logical-truth constant: that it is the conjunction of all logical truths and that it obeys t-deletion). Égré [12], pp. 32–35 notes the fact that, in $\mathbf{K3LP^T}$, λ does not obey its own version of deletion, and proposes to explain it with the idea that λ is a "*half-truth*" (true *only in some sense*). As Égré himself observes, the introduction of a *dedicated semantic status for paradoxical sentences* generates new semantic paradoxes. For example, the proposed explanation appeals to the new idea that λ is a *half-truth* rather than to the old idea that λ is an *untruth* presumably because, in $\mathbf{K3LP^T}$, the latter doesn't really rule out that λ has all possible sorts of good semantic statuses. If so, it is crucial e.g. that *being half-true rules out being wholly true*. But then, consider a sentence λ' identical with 'λ' is not wholly true': since being true but not wholly true is tantamount to being half-true, it can be ruled out that λ' is wholly true, and so presumably λ' is in every sense not wholly true, and hence, since λ' says precisely that it is not wholly true, λ' is after all wholly true. Problem. Moreover, *a system is sound to the extent to which what it proves is true*: therefore, if $\mathbf{K3LP^T}$ proves half-truths, it is at best half-sound, and we'd better look for sounder systems (similarly, if a system of geography proves the half-truth that France is a South-American country (hello French Guiana!), we'd better look for a sounder system). Finally, it's important to note that λ is far from being the only case of a logical truth that is absurd, since, as I've already observed in fn 25, by courtesy of λ itself and its relatives *every unrestricted logical and truth-theoretic law has instances that are absurd in* $\mathbf{K3LP^T}$ *and, presumably, should thus also count—alongside their negations—as half-truths*. In effect, the proposed explanation is then subject to variations of all the problems mentioned in the fourth paragraph of this section (except for the dubiousness problem, since there is ample linguistic evidence that half-truths exist); in particular, if a defender of $\mathbf{K3LP^T}$ accepts that the target logical and truth-theoretic laws are only half-true—not a very alluring semantic status, as witnessed e.g. by the claim that France is a South-American country—it becomes unclear how fit a contender her view is in the debate on the semantic paradoxes. Thanks to an anonymous referee for bringing to my attention the relevance of Égré [12] for the issue of failure of deletion in $\mathbf{K3LP^T}$ and to Paul Égré for feedback on an earlier version of this discussion.

[42]A dual argument for the same conclusion is available by focusing on *absurdity* rather than on *logical truth* and appealing to the compelling metaentailment of f-*deletion*/f-*insertion* that $\Gamma \vdash \Delta, \mathsf{f}$ is valid iff $\Gamma \vdash \Delta$ is: left to right (and crucially for the present argument), *entailment save for the bans of logic is entailment* and, right to left (but not crucially for the present argument), *entailment is entailment save for the bans of logic*.

4 The Theory of Validity

I hope that the foregoing considerations will help in advancing our understanding of the *philosophical* features of $\mathbf{K3LP^T}$ (which, in my personal view, has somewhat lagged behind our understanding of its *formal* features). Let's now however move on from truth to what is, for many of us, another important semantic notion (e.g. *qua* having something to do with *truth preservation* broadly understood): *validity*.[43] The absolutely central feature of *naive validity* is the principle of *manifestation* of the *metatheoretic facts of entailment* by the *object-theoretic facts of validity*. With V a validity predicate, focusing on a single system and on single-premise, single-conclusion validities, manifestation can be formulated as being to the effect that $\Gamma, \varphi \vdash \Delta, \psi$ is valid iff $\Gamma \vdash \Delta, V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is, where all the elements of Γ and Δ are "validity sentences" (for the purposes of this paper, atomic V-sentences as well as t and f[44] are *validity sentences*, and so are negations, conjunctions, disjunctions and implications of validity sentences, see Zardini [46], p. 365).

It is possible to extend **K3LP** to a system (henceforth '$\mathbf{K3LP^V}$') where *manifestation holds* (and it is possible to do this in such a way that $\mathbf{K3LP^V}$ also contains a *classical theory of syntax*, which we'll henceforth assume). Manifestation gives rise to a naive-validity version of *Curry's paradox* (see Beall and Murzi [4] for a presentation and Zardini [44, 46] for some critical discussion). Letting κ be identical with $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$, such a version of Curry's paradox still shows that both $\varnothing \vdash_{\mathbf{K3LP^V}} \kappa$ and $\kappa \vdash_{\mathbf{K3LP^V}} \varnothing$ are valid[45] (just as the original, naive-truth version of Curry's paradox (where κ' is identical with $T(\ulcorner\kappa'\urcorner) \rightarrow \mathsf{f}$) did with respect to κ')[46]. However, while (TRANS), though not included as a defining principle of **K3LP**, was

[43]For example, one can understand the claim that the entailment from φ to ψ is *valid* as a *conditional truth attribution of a special type*: that ψ *is guaranteed to be true given* φ. Thanks to José Martínez for eliciting this point.

[44]In Section 3 (especially in fn 41), I've emphasised the incompatibility in $\mathbf{K3LP^T}$ of the thesis that the logical-truth constant is the conjunction of all logical truths and the thesis that the logical-truth constant obeys t-deletion (similarly for the absurdity constant). In this section, I'm going to understand t and f so that *they do obey* t-*deletion/*t-*insertion and* f-*deletion/*f-*insertion respectively*. *Model-theoretically*, you can think of this as t and f *always getting values 1 and 0* respectively; *informally*, you can think of this as t and f being the *trivial truth* ('For some P, it is the case that P') and the *trivial falsity* ('For every P, it is the case that P') respectively. Thanks to an anonymous referee for pointing out the need for this clarification.

