Synthese
https://doi.org/10.1007/s11229-018-1711-1

Paradoxical hypodoxes

Alexandre Billon[1,2]

Received: 9 March 2017 / Accepted: 31 January 2018

Abstract Most paradoxes of self-reference have a dual or 'hypodox'. The Liar paradox (Lr = 'Lr is false') has the Truth-Teller (Tt = 'Tt is true'). Russell's paradox, which involves the set of sets that are not self-membered, has a dual involving the set of sets which are self-membered, etc. It is widely believed that these duals are not paradoxical or at least not as paradoxical as the paradoxes of which they are duals. In this paper, I argue that some paradox's duals or hypodoxes are as paradoxical as the paradoxes of which they are duals, and that they raise neglected and interestingly different problems. I first focus on Richard's paradox (arguably the simplest case of a paradoxical dual), showing both that its dual is as paradoxical as Richard's paradox itself, and that the classical, Richard-Poincaré solution to the latter does not generalize to the former in any obvious way. I then argue that my argument applies mutatis mutandis to other paradoxes of self-reference as well, the dual of the Liar (the Truth-Teller) proving paradoxical.

Keywords Paradoxes of self-reference · Hypodoxes · Richard's paradox · Truth-teller · Circularity · Definitions

Most paradoxes of self-reference have a dual or 'hypodox'.[1] The Liar paradox (Lr = 'Lr is false') has the Truth-Teller (Tt = 'Tt is true'). Russell's paradox, which involves

[1] I use the phrase 'paradoxes of self-reference' to fix the reference to paradoxes that are of the same kind as the Liar paradox, whether or not they actually involve self-reference.

✉ Alexandre Billon
alexandre.billon@univ-lille3.fr

1 Université Lille Nord de France, 59000 Lille, France

2 UdL3, STL, 59653 Villeneuve d'Ascq, France

Published online: 16 February 2018

the set of sets that are not self-membered, has a dual involving the set of sets which are self-membered, etc.

Arguably, the difference between the paradoxes and their duals is that while the former hinge on an object which seems *overdetermined*, the latter hinge on an object which seems *underdetermined*. By a well-worn argument, if the Liar sentence Lr is true, it is false, and if it false it is true. Lr accordingly seems to be both false and true, and hence (semantically) overdetermined. The Truth-Teller sentence Tt seems, on the other hand, underdetermined. It says that it is true but it could be wrong, and nothing, it seems, could decide the issue. Similarly, the extension of Russell's set seems overdetermined (it should both include and exclude Russell's set) while that of the set of sets that are self-membered seems underdetermined (it is hard to see what could make it the case that it is, or that is not, itself self-membered).[2]

It is widely believed that because they hinge on an object's apparent underdetermination, the paradoxes' duals are less problematic than the paradoxes of which they are duals, and that they are so to speak 'under-paradoxical'. Peter Eldridge-Smith, who has coined the term 'hypodox' and provided their most extensive analyses (he has argued that hypodox involve under-determination while paradoxes involve over-determination and conjectured that each paradox has a paradoxical dual [see Eldridge-Smith (2007, en. 3, 2008)] says, for example, that "hypodoxes seem to be proto-paradoxes (Eldridge-Smith 2012)". Mackie (1973, p. 241) likewise explained that "in the ordinary paradoxes you have both undecidability and contradiction; in the truth-teller-variants [s.c. the duals] undecidability without contradiction". We might put the same point by saying that duals seem pathological but not paradoxical, or again, *merely pathological*.[3]

Connectedly, there is a wide consensus to the effect that whereas a treatment of the paradoxes requires one to debunk the reasoning that leads to a contradiction *and* to find out a proper characterization of the paradoxical object, a treatment of their

[2] For general presentations of the duals of classical paradoxes, see Mackie (1973), Eldridge-Smith (2008) and Armour-Garb and Woodbridge (2012). It is sometimes said that the Gödel sentence, on which Gödel's Incompleteness theorems hinge expresses its own unprovability, and can be considered as a 'provability version' of the Liar [see e.g. Gaifman (2006) on this analogy]. The Henkin sentence, which can be described as expressing its own provability [see Halbach et al. (2014)] is likewise a provability version of the Gödel sentence and might be taken to be the dual of the Gödel sentence.

[3] It is, to my knowledge, Herzberger (1970, pp. 148–149), who coined the term "pathological" to characterize the Truth-Teller. He called the classical paradoxes "negative pathologies", their duals "positive pathologies" and claimed that only the former give rise to paradoxes. Mackie (1973) investigated many duals, calling them "truth-teller variants of the paradoxes", and claimed that they involve undecidability without contradiction, while the paradoxes involve undecidability and contradiction. Woodbridge (2004) and Woodbridge and Armour-Garb (2005) claim that *semantic* paradoxes involve "apparent inconsistency" while their duals involve "apparent indeterminacy". Eldridge-Smith (2008) argues that duals involve underdetermination while paradoxes involve overdetermination. As emphasized by an anonymous referee for this journal, these claims are close but they are really different. Undecidability without contradiction, for example, is not the same as underdetermination (the former does not imply the latter). Nor is it exactly the same as as apparent indetermination (something might be indeterminate on first appareances but prove decidable). Importantly, these claims also differ in scope. To my knowledge, Eldridge-Smith's claim has the widest scope: he has conjectured that the duality between hypodoxes and paradoxes concerns all the paradoxes of self-reference, he has shown that it concerns most of them [including some non-classical ones that he himself discovered, see Eldridge-Smith (2015)] and he has convincingly argued that it even generalizes to time-travel paradoxes (Eldridge-Smith 2007).

duals only requires one to find out a proper characterization of the dual's object: no debunking would be needed in their case.

In this article, I argue that this consensus is misguided, some of the paradoxes' duals being paradoxical, and that some of the 'paradoxical duals' raise problems that have been neglected and are interestingly different from those raised by the paradoxes of which they are duals. I first focus on Richard's paradox, which, for reasons that will appear soon, is arguably the simplest case of a paradox with a paradoxical dual. After a brief presentation of this paradox and its dual (Sect. 3), I show that the latter is paradoxical (Sect. 4) and even as paradoxical as the former (Sect. 5). Interestingly, the classical, Richard-Poincaré solution to Richard's paradox does not generalize to this new paradox (Richard's paradoxical dual) in any obvious way (Sect. 6). In the last section, I show that my argument applies *mutatis mutandis* to a simple version the Liar, showing that its dual, the Truth-Teller is paradoxical as well (Sect. 7). Before doing all that, however, I need to say a few words about the difference between genuine paradoxes and proto-paradoxes or mere pathologies (Sect. 1), and about the logic and notations of definitions (Sect. 2).