[45]*Reason.* By reflexivity, $\kappa \vdash_{\mathbf{K3LP^V}} \kappa$ (i.e. $\kappa \vdash_{\mathbf{K3LP^V}} V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$) is valid, and so, by manifestation (right to left), $\kappa, \kappa \vdash_{\mathbf{K3LP^V}} \mathsf{f}$ is valid, and hence, by contraction, $\kappa \vdash_{\mathbf{K3LP^V}} \mathsf{f}$ (as well as, by f-deletion, $\kappa \vdash_{\mathbf{K3LP^V}} \varnothing$) is valid, and thus, by manifestation (left to right), $\varnothing \vdash_{\mathbf{K3LP^V}} V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ (i.e. $\varnothing \vdash_{\mathbf{K3LP^V}} \kappa$) is valid.

[46]*Reason.* By the entailment of *modus ponens* ($\varphi, \varphi \rightarrow \psi \vdash \psi$), $T(\ulcorner\kappa'\urcorner), T(\ulcorner\kappa'\urcorner) \rightarrow \mathsf{f} \vdash_{\mathbf{K3LP^T}} \mathsf{f}$ is valid, and so, by full intersubstitutability, $T(\ulcorner\kappa'\urcorner), T(\ulcorner\kappa'\urcorner) \vdash_{\mathbf{K3LP^T}} \mathsf{f}$ is valid, and hence, by contraction, $T(\ulcorner\kappa'\urcorner) \vdash_{\mathbf{K3LP^T}} \mathsf{f}$ (as well as, by f-deletion and full intersubstitutability, $\kappa' \vdash_{\mathbf{K3LP^T}} \varnothing$) is valid, and thus, by the metaentailment of the *deduction theorem* (if $\Gamma, \varphi \vdash \Delta, \psi$ is valid, $\Gamma \vdash \Delta, \varphi \rightarrow \psi$ is valid), $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\kappa'\urcorner) \rightarrow \mathsf{f}$ (i.e. $\varnothing \vdash_{\mathbf{K3LP^T}} \kappa'$) is valid.

nevertheless admissible in it, it is no longer admissible in $\mathbf{K3LP^V}$, so that $\varnothing \vdash_{\mathbf{K3LP^V}} \varnothing$ is not valid. That is essentially supposed to be $\mathbf{K3LP^V}$'s solution to the naive-validity version of Curry's paradox. For reasons analogous to those seen in the case of λ, this is not a particularly effective solution to the paradox, but I'm not going to insist on those.[47,48]

I'm going to develop instead an interesting twist to which, contrary to the original, naive-truth version of Curry's paradox and also contrary to the Liar paradox, the naive-validity version of Curry's paradox readily (but not exclusively, see fn 58) offers itself: such development will constitute the main argument of this section and will culminate in the advertised "final cut". Notice that, in $\mathbf{K3LP^V}$, both the notion of

[47] Shapiro [35] touches on something like the fact that, in $\mathbf{K3LP^V}$, *something is supposed both to hold but also to entail triviality* (fn 55 places and then eliminates an obstacle to such a fact's posing a problem for $\mathbf{K3LP^V}$). However, contrary to what Shapiro would seem to assume, such a kind of commitment is not a novel feature of $\mathbf{K3LP^V}$ *vis-à-vis* $\mathbf{K3LP^T}$. As we've seen in Section 3 (especially in its fourth paragraph), *the same kind of commitment must already be undertaken in the case of* $\mathbf{K3LP^T}$*'s treatment of the Liar paradox*. Shapiro does not discuss the style of consideration put forward towards the end of that paragraph to pin down that kind of commitment and is instead struck by the point that, if one accepts that $\kappa \vdash \mathsf{f}$ is valid, *since* $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ *is supposed to say just that one must accept* $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$. But, even on the latter line of thought, *the same kind of point already applies in the case of* $\mathbf{K3LP^T}$*'s treatment of the Liar paradox*, given that absurdities are those sentences that *cannot logically be true* (cf Zardini [47], p. 473): if one accepts that $\lambda \vdash \mathsf{f}$ is valid (i.e. that λ is absurd), *since 'λ cannot logically be true' is supposed to say just that one must accept 'λ cannot logically be true'*, and so, by law **T** for logical modality (in the version $\neg\Diamond\varphi \rightarrow \neg\varphi$), one must accept $\neg T(\ulcorner\lambda\urcorner)$. Thanks to an anonymous referee for feedback that led to a rewriting of the material in this fn.