1 Paradoxes and mere pathologies

The claim of this paper—that some duals are full paradoxes rather than not mere pathologies—could easily be misunderstood and wrongly toned down. After all, one might think, everyone admits that duals, such as the Truth-Teller, are queer, even though not exactly in the same way paradoxes are. And calling the Truth-Teller "paradoxical" or "pathological", one might pursue, is a just terminological decision as long as one understands the differences between the Truth-Teller (which involves apparent semantic underdetermination) and the Liar (which involves apparent semantic overdetermination).

I disagree. There is a natural way to draw a line between full paradoxicality and mere pathlogy, and a way to frame the main claim of this article that makes it inconsistent with the common take on the Truth-Teller and other duals. A paradox, I take it, is the apparently valid derivation of an apparently unacceptable conclusion from apparently acceptable premises. As noticed by Sorensen (2001, p. 1) suggests that there are degrees of paradoxicality: the more plausible the premises and the derivation, the more implausible the conclusion, the stronger the paradox. In the classical paradoxes of self-reference, such as the Liar and Richard's paradox we derive a contradiction (very implausible, arguably) from very plausible premises. We thus have a full paradox for sure.

Now it is true that we can also derive a contradiction from certain claims about duals. This is what Eldridge-Smith (2012) (quoted above) means when he claims that that "by adding principles, a hypodox may become a paradox". We may indeed, with the help of some extra premises, derive a contradiction (or something as intuitively implausible as a contradiction) from duals. If Eldridge-Smith says that duals are "proto-paradoxes" nevertheless, it is because he believes that these additional premises are less plausible than the ones that give rise to genuine paradoxes. Consider the Truth-Teller, for example. The general opinion on the Truth-Teller is that it is apparently indeterminate whether is is true or fase and that no non-semantic fact can decide its truth-value. It is thus commonly admitted that the Truh-Teller threatens a principle called "the supervenience of semantic facts on non-semantic facts" [see Kremer (1988) as well as Gupta and Belnap (1993, pp. 31, 89–91, 104–9)], to the effect that there can be no difference in semantic facts without a difference in non semantic facts (we suppose that the interpretation of non-semantic vocabulary is fixed). If this supervenience claim were as intuitively plausible, or nearly as intuitively plausible, as the "premises" of the Liar paradox, the Truth-Teller would *ipso facto* be paradoxical. It is not however, nearly as plausible as these, and this explains why the Truth-Teller is not ususally considered a paradox.[4]

Consider, to take another example, the 'Open Pair' [$L1$ = '$L2$ is not true', $L2$ = '$L1$ is not true', see Woodbridge and Armour-Garb (2005)]. As the reader will check, it can easily yield a contradiction if it is assumed that $L1$ and $L2$ have the same truth-value. This last premise seems however quite plausible to many, and this is the reason why, although the Open Pair displays some similarities with the Truth-Teller, it is often deemed genuinely paradoxical (Sorensen (2001) calls it, for example, the 'No-No' paradox).

We are now in a position to make the claim of this paper more precise: I will argue that some duals are *as paradoxical* or at least *nearly as paradoxical*, as the paradoxes of which they are duals, and more specifically, that Richard's dual falls in the first category and the Truth-Teller in the second.[5] This is *not* a widespread claim. Almost all contemporary logicians and philosophers consider, for instance, that the Truth-Teller is not paradoxical or at least not nearly as paradoxical as the Liar. To take some examples among many, Herzberger (1970, pp. 148–149), Kripke (1975, p. 693), Simmons (1993, p. 3), Gupta and Belnap (1993, pp. 188–189), Yaqub (1993, p. 41), Goldstein (2006) or Scharp (2013, p. 255) all claim that the Truth-Teller is not paradoxical.[6] As already noted, as they take duals not to be paradoxical, most

[4] Drawing on Kripke (1975)'s semantics, Kremer (1988) argues that one can ascribe truth, falsity or "neither true or false" to the truth-teller and that the latter accordingly refutes the supervenience of the semantic. Gupta and Belnap (1993)'s develop a revision theory of truth that allows restoring the supervenience claim by attributing to the truth-teller a definite semantic characterization in terms of revision rules. Yaqub (1993) has argued on more general grounds that the supervenience of semantics does not hold (see pp. 39–42) and he has developped an interpretation of the revision-theoretic formalism in which the supervenience of the semantics fails.

[5] In order to do that, moreover, I will not rely on the claim that the supervenience of semantic facts is nearly as plausible as the principles on which we rely to derive a contradiction from the Liar. I will rely, instead, on the logic of definitions.

[6] In fact, the only researchers I know who disagree are Austin (1950), who claims that "a statement to the effect that it is itself true is every bit as absurd as one to the effect that it is itself false (p. 122, fn. 22)", and Priest and Mortensen (1981), who grant that the Truth-Teller is genuinely paradoxical. Both Austin (1950) and Priest and Mortensen (1981) only consider the Liar's dual, though, and the reasons they invoke are quite different from those presented here. The former claims that "a statement to the effect that it is itself true is every bit as absurd as one to the effect that it is itself false (p. 122, fn. 22)", and that "It takes two to make a truth. (…) Hence (…) a statement cannot without absurdity refer to itself (p. 118, fn. 13)". Priest and Mortensen (1981), argue that even though we have reasons to believe that the Truth-Teller is neither true nor false, (for there is no proof that the Truth-Teller is true, and no proof that it is false) there is a proof that it is either true or false. Finally, Beall (2009, pp. 14–5) recently suggested that the Truth-Teller is both true and false, admitting that this claim is "ultimately defeasible".

reasearchers also believe that their treatment does not involve debunking a paradoxical reasoning.

2 On definitions

Following a suggestion of Richard (1905), Poincaré (1906), Russell (1906) and others have claimed that the classical paradoxes hinge on certain viciously circular definitions. We will soon make explicit the role of circular definitions in Richard's paradox and its paradoxical dual. A few preliminary words on the notations and the logic of definitions will prove useful. Paraphrasing Gupta (2015), we can say that definitions involve three elements:

- a defined term 'X' (and a corresponding defined object X),
- an expression containing the defined term '$\ldots X \ldots$' which is called the *definiendum* ,
- and another expression '$-------$' that is 'equated' by the definition with the definiendum, and which is called the *definiens*.

They can be represented thus[7]:

$$X : \ldots X \ldots =_{df} -------$$

The definiens and the definiendum must be of the same logical category, but the defined term need not be identical to the definiens. I can for example define x by "$x : 2x =_{def} 2$" (this should read "x is defined by the fact that $2x$ is identical to 2") or even by a sentential definition such as "$x : (y = x) =_{def} (y = 1)$" (this should read "$x$ is defined by the fact that $y = x$ is equivalent to $y = 1$).

It is common to work with embedded definitions of this form:

$$\begin{array}{lll} X : \ldots X \ldots & =_{df} & \ldots - Y - \ldots \\ Y : -Y- & =_{def} & *Z* \\ Z : \ldots & \ldots & \ldots \end{array}$$

In such a case we can, so to speak 'unfold' the first definition and conclude that X is defined in terms of Z[8]:

$$X : \ldots X \ldots =_{df} \ldots * Z * \ldots$$

This move, which typically allows one to eliminate terms that are not semantically primitive, is arguably essential to our use of definitions. It is usually justified by a rule,

[7] The defined term is often omitted when it is specified by the context. The definition symbol '$=_{def}$' should never be confused with the mere identity symbol'='. Just like the defined term corresponds to a defined object the definiendum corresponds to an 'identifying object' and the definiens to 'defining object'.

[8] Notice that even if the two definitions above were stipulative, this would not be the case of the one below. X has already been introduced (and defined) above, so the definition of X below cannot be a stipulative definition. It is a so-called descriptive or explanatory definition: it specifies the object or the meaning of the term that has already been defined.

called *Definiendum-Elimination*, which is part of what Gupta (2015, §2–2) dubs "the traditional account of definitions" as "generalized identities," to the effect that we can replace the occurrence of a definiendum (here the definiendum of Y, $-Y-$) by its definiens (here the definiens of Y, $*Z*$).

In what follows, in order to derive a new definition D' of an object x (or of the corresponding term 'x') from another definition D of that object, I will only use *Definiendum-Elimination*, replacing the occurrences of a definiendum in the definiens of D. [9] I will also suppose that the first definition succeeds in picking at least (or at most) one object iff the second does.[10]

Now, we say that a definition of the form

$$X : X =_{df} - - - - - - - -$$

is (explicitly) circular if the definiens '$- - - - - - - -$' contains the defined term, or if it generalizes over totality that contains the defined object.[11] It is implicitly circular, on the other hand, if it can yield a circular definition through the elimination of non-primitive terms in its definiens.

Circular definitions are then those whose definiens refer directly (through a name or a quantifier) or indirectly, (through expressions whose definiens refer directly to the

[9] A referee for this journal alerted me of an interesting caveat. Suppose a is defined by

$$a : a =_{def} b$$

Replacing the definiendum 'a' by the definiens 'b' *in this very definition*, we would get that $a : \quad b =_{def} b$, but a is certainly not defined like that. Wagner (2017, §6) puts forward other counter-examples involving quotations. Suppose the number one is defined by the fact that it is the least (strictly) positive interger. 'One' is three letters long. It does not follow that 'the least (strictly) positive integer' is three letters long. Interestingly, however, none of these counter-examples threaten our practice of unfolding embedded definitions (this practice does not require replacing occurrences of a definiendum in the very definiendum of a definition, and we can deal with the occurrences of a definiendum within quotation marks by separating its contribution to the morphology and the meaning of the resulting, quoted expression). In what follows, I will only use a very restricted *Definiendum-Elimination* rule, to the effect that *in the definiens of a given definition*, we can replace occurrences of a definiendum by its definiens when they are not under quotation marks. We shall see later that some researchers expressed important reservations about Definiendum-Elimination, even in this restricted form, precisely because some definitions are problematic and seem to generate paradoxes (p. 28). It should be noticed, finally, that everything I have said about *Definiendum-Elimination* applies *mutatis mutandis* to the dual rule of *Definiendum-Introduction* to the effect that we can substitute a definiendum to its definiens.

[10] This last supposition can be be derived from *Definiendum-Elimination* and from its converse (and equally plausible) principle, namely *Definiendum-Introduction*. Together, these principles allow one to derive D' from D and D from D'.

[11] There is a close connection between the two disjuncts of this definition. Consider the claim that if $X \in D$

$$X : X =_{df} - - - \quad \forall x \in D,\ \Phi(x) - - - -$$

then

$$X : X =_{df} - - - \Phi(X) \wedge \quad \forall x \in D,\ \Phi(X) - - - -$$

Provided that this plausible claim is accepted, a definition whose definiens generalizes over a totality containing the defined object will yield one whose definiens contains the defined term.

defined object) to the defined object.[12] Circular definitions have been incriminated as generating Richard's and other paradoxes and some believe that they should simply be banned. It should however be emphasized that they are generally quite benign. Consider the following circular definitions of real numbers through equations:

$$\begin{aligned} a &: a =_{def} 2a - 1 \\ b &: b =_{def} b + b^2 + 1 \\ c &: c =_{def} c \end{aligned}$$

The first succeeds in picking 1. The second, it is true, fails to pick a unique real number because it picks no real number and the last fails to pick a unique real number because it picks too many. But even if they fail, these last two definitions are quite benign in that they fail, so to speak, trivially: we have no good reason to believe that they should succeed and their failures do not accordingly generate any paradox. Moreover, be they inductive, recursive or impredicative, circular definitions are ubiquitous in mathematics, and it is far from clear that they could de dispensed with (Feferman 2005).[13,14] Finally, even if we could ban explicitly circular definitions, it is hard to see how we could, in our ordinary practice, ever be sure of avoiding *implictly circular definitions*, definitions, that is, that are not explictly circular but yield circular definitions when the non-primitive terms in their definiens are eliminated through Definiendum-Elimination. Yet implictly circular definitions seem to raise the same difficulties as their explicitly circular counterparts. For all these reasons it seems hard to put a general ban on circular definitions.[15] Gupta and Belnap (1993) and others have

---

[12] To my knowledge, there is no canonical way to define circular definitions. A definition which is circular in my sense is circular, intuitively, in that we seem to run into circles when we try to use the definition to find out what it refers to. It might however be reducible to a definition that is not circular, and need not, in that sense, be *essentially* circular [see Gupta (2015, §2–7)].

[13] The definition of a function (or predicate) on natural integers is typically called 'recursive' when it contains (i) an initialization clause, defining its initial value, (ii) and an inductive clause, specifying the relationship between the functions's value on n and on $n + 1$ for an arbitrary n. Inductive definitions are, more broadly, definitions appealing to an initialization clause and an inductive clause. Impredicative definitions have a definiens generalising over a totality to which the defined object belongs. It should be emphasized however that many authors use inductive and recursive interchangeably, and that the concept of impredicativity have been used to refer to different (if related) phenomena [cf. Feferman (2005)]

[14] Circular definitions seem essential to ordinary arithmetics (cf. e.g. the definition of natural numbers), classical analysis (cf. e.g. the least upper bound principle) and set theory (cf. e.g. the axiom of separation). They are also integral to the practice of defining objects as the fixed points of certain functions. Feferman has conjectured that the part of mathematics *that is indipensable to the natural sciences* might dispense with predicative definitions *provided that it already accepts arithmetics* [see Feferman (2005) and the works cited there].