[48] Barrio et al. [3] rely on the fact that $\kappa \vdash_{\mathbf{K3LP^V}} \varnothing$ is valid to show that $\varnothing \vdash_{\mathbf{K3LP^V}} (V(\ulcorner\mathsf{t}\urcorner, \ulcorner\kappa\urcorner)\&\kappa) \rightarrow V(\ulcorner\mathsf{t}\urcorner, \ulcorner\mathsf{f}\urcorner)$ (i.e. $\varnothing \vdash_{\mathbf{K3LP^V}} (V(\ulcorner\mathsf{t}\urcorner, \ulcorner\kappa\urcorner)\&V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)) \rightarrow V(\ulcorner\mathsf{t}\urcorner, \ulcorner\mathsf{f}\urcorner)$) is valid, which they find is in tension with the fact that, according to $\mathbf{K3LP^V}$, $\mathsf{t} \vdash \kappa$ and $\kappa \vdash \mathsf{f}$ are valid but $\mathsf{t} \vdash \mathsf{f}$ isn't. However, contrary to what Barrio, Rosenblatt and Tájer would seem to assume, such a kind of tension is not a novel feature of $\mathbf{K3LP^V}$ *vis-à-vis* $\mathbf{K3LP^T}$. As we've come close to seeing in Section 3 (especially in its fifth paragraph), *the same kind of tension is already present in the case of* $\mathbf{K3LP^T}$*'s treatment of the Liar paradox*. To wit, $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\lambda\urcorner) \rightarrow T(\ulcorner\mathsf{f}\urcorner)$ is valid, which you might find is in tension with the fact that, according to $\mathbf{K3LP^T}$, λ is true but f isn't (for more tension, add the fact that it is not the case that, according to $\mathbf{K3LP^T}$, f is true). The diagnosis of the general problem here would seem clear: the problem is due to the fact that, in these systems, *some sentences are both logical truths and absurdities*. In Section 3, we've already seen some kinds of problems this fact generates; Barrio, Rosenblatt and Tájer's observation (and the one in this fn) points to a further such kind. (Hlobil [21] addresses Barrio, Rosenblatt and Tájer's observation by in effect so characterising V that a sentence like κ is no longer absurd. *Such a general type of strategy* would not seem to go *far enough*, since it does not address the related problem observed in this fn. Moreover, *the particular way in which Hlobil implements the strategy* spawns additional concerns, for *it requires denying the right-to-left direction of manifestation*—indeed, it requires denying the entailment of "*soundness*" $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \vdash \varphi \rightarrow \psi$, given which it becomes unclear in what sense the resulting notion expressed by V is still a notion of *naive* validity: similarly to how a sentence's being really true (Section 3) entails that what the sentence says is the case, an entailment's being really valid presumably entails that what its premise says suffices for what its conclusion says (Zardini [44], p. 637, fn 7; a concern that is reinforced by the circumstance that the characterisation of V offered by Hlobil consists of principles that are *just as plausible* if $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is understood instead as '**K3LP** validates or is so extended as to validate $\varphi \vdash \psi$'—a *notion* that is very different from that of naive validity, even by the lights of those that would accept that they are *coextensional*!).) Thanks to Ulf Hlobil for inspiring this fn.

what is naively valid and *what is valid in* **K3LP$^{\mathbf{V}}$** are expressible.[49] Suppose now, for dramatic purposes, that, insofar as its target subject matter is concerned, you adopt **K3LP$^{\mathbf{V}}$**. What relations should you think obtain between those two notions (keeping in mind that, as *per* Section 3, naive validity is real validity)?

In one direction, you should presumably still think that *there is a gap from naive validity to validity in* **K3LP$^{\mathbf{V}}$**. For, after all, there are presumably entailments that are naively valid in virtue of certain notions but that are not valid even in **K3LP$^{\mathbf{V}}$** simply because **K3LP$^{\mathbf{V}}$** is not supposed to be a logic for those notions (for example, the entailment from 'It is necessary that φ' to φ, which is valid in virtue of the notion of necessity). In the other direction, however, you should presumably now think that *there is no gap from validity in* **K3LP$^{\mathbf{V}}$** *to naive validity*, in the sense that an entailment's being valid in **K3LP$^{\mathbf{V}}$** is *inseparable* (in all contexts that matter, cf fn 37) from its being naively valid (with the consequence e.g. that you should only think that an entailment is valid in **K3LP$^{\mathbf{V}}$** if you also think that it is naively valid). For, after all, we're supposing that you think that, insofar as its target subject matter is concerned, **K3LP$^{\mathbf{V}}$** gets things *just right*, and so that it gets them *just right* in particular whenever it declares an entailment valid (whereas, if e.g. you think that an entailment is valid in **K3LP$^{\mathbf{V}}$** without thinking that it is naively valid, you're not letting your judgements of naive validity *conform* to your judgements of validity in **K3LP$^{\mathbf{V}}$**, and so, even insofar as its target subject matter is concerned, you're not after all *adopting* **K3LP$^{\mathbf{V}}$**).

In a normal situation, the absence of such a gap would be guaranteed by an entailment (or implication, or something like that) from an entailment's being valid in **K3LP$^{\mathbf{V}}$** to its being naively valid. But, relatedly to what I've already noted in another dialectical context in Section 3, the situation is not normal: in **K3LP$^{\mathbf{V}}$**, entailment has nothing like that force (for example, in **K3LP$^{\mathbf{V}}$**, the entailment from λ to f is valid, although the former is accepted and the latter rejected). The natural diagnosis of this problem is that it arises from the general failure of (TRANS) in **K3LP$^{\mathbf{V}}$**: letting $V_{\mathbf{K3LP^V}}$ express validity in **K3LP$^{\mathbf{V}}$**, if (TRANS) fails, then, if e.g. $\Gamma \vdash V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is valid, even assuming that $V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \vdash V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is valid that does not entail that $\Gamma \vdash V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is valid, so that, under the supposition of Γ, it is the case that [$V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ holds but $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ might not]. In such a context, $V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is then separable from $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ (with the consequence e.g. that, under the supposition of Γ, you may think that

[49] To see the latter, consider e.g. that **K3LP$^{\mathbf{V}}$** contains a classical theory of syntax and that **K3LP$^{\mathbf{V}}$** itself can be defined proof-theoretically (e.g. by taking a proof-theoretic definition of **K3LP** and extending it with rules suitable for a classical theory of syntax as well as with the two-way rule corresponding to manifestation) and so in a way that can straightforwardly be represented by a classical theory of syntax. (Indeed, even if **K3LP$^{\mathbf{V}}$** were such that, for whatever reason, the notion of what is valid in **K3LP$^{\mathbf{V}}$** is not capturable by a classical theory of syntax, there would still be a *sound (but incomplete) proof-theoretic approximation of that notion that does contain manifestation and is capturable by a classical theory of syntax*, and the main argument of this section could then be run by substituting derivability in that approximation for validity in **K3LP$^{\mathbf{V}}$**.)