[15] The problem here is somewhat analogous to that confronting the prospects of banning self-reference on the ground that it has been accused of generating paradoxes. Self-reference often seems benign, and even if we could ban direct self-reference, there seems to be no way to be safe from indirect, empirical self-reference [see Kripke (1975, pp. 690–693)'s classical remarks]. It should also be emphasized that even if definitions in so called 'normal form' [see Gupta (2015)] are guaranteed to be non-circular, such definitions are much too restrictive. They not only forbid circular definitions (which, as we have seen seem essential to mathematics and to our ordinary practice of defining objects and terms) they also forbid so called implicit definitions such as the implicit definition of a set by the axioms of set theory.

argued that we could, and that we should embrace all circular definitions (see p. 28). My claim here is more modest and tentative, however. It is simply that that unless it has proven impossible to do so without too much trouble, we should generally accept instances of circular definitions.

3 Richard's paradox and its dual

Just like there are many versions of the Liar paradox, there are many versions of Richard's paradox, and many ways to regiment (and expand on) Richard (1905)'s very concise presentation of the eponymous paradox [see Simmons (1993) and Priest (2006b) for two influential versions of the paradox]. In this section, I put forward what I take to be the gist of the paradox. My presentation tries to remain faithful to Richard's idea that his paradox really is a *definability* paradox (rather, say, than a denotation one), a paradox that hinges explicitly, accordingly, on the logic of definitions (rather, say, than on the logic of denotation and definite descriptions). It differs, in that respect, from other influential presentations—upon which I shall only touch in footnotes.

Let E be the set of real numbers that can be defined in a finite number of words. Finite definitions of real numbers are countably infinite. Assuming that the definition of a real number defines at most one real number, E is at most countably infinite. As each natural number can be defined by the words "the number n", E is in fact countably infinite. Accordingly, there is a bijective / one-one enumeration from $\mathbb{N}^*$ to E. Let h be one such enumeration.[16]

Let us write $(h(i))_j$ the jth decimal of the normalized decimal representation of $h(i)$ (the normalized representation is the one that prefers '0.1' to te coreferring '0.099999...') (Table 1). Let g be a function from the ring of integers modulo 10 ($\mathbb{Z}/10\mathbb{Z}$) to itself.[17] We can define *the number* α *whose integral part is* 0 *and whose* ith *decimal is obtained by applying* g *to the* ith *digit of the* ith *row* $((h(i))_i)$ *of the enumeration of the set E:*

$$\alpha : \alpha =_{def} \sum_{i \in \mathbb{N}^*} \alpha_i 10^{-i}$$

where for all i in $\mathbb{N}^*$,

$$\alpha_i : \alpha_i =_{def} g(h(i)_i) \qquad (1_{def})$$

[16] As noticed by Simmons (1993, p. 27), Richard explicitly constructs a surjective/onto enumeration h, but he wrongly assumes that the enumeration he has constructed is bijective. Although it is not the one Richard picked out, there is in fact a bijective enumeration of E. To explicitly construct a bijective enumeration, one can list all the finite definitions of numbers in lexicographic order and take h to be the function defined recusrively as follows: (i) it associates to the first definition in the list, the number in E defined by this definition, (ii) and it associates to the nth definition the smallest number in E that is not in the image of $\{1, \ldots, n-1\}$ by h. Even though, I think we can obtain Richard's paradox (and its dual) without relying on the fact that h is injective , relying on this fact makes things slightly simpler.

[17] The ring of integers modulo 10 ($\mathbb{Z}/10\mathbb{Z}$) is the ten first numbers 0, 1, ... 9 endowed with the addition and multiplication operations, and where (to put it rather roughly) it is assumed that for all number x, $x = x + 10$.

Table 1 h is an exhaustive enumeration of E. The decimal representation of $h(i)$ is $h(i)_0.h(i)_1h(i)_2 \dots h(i)_j \dots$ The number α is obtained by applying g to each decimal of the diagonal number $0.h(1)_1h(2)_2..h(i)_i \dots h(j)_j \dots$

E	Decimal representations of the members of E							
	$\underline{0.}$							
$h(1)$	$h(1)_0.$	$\underline{h(1)_1}$	$h(1)_2$	$\dots$	$h(1)_i$	$\dots$	$h(1)_j$	$\dots$
$h(2)$	$h(2)_0.$	$h(2)_1$	$\underline{h(2)_2}$	$\dots$	$h(2)_i$	$\dots$	$h(2)_j$	$\dots$
$\vdots$	$\vdots$							
$h(i)$	$h(i)_0.$	$h(i)_1$	$h(i)_2$	$\dots$	$\underline{h(i)_i}$	$\dots$	$h(i)_j$	$\dots$
$\vdots$	$\vdots$							
$h(j)$	$h(j)_0.$	$h(j)_1$	$h(j)_2$	$\dots$	$h(j)_i$	$\dots$	$\underline{h(j)_j}$	$\dots$
$\vdots$	$\vdots$							

α is defined in a finite number of words (by the expression in italics above) so it should belong to the set E of the number so defined.[18]

As h is a bijective enumeration of the numbers of E and as α is in E, there is exactly one natural integer x such that $h(x) = \alpha$. We can define, by stipulation, n_α as being this number[19]:

$$n_\alpha : h(n_\alpha) =_{def} \alpha \qquad (2_{def})$$

(1_{def}) is a series of definitions whose n_αth element (take $i = n_\alpha$) is:

$$\alpha_{n_\alpha} : \alpha_{n_\alpha} =_{def} g(h(n_\alpha)_{n_\alpha}) \qquad (3_{def})$$

By Definiendum-Elimination, we can replace '$h(n_\alpha)$' with 'α' in the definiens of (3_{def}), yielding:

$$\alpha_{n_\alpha} : \alpha_{n_\alpha} =_{def} g(\alpha_{n_\alpha}) \qquad (4_{def})$$

Richard's paradox

(4_{def}) entails that:

$$\alpha_{n_\alpha} = g(\alpha_{n_\alpha}) \qquad (4_{=})$$

18 I follow Jules Richard in construing the eponymous paradox as a paradox of *definability* rather than *denotation*. Under that denotation construal, the paradox relies on the denotation schema, to the effect that 'a' refers (if at all) to a, and on the principles governing definite descriptions [see Priest (2006b) for the denotation construal]. On the definability construal, it relies on the logic of definitions, and in particular on the claim that the *definiens* of a definition defines (if anything at all) the *definiendum* of that definition. We shall see that it also relies on Definiendum-Elimination.

19 To get a definiens that is identical to the defined term below, we could also introduce the function h^{-1} (defined by the fact that for all x, $h(h^{-1})(x) = h^{-1}(h(x)) = x$). We would then have $n_\alpha : n_\alpha =_{def} h^{-1}(\alpha)$.

and thus, by existential generalization, that:
(i) g has *at least one* fixed point in $\mathbb{Z}/10\mathbb{Z}$: $\exists x \in \mathbb{Z}/10\mathbb{Z} : x = g(x)$.[20]
Richard's paradox stems from the fact that we can choose g such that $g = g_R$ (R stands for 'Richard'), where g_R provably has no fixed point in $\mathbb{Z}/10\mathbb{Z}$, thus contradicting (i). Just take $g_R(x) = (x + 1) \pmod{10}$, the function that associates 1 to 0, 2 to 1,... 0 to 9. Then α is defined as *the number α whose integral part is* 0 *and whose* ith *decimal is obtained by applying the function add 1 (mod 10) to the* ith *digit of the* ith *row* $((h(i))_i)$ *of the enumeration of the set E*. And α_{n_α}, must be a fixed-point of the function add 1 (mod 10) which has none.

The contradiction is rooted, here, in the fact that the definition of α_{n_α} (4_{def}) overdetermines its object whereas it should not: by (4_{def}), α_{n_α} must be both equal to α_{n_α} and to $g_R(\alpha_{n_\alpha})$ which differs from α_{n_α}. More deeply, it is rooted in the fact that the definition of α (rather than α_{n_α}) overdetermines its object: the latter must both belong to the set E and—as $g_R(\alpha_{n_\alpha}) \neq \alpha_{n_\alpha}$—not belong to that set.

Richard's Dual
We get Richard's dual if we choose g such that $g = g_{DR}$ ('DR' stands for 'Dual Richard') where g_{DR} provably has more than two fixed points in $\mathbb{Z}/10\mathbb{Z}$. We can choose, for example, g_{DR} as the identity function over $\mathbb{Z}/10\mathbb{Z}$, which has every number in $\mathbb{Z}/10\mathbb{Z}$ for fixed points. This is exactly what Mackie (1973, 298) did, presenting what is, to my knowledge the first instance of Richard's dual. In that case α is defined simply as *the number α whose integral part is* 0 *and whose* ith *decimal is the* ith *digit of the* ith *row* $((h(i))_i)$ *of the enumeration of the set E.*

The nature of α_{n_α} seems underdetermined here. Indeed, the only thing the identity ($4_=$) ('$\alpha_{n_\alpha} = g_{DR}(\alpha_{n_\alpha})$') tells us is that α_{n_α} is a fixed point of g_{DR}, and thus, if this function is identity, that α_{n_α} is self-identical. This is not very informative and we seem to be left with an unsolvable underdetermination.

4 The paradoxical character of Richard's dual

Richard's dual, I claim, is in fact paradoxical. This is because our definition (4_{def}) of α_{n_α} ('$\alpha_{n_\alpha} : \alpha_{n_\alpha} =_{def} g(\alpha_{n_\alpha})$') tells us more than the mere identity ($4_=$) ('$\alpha_{n_\alpha} = g(\alpha_{n_\alpha})$') which follows from it.

Indeed, unlike ($4_=$), (4_{def}) does not only entail that α_{n_α} is a number x in $\mathbb{Z}/10\mathbb{Z}$ such that $x = g(x)$, it entails that it is *the only such number*, and thus that there is only one such number. In other words, it does not only entail that α_{n_α} is a fixed point of g in $\mathbb{Z}/10\mathbb{Z}$, but also that it is the only such fixed-point.

We can make this point more explicit by putting our definitions in 'sentential form'. The definition (4_{def}) is equivalent to the following definition in sentential form:

$$\alpha_{n_\alpha} : (x = \alpha_{n_\alpha}) =_{def} (x \in \mathbb{Z}/10\mathbb{Z} \wedge x = g(x))$$

[20] A fixed point of a function f in a domain D is a member x of D such that $x = f(x)$.

This definition says that α_{n_α} is defined by the fact that for any x, $(x = \alpha_{n_\alpha})$ is equivalent to $(x \in \mathbb{Z}/10\mathbb{Z} \wedge x = g(x))$. It thus entails that this equivalence indeed holds and that:

$$\forall x,\ (x = \alpha_{n_\alpha}) \leftrightarrow (x \in \mathbb{Z}/10\mathbb{Z} \wedge x = g(x)) \tag{5}$$

And accordingly, it entails both that:

(i) g has *at least one* fixed point in $\mathbb{Z}/10\mathbb{Z}$: $\exists x \in \mathbb{Z}/10\mathbb{Z} : x = f(x)$ ($\alpha_{n_\alpha} = \alpha_{n_\alpha}$ entails that $\alpha_{n_\alpha} \in \mathbb{Z}/10\mathbb{Z} \wedge \alpha_{n_\alpha} = f(\alpha_{n_\alpha})$ by (5), which entails that $\exists x \in \mathbb{Z}/10\mathbb{Z} : x = g(x)$ by existential generalization)

(ii) g has *at most one* fixed point in $\mathbb{Z}/10\mathbb{Z}$: $\forall x_1, x_2 \in \mathbb{Z}/10\mathbb{Z}$, $(x_1 = g(x_1) \wedge x_2 = g(x_2)) \rightarrow x_1 = x_2$ ((5) entails that $\forall x \in \mathbb{Z}/10\mathbb{Z}$, $x = g(x) \rightarrow x = \alpha_{n_\alpha}$).

Now, just like we can choose g such that $g = g_R$ and has no fixed point, and thus contradict (i)—this was Richard's paradox—we can choose g such that $g = g_{DR}$ where g_{DR} has two or more fixed points, like the identity function, and thus contradict (ii)—this is Richard's dual. In the latter case, then, we get a contradiction as well, and we get it quite easily too. Hence the paradox.

In the case of Richard's paradox, the contradiction stems from the fact that the definition of α_{n_α} (and more fundamentally that of α) *overdetermines* its object whereas (given the logic of definitions) it should not. In the case of Richard's dual, it stems from the fact that the definition α_{n_α} (and more fundamentally that of α) *underdetermines* its object whereas (given, again, the logic of definitions) it should not.