$V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ holds without thinking that $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ holds). What is then required to close the ensuing gap is that (TRANS) not fail in the case where its second entailment (the one from the intermediate conclusion to the final conclusion) is $V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \vdash V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$, which is in effect tantamount to the requirement that $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ can be substituted for $V_{\mathbf{K3LP^V}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ as a conclusion.

Generalising this natural treatment, the absence of a gap from validity in $\mathbf{K3LP^V}$ to naive validity can still be guaranteed in a nontransitive framework by requiring that V can be substituted for $V_{\mathbf{K3LP^V}}$ in contexts that are traditionally taken to allow the substitution of weaker sentences for stronger ones (for example, a stand-alone sentence as a conclusion is in such a context) and, *vice versa*, $V_{\mathbf{K3LP^V}}$ can be substituted for V in contexts that are traditionally taken to allow the substitution of stronger sentences for weaker ones (for example, a sentence embedded under negation as a conclusion is in such a context). We're therefore led to requiring a principle of *one-way substitutability* of V for $V_{\mathbf{K3LP^V}}$ (henceforth simply 'one-way substitutability'): V can be substituted for $V_{\mathbf{K3LP^V}}$ in "upwards-monotonic contexts" and, *vice versa*, $V_{\mathbf{K3LP^V}}$ can be substituted for V in "downwards-monotonic" contexts (where, for the purposes of this paper, stand-alone sentences as conclusions (premises) are in an *upwards-monotonic context* (*downwards-monotonic context*); sentences in an upwards-monotonic context (downwards-monotonic context) further embedded under negation or as antecedents of an implication are in a downwards-monotonic context (upwards-monotonic context); sentences in an upwards-monotonic context (downwards-monotonic context) further embedded as consequents of an implication, or as conjuncts of a conjunction, or as disjuncts of a disjunction are in an upwards-monotonic context (downwards-monotonic context)).

Now, since we're assuming that $\mathbf{K3LP^V}$ contains a classical theory of syntax, one-way substitutability cannot be part of $\mathbf{K3LP^V}$ itself. The reason for this is *Gödelian*. By reflexivity and manifestation (right to left), $\mathsf{t}, V(\ulcorner\mathsf{t}\urcorner, \ulcorner\mathsf{f}\urcorner) \vdash_{\mathbf{K3LP^V}} \mathsf{f}$ is valid, and so, by f-deletion and the properties of negation in $\mathbf{K3LP^V}$, $\mathsf{t} \vdash_{\mathbf{K3LP^V}} \neg V(\ulcorner\mathsf{t}\urcorner, \ulcorner\mathsf{f}\urcorner)$ is valid. If one-way substitutability were part of $\mathbf{K3LP^V}$, it would follow that $\mathsf{t} \vdash_{\mathbf{K3LP^V}} \neg V_{\mathbf{K3LP^V}}(\ulcorner\mathsf{t}\urcorner, \ulcorner\mathsf{f}\urcorner)$ is valid, thereby crashing on *Gödel's second incompleteness theorem* [18]. Therefore, we must assume that one-way substitutability is only part of a *stronger* system $\mathbf{K3LP^{VS^{\Rightarrow}}}$.

Unfortunately, $\mathbf{K3LP^{VS^{\Rightarrow}}}$ is an untenable system. Since $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner) \vdash_{\mathbf{K3LP^V}} \varnothing$ is valid and $\mathbf{K3LP^{VS^{\Rightarrow}}}$ is stronger than $\mathbf{K3LP^V}$, $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner) \vdash_{\mathbf{K3LP^{VS^{\Rightarrow}}}} \varnothing$ is valid, and so, by one-way substitutability, $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner) \vdash_{\mathbf{K3LP^{VS^{\Rightarrow}}}} \varnothing$ is valid. Moreover, since $\kappa \vdash_{\mathbf{K3LP^V}} \mathsf{f}$ is valid and $\mathbf{K3LP^V}$ contains a classical theory of syntax, $\varnothing \vdash_{\mathbf{K3LP^V}} V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ is valid, and so, since $\mathbf{K3LP^{VS^{\Rightarrow}}}$ is stronger than $\mathbf{K3LP^V}$, $\varnothing \vdash_{\mathbf{K3LP^{VS^{\Rightarrow}}}} V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ is valid. Now, presumably, even in a $\mathbf{K3LP^{VS^{\Rightarrow}}}$-style system, (TRANS) should hold *when the cut formula is an elementary, provable mathematical claim*, such as '$2 + 2 = 4$' or... $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ (e.g. recall from fn 49 that $\mathbf{K3LP^V}$ can be defined proof-theoretically, so that $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ is then an elementary, provable proof-theoretic claim). But then, as the song goes, we should have the nerve to make the final cut on $\varnothing \vdash_{\mathbf{K3LP^{VS^{\Rightarrow}}}}$

$V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner)$ and $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner) \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} \varnothing$, and get that $\varnothing \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} \varnothing$ is valid. Therefore, $\mathbf{K3LP^{VS^\Rightarrow}}$ is totally trivial.[50]

At the same time, while $\mathbf{K3LP^{VS^\Rightarrow}}$ is stronger than $\mathbf{K3LP^V}$, it would seem *the system that a defender of* $\mathbf{K3LP^V}$ *should adopt, so that she can make sure that, in the object theory, there is no gap from validity in* $\mathbf{K3LP^V}$ *to naive validity*: coherently, one can only adopt the weaker system $\mathbf{K3LP^V}$ if one adopts the stronger system $\mathbf{K3LP^{VS^\Rightarrow}}$.[51] (That is similar to how, coherently, one can only adopt e.g. the weaker system **PA** if one adopts the stronger system obtained from **PA** by adding the reflection principle for provability in **PA**.)[52] Since one can't coherently adopt $\mathbf{K3LP^{VS^\Rightarrow}}$, one can't coherently adopt $\mathbf{K3LP^V}$ either.

Notice that, in a sense, the argument *generalises* to all systems where $\kappa \vdash \mathsf{f}$ is valid (and which contain a classical theory of syntax).[53] In turn, these include all

[50]Despite the title of this paper, "the final cut" is not really needed to bring out the core of the problem here: we might just as well only focus on the "right half" of the cut to the effect that $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner) \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} \varnothing$ is valid, which *already suffices to make* $\mathbf{K3LP^{VS^\Rightarrow}}$ *untenable*, since, to repeat, $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner)$ *is an elementary, provable mathematical claim*.

[51]From the fifth to the third last paragraphs, I've argued for this claim by arguing that the acceptance of the absence of a gap from validity in $\mathbf{K3LP^V}$ to naive validity is required by the adoption of $\mathbf{K3LP^V}$ (fifth last paragraph) and that one-way substitutability guarantees the absence of such a gap (fourth and third last paragraphs). A defender of $\mathbf{K3LP^V}$ may thus take my argument as issuing the challenge of either resisting the move from the adoption of $\mathbf{K3LP^V}$ to what can plausibly be taken to be the acceptance of the absence of a gap from validity in $\mathbf{K3LP^V}$ to naive validity or finding a principle that is weaker than one-way substitutability but that can still plausibly be taken to guarantee the absence of such a gap. Thanks to an anonymous referee for help with clarifying this dialectical situation.

[52]One can also go back and briefly compare a reflection principle with manifestation. Let **S** be a system that contains a classical theory of syntax and can express its own consequence relation, $V_{\mathbf{S}}$ express validity in **S**, $\mathbf{S^R}$ be the extension of **S** with the reflection principle $\varnothing \vdash V_{\mathbf{S}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \rightarrow (\varphi \rightarrow \psi)$ and $\mathbf{S^V}$ be the extension of **S** with manifestation. Then, just as $\varnothing \vdash_{\mathbf{S^R}} V_{\mathbf{S}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \rightarrow (\varphi \rightarrow \psi)$ is valid, so is $\varnothing \vdash_{\mathbf{S^V}} V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \rightarrow (\varphi \rightarrow \psi)$ (by reflexivity, manifestation (right to left) and the deduction theorem). This helps to bring out the *similarity* as well as the *dissimilarity* between a reflection principle and manifestation: while they both entail that a certain target consequence relation (validity in **S** and naive validity respectively) *is sound*, the target consequence relation of a reflection principle *does not itself contain* the reflection principle (validity in **S** does not itself contain the entailment $\varnothing \vdash V_{\mathbf{S}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) \rightarrow (\varphi \rightarrow \psi)$, and so $\mathbf{S^R}$ cannot be at least as weak as validity in **S**), whereas the target consequence relation of manifestation *does itself contain* manifestation (naive validity does itself contain the metaentailment that $\Gamma, \varphi \vdash \Delta, \psi$ is valid iff $\Gamma \vdash \Delta, V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ is, where all the elements of Γ and Δ are validity sentences, and so $\mathbf{S^V}$ can be at least as weak as naive validity). Thanks to an anonymous referee for recommending this comparison.

[53]*Reason.* Let **S** (for 'smart' or 'silly' you choose) be any such system and $\mathbf{S^{S^\Rightarrow}}$ its extension validating the corresponding version of one-way substitutability. Since $V(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner) \vdash_{\mathbf{S}} \varnothing$ is valid and $\mathbf{S^{S^\Rightarrow}}$ is stronger than **S**, $V(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner) \vdash_{\mathbf{S^{S\Rightarrow}}} \varnothing$ is valid, and so, by one-way substitutability, $V_{\mathbf{S}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner) \vdash_{\mathbf{S^{S\Rightarrow}}} \varnothing$ is valid. Moreover, since $\kappa \vdash_{\mathbf{S}} \mathsf{f}$ is valid and **S** contains a classical theory of syntax, $\varnothing \vdash_{\mathbf{S}} V_{\mathbf{S}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner)$ is valid, and so, since $\mathbf{S^{S^\Rightarrow}}$ is stronger than **S**, $\varnothing \vdash_{\mathbf{S^{S\Rightarrow}}} V_{\mathbf{S}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner)$ is valid. Now, presumably, even in an $\mathbf{S^{S^\Rightarrow}}$-style system, (TRANS) should hold *when the cut formula is an elementary, provable mathematical claim*. But then we can cut on $\varnothing \vdash_{\mathbf{S^{S\Rightarrow}}} V_{\mathbf{S}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner)$ and $V_{\mathbf{S}}(\ulcorner\kappa\urcorner, \ulcorner \mathsf{f}\urcorner) \vdash_{\mathbf{S^{S\Rightarrow}}} \varnothing$, and get that $\varnothing \vdash_{\mathbf{S^{S\Rightarrow}}} \varnothing$ is valid. Therefore, $\mathbf{S^{S^\Rightarrow}}$ is totally trivial. At the same time, while $\mathbf{S^{S^\Rightarrow}}$ is stronger than **S**, it would seem *the system that a defender of* **S** *should adopt, so that she can make sure that, in the object theory, there is no gap from validity in* **S** *to naive validity*: coherently, one can only adopt the weaker system **S** if one adopts the stronger system $\mathbf{S^{S^\Rightarrow}}$. Since one can't coherently adopt $\mathbf{S^{S^\Rightarrow}}$, one can't coherently adopt **S** either.