The semi-technicalities above should not mask the relative simplicity of Richard's dual. A common, plain English, presentation of Richard's paradox consists in saying that we can define a number (the number α whose integral part is 0 and whose ith decimal is obtained by adding 1 (mod 10) to the ith digit of the ith row of the enumeration of the set E) which cannot be defined on pain of contradiction. Richard's dual consists similarly in the fact we can define a number (the number α whose integral part is 0 and whose ith decimal is the ith digit of the ith row of the enumeration of the set E) which cannot be defined on pain of a contradiction. The main difference stems from the way the contradiction is derived from the claim that the definition succeeds. This derivation relies, in Richard's case, on the fact that no value of the defined term is such that the value of the definiens is identical to that of the definiendum. It relies, in the case of Richard's dual, on the fact that more than one value of the defined term is such that the value of the definiens is identical to that of the definiendum.

5 An equally paradoxical dual?

Richard's paradox and its dual in fact differ in two ways. First, as we have seen, while we can get Richard's paradox from the mere identity ($4_=$), its dual relies on the definition (4_{def}).

More interestingly, Richard's paradox and its dual hinge on two different assumptions regarding our initial definition of α (1_{def}). Richard's paradox hinges on the assumption that:

- *Existence* Our initial definition of α (1_{def}) succeeds in picking *at least one* number, in the sense that there is at least a real number x such that $\forall i \in \mathbb{N}^*,\ x_i = g(h(i)_i)$, (from which it follows that the definition of α_{n_α} (4_{def}) should pick at least one number in $\mathbb{Z}/10\mathbb{Z}$ and thus that g_R should have at least one fixed point.)

Richard's dual hinges, on the other hand, on the assumption that:

- *Uniqueness* Our initial definition of α (1_{def}) succeeds in picking *at most one* number, in the sense that there is at most one x such that $\forall i \in \mathbb{N}^*,\ x_i = g(h(i)_i)$ (from which it follows that the definition of α_{n_α} (4_{def}) should pick at most one number in $\mathbb{Z}/10\mathbb{Z}$ and that g_{DR} should have at most one fixed point)

Importantly, the premises on which Richard's dual relies to get a contradiction are arguably as plausible and as entrenched as those on which Richard's paradox relies. It is quite hard, first, to see how one could argue that the definitions (1_{def}) and (2_{def}) do entail ($4_=$), but not (4_{def}). (4_{def}) is in that sense as plausible as ($4_=$).

In the same way, *Uniqueness* seems to be at least as intuitively plausible as *Existence*.

Now, a paradox is, as we have seen, the apparently valid derivation of an apparently unacceptable conclusion from apparently acceptable premises, and the more plausible the premises and the derivation, the more implausible the conclusion, the stronger the paradox. As in Richard's paradox and its dual we derive a contradiction from equally acceptable premises, and the derivation is in each case equally plausible, the paradox and its dual are equally paradoxical. This implies that the dual equally needs to be treated and that its treatment cannot merely consist in characterizing α: it should also allow us to debunk the paradoxical reasoning above. Unfortunately, we shall see that the classical solution to Richard's paradox is unable to do that.

6 Richard's paradox, its dual and vicious definitions

The classical solution to Richard's paradox denies *Existence* and claims that our initial definition of α (1_{def}) fails to pick a number. This was already part of Richard (1905)'s own solution: he argued that such a definition fails to pick a number because it is meaningless. Despite the fact that it is well-formed and that we seem to understand it very well, the uninterpreted expression defining α ('the number whose ith decimal is …') would be associated with no interpretation in our language. The classical solution does not obviously generalize to Richard's dual. Indeed, it does *not* allow one to reject *Uniqueness* (a definition that picks no number can, and in fact must pick at most one number). It does not accordingly prevent one from deriving the conclusion that the definition of α picks at most one number and that g_{DR} should accordingly have at most one fixed point. In order to solve Richard's dual, it seems that we should deny *Uniqueness* rather than *Existence*, and claim that the uninterpreted expression defining α ('the number whose ith decimal is …') should be associated at once *with more than two* interpretations and references (rather than none) in our language. Only this would allow g_{DR} to have more than two fixed points.

Finally, Richard (1905), quickly followed by Poincaré (1906) and Russell (1906), argued that the definition of α is *vicious* because of some form of *circularity*. On the present account of the paradox, the definition of α and more specifically of its n_α's decimal is indeed implicitly circular and vicious. It is implicitly *circular* as eliminating the non-primitive terms of (the definiens of) α_{n_α}'s definition (3_{def}) yields the definition of α_{n_α} (4_{def}) whose *definiens* contains 'α_{n_α}'.

This definition is also *vicious*, both in the case of Richard's paradox and of its dual, because it allows us to derive a contradiction that:

- does not involve the newly defined term (contradiction to the effect that g has and does not have at least (respectively, at most) one fixed point)
- and was not arguably derivable before this term was defined.

It thus violates what Suppes (1957) calls the 'Constraint of Non-Creativity' on good definitions to the effect that the definition of a new term should not allow one to derive new claims that do not involve that term [see also Belnap (1993, pp. 122–138) and Gupta (2015) who call this constraint 'Conservativeness'].[21]

In the case of Richard's paradox the definition of α is vicious because it violates *Existence*, in the case of its dual, it is vicious because it violates *Uniqueness*.[22]

Richard's dual is as paradoxical as Richard's paradox, and the classical solution the the latter does not generalize to the former. The new paradox, in other words, appeals to new moves. This is already an interesting result.[23] There is, however, more to come.

7 Other paradoxical duals? The example of the liar

I have focused on Richard's paradox because my argument for the paradoxical nature of duals crucially depends on the role of (at least implicitly) circular definitions in generating the paradoxes of self-reference. This role was first emphasized by Jules Richard, and it is particularly obvious in the eponymous paradox. I would like to show, however, that the paradoxical nature of duals is not restricted to the case of Richard's paradox. In this section, I will argue that my argument applies *mutatis mutandis* to a

[21] Quite generally, if the logic is classical, in order not to violate this 'Constraint of Non-Creativity' and to be benign, the definition of a name must satisfy both *Existence* (i.e. pick at least one object) and *Uniqueness* (i.e. pick at most one object). If it did not satisfy Existence, the defined term would not refer (which would contradict, for example, the classical rule of Existence-Generalization). If it violated Uniqueness, the defined term would refer to more than one object (which would entail a contradiction to the effect that that these two different objects are identical). See the related remarks in Gupta (2015§2-4)'s discussion of definitions in 'normal form'.

[22] Interestingly, however, it seems that circularity as such is to blame for the definition's vicious character. It can indeed be checked that if $g = g_{BR}$ ('BR' stands for Benign Richard) is a function that has only one fixed point in $\mathbb{Z}/10\mathbb{Z}$ (say $g_{BR}(x) = 2x + 1 \pmod{10}$, which has 9 for unique fixed point), the definition of α will still be circular, but α_{n_α} will pick a unique and definite number (the fixed point of g_{BR}) and the same will arguably be true of α (arguably, that is, if, as is commonly assumed in discussions of Richard's paradox, the other decimals of α are not problematic).

[23] Assuming that paradoxes of the same kind must admit a solution of the same kind (this is what Priest (1994) calls the "principle of a uniform solution"), this implies, moreover, either that the classical solution to Richard's paradox is not a good solution, or that Richard's paradox and its dual, depsite their resemblance and their striking symmetry, are not of the same kind. Given that it shows that Richard's dual is paradoxical too and arises from very similar principles as Richard's paradox itself, the present article provides some reasons to favour the first option. Even if I won't come back to it, similar remarks could apply to the Liar and its Truth-Teller dual, with which I deal below.

common version of the Liar and its Truth-Teller dual. This version, which makes the argument simpler, relies on natural language and achieves self-reference through the use of quotations.

Suppose that we define the P-Teller sentence Ptr as follows:

$$Ptr : Ptr =_{def} \text{`}Ptr\,is\,P' \tag{6_{def}}$$

where 'P' is substituted by the sentential predicate 'true' or 'false'. We assume that the language is fixed and that all sentences are interpreted. Such a stipulative, explicitly circular definition, defining Ptr by the fact that it is identical to 'Ptr is P' is, as I have said, a very common way to introduce the Liar and the Truth-Teller, and it provides us with a very common version of these pathologies.

Let us write $V(s)$ the semantic value of a sentence s and $M(s)$ its morphology, that is, the sequence of mark associated with the interpreted sentence ('V' and 'M' are function terms of the metalanguage).

I will consider that V is a *unary* function that takes just a sentence as argument and whose value is a semantic value, rather than a binary function that depends of 'circumstances of evaluation' as well as sentences. This means, for example, that if the sentences we consider are context-sensitive, the semantic-values will be functions from circumstances of evaluation to truth-values (or intensions) rather than truth-values (or intensions). Except in fn. 30, however, I will also suppose, for simplicity, that our sentences are not context-sensitive and that the setting is extensional, so that semantic-values are just truth-values.

An interpreted sentence s is defined as the couple of a morphology $M(s)$ and a semantic value $V(s)$, so our definition (6_{def}) of 'Ptr' says that the morphology of Ptr is defined by its identity with M'(Ptr is P') and its semantic value is defined by its identity with V('Ptr is P'). It abbreviates the following definition of an ordered pair:

$$\begin{bmatrix} M(Ptr) \\ V(Ptr) \end{bmatrix} : \begin{bmatrix} M(Ptr) \\ V(Ptr) \end{bmatrix} =_{def} \begin{bmatrix} M(\text{`}Ptr \text{ is } P\text{'}) \\ V(\text{`}Ptr \text{ is } P\text{'}) \end{bmatrix} \tag{7_{def}}$$

(I use square brackets for tuples and series).

We will say that a sentential predicate 'P' is *functionally defined* if there is a function g_P such that:

$$V\,(\text{``s' is P'}) : V\,(\text{``s' is } P\text{'}) =_{def} g_{\mathcal{P}}\,(V\,(\text{`s'})) \tag{8_{def}}$$

(where 's' is substituted with a sentence)

The Truth and Falsity predicates seem to be examples of functionally defined predicates. It is indeed arguable that the following version of the Truth-Schema holds:

$$V\,(\text{``s' is true'}) : V\,(\text{``s' is true'}) =_{def}\,(V\,(\text{`s'})) \tag{T_{def}}$$

If this assumption, which we shall discuss in due course, is granted, then 'true' is functionally defined, and g_{true} is the identify function over truth and falsity.

It is similarly arguable that the following version of the Falsity-Schema holds:

$$V(\text{“s' is false'}) : V(\text{“s' is false'}) =_{def} 1 - (V(\text{‘s'})) \qquad (F_{def})$$

If so, ‘false’ is functionally defined and g_{false} is the permutation of truth with falsity.

By Definiendum-Elimination and (8_{def}), we can eliminate ‘V(‘Ptr is P’)’ from the definition (7_{def}) and get:

$$\begin{bmatrix} M(Ptr) \\ V(Ptr) \end{bmatrix} : \begin{bmatrix} M(Ptr) \\ V(Ptr) \end{bmatrix} =_{def} \begin{bmatrix} M(\text{‘}Ptr \text{ is } P\text{'}) \\ g_{\mathcal{P}}(V(Ptr)) \end{bmatrix} \qquad (9_{def})$$

that is,

$$Ptr : Ptr =_{def} f_{\mathcal{P}}(Ptr)$$

where $f_{\mathcal{P}}$ is the function that associates with the sentence (x, y) whose morphology is x and whose semantic value is y, the sentence $(M(\text{‘}Ptr \text{ is } P\text{'}), g_{\mathcal{P}}(x, y))$. This implies that $f_{\mathcal{P}}$ has a unique fixed point. Now, as M(‘Ptr is P’) does not depend on Ptr (it depends on the sequence of marks M(‘Ptr’), not on Ptr itself), $f_{\mathcal{P}}$ has a unique fixed point iff $g_{\mathcal{P}}$ has a unique fixed point. We can accordingly conclude that $g_{\mathcal{P}}$ has exactly one fixed point.[24]

Like in the case of Richard’s paradox and its dual, this implies not only that $g_{\mathcal{P}}$ has at least one fixed point but also that it has at most one:

(i) $g_{\mathcal{P}}$ has at least one fixed point.
(ii) $g_{\mathcal{P}}$ has at most one fixed point.

The Liar

We get a version of the Liar paradox if ‘P’ is substituted with ‘false’. We have seen indeed that g_{false} is the permutation of truth with falsity. The latter has no fixed point (on pain of triviality), which contradicts (i). Hence the paradox.

The Truth-Teller

We get a version of the Truth-Teller if ‘P’ is substituted with ‘true’. As we have seen, g_{true} is the identity over truth and falsity. The latter has at least two fixed points (on pain of triviality again), which contradicts (ii). Hence the paradox.