the systems where reflexivity is valid, manifestation (right to left) holds and the metaentailment of *contraction*:

(CONTR) If $\Gamma, \varphi, \varphi \vdash \Delta$ is valid, $\Gamma, \varphi \vdash \Delta$ is valid, and, if $\Gamma \vdash \Delta, \varphi, \varphi$ is valid, $\Gamma \vdash \Delta, \varphi$ is valid

holds (first conjunct).[54]

However, the *transitive* systems where $\kappa \vdash \mathsf{f}$ is valid are arguably shown to be untenable *already by the old naive-validity version of Curry's paradox*. For, if one accepts that $\kappa \vdash \mathsf{f}$ is valid, one accepts that the entailment from κ to f is really valid (fn 4), and so one is in effect willy-nilly accepting κ, which says just that (Section 3, see also fn 47). Therefore, one accepts a sentence that one accepts entails f, and so, by transitivity, one is committed to accepting f.[55] For those systems, the new naive-validity version of Curry's paradox relying on one-way substitutability is thus an overkill. Indeed, it is very much arguable that, given the kind of conception of validity that—for better or worse—*contractive transitive* systems are already committed to in view of the original, naive-truth version of Curry's paradox, the component of manifestation that they should reject is precisely the right-to-left direction (Zardini [44], pp. 636–638; cf Field [17]). For such a version of Curry's paradox shows that, in such systems, keeping fixed *modus ponens* for $\rightarrow$, it follows that $\kappa' \vdash \mathsf{f}$ is valid but $\varnothing \vdash \kappa' \rightarrow \mathsf{f}$ isn't: such a failure of the deduction theorem entails that, in such systems, the validity of the entailment from φ to ψ cannot be understood as requiring that φ be a *sufficient condition* for ψ, thereby putting into question $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner), \varphi \vdash \psi$ and so the right-to-left direction of manifestation.

At this juncture, it is worth noting that *noncontractive* systems can afford a much tighter connection between entailment and object-theoretic notions like implication and validity, and so can validate manifestation. On a noncontractive system validating manifestation (henceforth generically '$\mathbf{LW^V}$'; the details of the system won't matter for the purposes of this paper but see e.g. Zardini [46] for a specific proposal), $\kappa, \kappa \vdash_{\mathbf{LW^V}} \mathsf{f}$ is valid but $\kappa \vdash_{\mathbf{LW^V}} \mathsf{f}$ isn't. That is sufficient for solving the old naive-validity version of Curry's paradox, but could we use something like the ideas of the fourth last paragraph concerning a "final cut" to argue also for a "final contraction"? Well, by considerations analogous to those made for $\mathbf{K3LP^V}$, we can extend $\mathbf{LW^V}$ to a stronger system $\mathbf{LW^{VS^{\Rightarrow}}}$ where

[54] *Reason.* By reflexivity, $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner) \vdash V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner)$ is valid, and so, by manifestation (right to left), $V(\ulcorner\kappa\urcorner, \ulcorner\mathsf{f}\urcorner), \kappa \vdash \mathsf{f}$ (i.e. $\kappa, \kappa \vdash \mathsf{f}$) is valid, and hence, by (CONTR) (first conjunct), $\kappa \vdash \mathsf{f}$ is valid.

[55] The argument is not so direct for *nontransitive* systems, since, in such systems, it is not generally the case that, if one accepts φ and accepts that $\varphi \vdash \psi$ is valid, one is committed to accepting ψ (see Zardini [50], pp. 255–265 for an extended discussion of the normativity of nontransitive logical consequence, cf fn 27). Still, even setting aside the problems that have emerged in Section 3 for accepting both those claims in the special case where ψ is f, the general principle just mentioned is supposed to fail in a nontransitive system only if one accepts φ because it is the conclusion of an entailment of type X (fn 27), a circumstance that is presumably not realised in the case of one's acceptance that $\kappa \vdash \mathsf{f}$ is valid (as opposed to one's acceptance of, say, λ because it is the conclusion of $\varnothing \vdash \lambda$), and so, for the reason pointed out in the main text, presumably not realised in the case of one's acceptance of κ. As far as I'm concerned, that's actually decisive; however, be that as it may with that dialectic, the main argument of this section is supposed to show that, once one reflects on certain features connected with the fact that $\kappa \vdash \mathsf{f}$ is valid in a nontransitive system like $\mathbf{K3LP^V}$, ultimately not even such a system is off the hook.

one-way substitutability holds. Since $V(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner), V(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner) \vdash_{\mathbf{LW}^{\mathbf{V}}} \varnothing$[56] is valid and $\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}$ is stronger than $\mathbf{LW}^{\mathbf{V}}$, $V(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner), V(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner) \vdash_{\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}} \varnothing$ is valid, and so, by one-way substitutability, $V_{\mathbf{LW}^{\mathbf{V}}}(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner), V_{\mathbf{LW}^{\mathbf{V}}}(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner) \vdash_{\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}} \varnothing$ is valid. Now, presumably, even in an $\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}$-style system, (CONTR) should hold *when the to-be-contracted formula is an elementary, disprovable mathematical claim*, such as '2 + 2 = 5' or… $V_{\mathbf{LW}^{\mathbf{V}}}(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner)$. But then we can contract on $V_{\mathbf{LW}^{\mathbf{V}}}(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner), V_{\mathbf{LW}^{\mathbf{V}}}(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner) \vdash_{\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}} \varnothing$, and get that $V_{\mathbf{LW}^{\mathbf{V}}}(\ulcorner\kappa\urcorner, \ulcorner\mathrm{f}\urcorner) \vdash_{\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}} \varnothing$ is valid. But, contrary to the case of $\mathbf{K3LP}^{\mathbf{VS}^{\Rightarrow}}$, this conclusion is *wholly benign* (indeed, it's just what you'd expect from the kind of stronger system $\mathbf{LW}^{\mathbf{VS}^{\Rightarrow}}$ is). Since there is a similar reason for expecting that also nonreflexive systems are not affected by the new naive-validity version of Curry's paradox, concluding by putting together what we've seen in these last three paragraphs *the only system that is straightforwardly affected by the new naive-validity version of Curry's paradox but not by the old one is a nontransitive system like* $\mathbf{K3LP}^{\mathbf{V}}$. Therefore, the new naive-validity version of Curry's paradox targets *specifically* a nontransitive system like $\mathbf{K3LP}^{\mathbf{V}}$.

There is more to naive validity than manifestation. In particular, a system should not only connect the *metatheoretic facts of its entailments* with the *object-theoretic facts of naive validity*—the driving force behind naivety (Section 3) requires that the system should also be strengthened to connect the latter facts with the *object-theoretic facts of what is valid in the system*. Naive validity should also include one-way substitutability. However, although $\mathbf{K3LP}^{\mathbf{V}}$ solves the old naive-validity version of Curry's paradox (sort of, see the second paragraph of this section) and validates manifestation, $\mathbf{K3LP}^{\mathbf{VS}^{\Rightarrow}}$ crashes on the new naive-validity version of Curry's paradox, and so $\mathbf{K3LP}^{\mathbf{V}}$ cannot be strengthened to validate one-way substitutability.[57] Matters of naive validity cannot conform to matters of validity in $\mathbf{K3LP}^{\mathbf{V}}$.[58]

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

---

[56]When discussing a noncontractive system, it becomes essential that premises and conclusions be combined in *collections that are finer-grained than sets*; correspondingly, in such a discussion, '∅' refers to the *empty collection at the relevant level of fineness of grain*.

[57]Perhaps, one useful way of looking at the overall situation is that, while it's relatively easy to make the first connection mentioned above in the main text by dint of free-floating, "garbage" sentences (such as $V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$) that allow for a promiscuous behaviour both as logical truths and as absurdities, the problem is that the second connection mentioned above in the main text requires linking those sentences with highly constrained, "serious" sentences (such as $V_{\mathbf{K3LP}^{\mathbf{V}}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$) that don't admit of any such behaviour. Thanks to Filippo Ferrari for inviting such a reflection.

[58]That said, it's interesting to observe that the essence of the main argument of this section doesn't really require the notion of naive validity (which, actually, is in many ways problematic, see e.g. Zardini [44, 46]). We can start with the original, naive-truth version of Curry's paradox and, supposing that, insofar as its target subject matter is concerned, you adopt $\mathbf{K3LP}^{\mathbf{T}}$, observe instead that it is just as compelling to claim that you should now think that *there is no gap from validity in* $\mathbf{K3LP}^{\mathbf{T}}$ *to truth preservation*, and so that you should accept one-way substitutability of $T(\ulcorner\varphi\urcorner) \rightarrow T(\ulcorner\psi\urcorner)$ for $V_{\mathbf{K3LP}^{\mathbf{T}}}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$. Appropriately modified, the main argument of this section will then go through. Thanks to José Martínez for eliciting this point.

References

1. Asenjo, F. (1966). A calculus of antinomies. *Notre Dame Journal of Formal Logic*, *7*, 103–105.
2. Barrio, E., Pailos, F., & Szmuc, D. (2020). A hierarchy of classical and paraconsistent logics. *Journal of Philosophical Logic*, *49*, 93–120.
3. Barrio, E., Rosenblatt, L., & Tájer, D. (2021). Capturing naive validity in the cut-free approach. *Synthese*, *199*, 707–723.
4. Beall, J. C., & Murzi, J. (2013). Two flavors of Curry's paradox. *The Journal of Philosophy*, *110*, 143–165.
5. Carnap, R. (1931). Überwindung der Metaphysik durch logische Analyse der Sprache. *Erkenntnis*, *2*, 219–241.
6. Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2012). Tolerant, classical, strict. *Journal of Philosophical Logic*, *41*, 347–385.
7. Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind*, *122*, 841–866.
8. Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2015). Pragmatic interpretations of vague expressions: Strongest meaning and nonmonotonic consequence. *Journal of Philosophical Logic*, *44*, 375–393.
9. Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2020). Inferences and metainferences in ST. *Journal of Philosophical Logic*, *49*, 1057–1077.
10. Da Ré, B. (2021). Structural weakening and the paradoxes. *Notre Dame Journal of Formal Logic*, *62*, 369–398.
11. Dragalin, A. (1979). *Matematičeskij intuicionizm. Vvedenie v teoriju dokazatel'stv*. Moscow: Nauka.
12. Égré, P. (2021). Half-truths and the liar. In C. Nicolai, & J. Stern (Eds.) *Modes of truth* (pp. 18–40). London: Routledge.
13. Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.
14. Field, H. (1973). Theory change and the indeterminacy of reference. *The Journal of Philosophy*, *70*, 462–481.
15. Field, H. (1994). Deflationist views of meaning and content. *Mind*, *103*, 249–285.
16. Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.
17. Field, H. (2017). Disarming a paradox of validity. *Notre Dame Journal of Formal Logic*, *58*, 1–19.
18. Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, *38*, 173–198.
19. Gupta, A., & Standefer, S. (2018). Intersubstitutivity principles and the generalization function of truth. *Synthese*, *195*, 1065–1075.
20. Heck, R. (2012). A liar paradox. *Thought*, *1*, 36–40.
21. Hlobil, U. (2019). Faithfulness for naive validity. *Synthese*, *196*, 4759–4774.
22. Horwich, P. (1998). *Truth* (2nd edn). Oxford: Clarendon Press.
23. Incurvati, L., & Schlöder, J. (2019). Weak assertion. *The Philosophical Quarterly*, *69*, 741–770.
24. Kleene, S. (1938). On a notation for ordinal numbers. *The Journal of Symbolic Logic*, *3*, 150–155.
25. López de Sa, D., & Zardini, E. (2007). Truthmakers, knowledge and paradox. *Analysis*, *67*, 242–250.
26. Makinson, D. (1965). The paradox of the preface. *Analysis*, *25*, 205–207.
27. Oms, S., & Zardini, E. (2019). An introduction to the sorites paradox. In S. Oms, & E. Zardini (Eds.) *The sorites paradox* (pp. 3–18). Cambridge: Cambridge University Press.
28. Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, *8*, 219–241.
29. Priest, G. (2006a). *In contradiction* (2nd edn). Oxford: Oxford University Press.
30. Priest, G. (2006b). *Doubt truth to be a liar*. Oxford: Oxford University Press.
31. Read, S. (1988). *Relevant logic*. Oxford: Blackwell.
32. Restall, G. (1996). Truthmakers, entailment and necessity. *Australasian Journal of Philosophy*, *74*, 331–340.
33. Ripley, D. (2012). Conservatively extending classical logic with transparent truth. *The Review of Symbolic Logic*, *5*, 354–378.

34. Ripley, D. (2013). Paradoxes and failures of cut. *Australasian Journal of Philosophy*, *91*, 139–164.
35. Shapiro, L. (2013). Validity Curry strengthened. *Thought*, *2*, 100–107.
36. Tarski, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, *4*, 341–376.
37. Weir, A. (1993). Putnam, Gödel and mathematical realism. *International Journal of Philosophical Studies*, *1*, 255–285.
38. Weir, A. (2015). A robust non-transitive logic. *Topoi*, *34*, 99–107.
39. Yablo, S. (1985). Truth and reflection. *Journal of Philosophical Logic*, *14*, 297–348.
40. Zardini, E. (2008a). A model of tolerance. *Studia logica*, *90*, 337–368.
41. Zardini, E. (2008b). *Living on the slippery slope. The nature, sources and logic of vagueness*. PhD thesis, University of St Andrews.
42. Zardini, E. (2011). Truth without contra(di)ction. *The Review of Symbolic Logic*, *4*, 498–535.
43. Zardini, E. (2013a). It is not the case that [P and 'It is not the case that P' is true] nor is it the case that [P and 'P' is not true]. *Thought*, *1*, 309–319.
44. Zardini, E. (2013b). Naive logical properties and structural properties. *The Journal of Philosophy*, *110*, 633–644.
45. Zardini, E. (2014a). Evans tolerated. In K. Akiba, & A. Abasnezhad (Eds.) *Vague objects and vague identity* (pp. 327–352). Dordrecht: Springer.
46. Zardini, E. (2014b). Naive truth and naive logical properties. *The Review of Symbolic Logic*, *7*, 351–384.
47. Zardini, E. (2015a). Getting one for two, or the contractors' bad deal. Towards a unified solution to the semantic paradoxes. In T. Achourioti, K. Fujimoto, H. Galinon, & J. Martínez (Eds.) *Unifying the philosophy of truth* (pp. 461–493). Dordrecht: Springer.
48. Zardini, E. (2015b). The opacity of truth. *Topoi*, *34*, 37–54.
49. Zardini, E. (2015c). $\forall$ and ω. In A. Torza (Ed.) *Quantifiers, quantifiers, and quantifiers: Themes in logic, metaphysics, and language* (pp. 489–526). Dordrecht: Springer.
50. Zardini, E. (2015d). Breaking the chains. Following-from and transitivity. In C. Caret, & O. Hjortland (Eds.) *Foundations of logical consequence* (pp. 221–275). Oxford: Oxford University Press.
51. Zardini, E. (2016). Restriction by noncontraction. *Notre Dame Journal of Formal Logic*, *57*, 287–327.
52. Zardini, E. (2018). Generalised Tarski's thesis hits substructure. In N. Kellen, N. Pedersen, & J. Wyatt (Eds.) *Pluralisms in truth and logic* (pp. 237–276). Basingstoke: Palgrave MacMillan.
53. Zardini, E. (2019a). Instability and contraction. *Méditations hégéliennes* I. *Journal of Philosophical Logic*, *48*, 155–188.
54. Zardini, E. (2019b). Non-transitivism and the sorites paradox. In S. Oms, & E. Zardini (Eds.) *The sorites paradox* (pp. 168–186). Cambridge: Cambridge University Press.
55. Zardini, E. (2021). Substructural approaches to paradox: An introduction to the special issue. *Synthese*, *199*, 493–525.
56. Zardini, E. (2022a). Neoclassical logic, cloisonnist vagueness and nonobjectual truth. ms.
57. Zardini, E. (2022b). Paradox and substructurality. ms.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Elia Zardini[1,2]

[1] Department of Logic and Theoretical Philosophy, Faculty of Philosophy, Complutense University of Madrid, Madrid, Spain

[2] International Laboratory for Logic, Linguistics and Formal Philosophy, School of Philosophy and Cultural Studies, National Research University Higher School of Economics, Moscow, Russian Federation