[24] One could have derived this conclusion more simply by appealing to the fact that both lines of the definiton (7_{def}) are independant of each other (because the morphological part does not depend on Ptr, as opposed to the sequence of marks M(‘Ptr’)), which seems to entail that:

$$V(Ptr) : V(Ptr) =_{def} V(\text{‘}Ptr\, is\, P\text{'}) \qquad (10_{def})$$

and hence, if P is functionally defined, that $V(Ptr) : V(Ptr) =_{def} g_{\mathcal{P}}(V(Ptr))$ and that $g_{\mathcal{P}}$ has a unique fixed point. However, in order to properly justify the transition from (7_{def}) to (10_{def}) we would have had to demonstrate a lemma relating the definition of a vector to that of its parameters when one of them is independent of the others.

This version of the Liar and the Truth-Teller thus shows that the latter is also paradoxical.[25] The latter is arguably as paradoxical as the former. While the Liar relies on the assumption that:

- *Existence* Our initial definition of Ptr (6_{def}) succeeds in picking at least one interpreted sentence.

The Truth-Teller paradox relies on the assumption that:

- *Uniqueness* Our initial definition of Ptr (6_{def}) succeeds in picking at most one interpreted sentence.

But *Uniqueness* is at least as intuitively plausible as *Existence*.

Similarly, the instance of (8_{def}) where 'P' is substituted with 'true' (T_{def}) is simpler and at least as plausible as the instance of (8_{def}) where 'P' is substituted with 'false' (F_{def}).

It should be emphasized, moreover, that we haven't relied anywhere on the supervenience of semantic facts over non semantic facts. As already stated (see p. 4 and fn. 4), Kremer (1988) has argued that this principle is threatened by the (semantics of the) Truth-Teller. If this is right, we could naturally derive a contradiction from semantic

[25] This are natural language versions of the Liar and the Truth-Teller in which the Liar and Truth-Teller sentences are defined in a stipulative and explicitly circular way. It is well-known that there are versions of the Liar and the Truth-Teller that do not rely on such definitions. Consider this demonstrative version:

$$DPtr : Dtr =_{def} \text{`this very sentence is } P\text{'}$$

Or this 'Quinean' version using a 'syntactic function' and meant to avoid explictly circular definitions (Quine 1953, pp. 133–134):

$$QPtr : QPtr =_{def} \text{``yields a sentence that is P when applied to itself' yields a sentence that is P'}$$

[See also Gaifman (2004) who neatly relates this version to Gödel's diagonalization lemma, and claims that it does not really avoid circularity.] These two versions can however be treated like Ptr. The reason why, roughly, is that even though they are not *explcitly* circular, the definitions of $DPtr$ and $QPtr$ are still implicit circular. Consider $DPtr$. Just like we derived (9_{def}), we can derive

$$\begin{bmatrix} M(DPtr) \\ V(DPtr) \end{bmatrix} : \begin{bmatrix} M(DPtr) \\ V(DPtr) \end{bmatrix} =_{def} \begin{bmatrix} M(\text{`this sentence is } P\text{'}) \\ g_{\mathcal{P}}(\text{`this sentence is } P\text{'}) \end{bmatrix} \tag{11_{def}}$$

Now, the semantic value of the demonstrative expression 'this very sentence' is arguably defined by the fact that it is the same as that of the sentence in which it occurs:

$$V(\text{`this very sentence is P'}) : V(\text{`this very sentence is P'}) =_{def} V(\text{`DPtr is P'})$$

From this we get that:

$$\begin{bmatrix} M(DPtr) \\ V(DPtr) \end{bmatrix} : \begin{bmatrix} M(DPtr) \\ V(DPtr) \end{bmatrix} =_{def} \begin{bmatrix} M(\text{`this sentence is } P\text{'}) \\ g_{\mathcal{P}}(\text{`}DPtr\text{ is } P\text{'}) \end{bmatrix} \tag{12_{def}}$$

Deriving a contradiction from that is exactlty the same as deriving a contradiction from (9_{def}) (just replace 'Ptr' by '$DPtr$' in the main text's derivation). The case $QPtr$ is not different, provided that we grant that V(“yields a sentence that is P when applied to itself' yields a sentence that is P') : V(“yields a sentence that is P when applied to itself' yields a sentence that is P') $=_{def}$ V('QPtr is a sentence that is P').

behaviour of the Truth-Teller provided that we assume the supervience of the semantics. The present argument, however, derived a contradiction from the behavior of the Truth-Teller *without* relying on the supervenience of the semantics. It only relied, instead, on the logic of definitions.

We have seen that while Richard's dual directly depends on the definition (4_{def}), Richard's paradox depends on the mere identity ($4_{=}$) that follows from it. Something exactly similar is happening here: while we need the definition (8_{def}) to argue that the Truth-Teller is paradoxical, the mere identity ($8_{=}$) would be sufficient to show that the Liar is paradoxical:

$$V(\text{“s' is P'}) = g_{\mathcal{P}}(V(\text{‘s'})) \tag{$8_{=}$}$$

Now someone who contends that the identity versions of the Truth and Falsity schemata:

$$V(\text{“s' is true'}) = V(\text{‘s'}) \tag{$T_{=}$}$$

$$V(\text{“s' is false'}) = 1 - V(\text{‘s'}) \tag{$F_{=}$}$$

are more plausible than their definitional variants (T_{def}) and (F_{def}) could *eo ipso* claim that ($8_{=}$) is more plausible than (8_{def}) and that the Truth-Teller is not accordingly as paradoxical as the Liar. I am ready to grant that there is room for quibbling here, and the equal plausibility of the definitional and the equality premises is not as clear as in the case of Richard's paradox and its dual. However, it should be emphasized that (T_{def}) and (F_{def}) are quite weak. In particular, accepting them does not entail accepting the view, already suggested by (Tarski 1943, pp. 342–4) and endorsed by some deflationists about truth, to the effect that truth is defined by the sum of the instances of the T-schema ($T_{=}$)—following Yaqub (1993) we can call this view (TS)[26]:

$$\text{true} : \begin{bmatrix} \vdots \\ V(\text{“}s_i'\text{is true'}) \\ \vdots \end{bmatrix} =_{def} \begin{bmatrix} \vdots \\ V(\text{‘}s_i'\text{)} \\ \vdots \end{bmatrix} \tag{TS}$$

(where (‘s_i’)$_i$ is substituted with a series of sentences ranging over all the sentences of our language, and the brackets represent a series.)

Assuming that definitions should explain our understanding of worlds or concepts, (TS) implies that our understanding of truth is explained by *the sum of all the instances* of the Truth-Schema. As noticed by Gupta (1993, pp. 365–6), this is implausible because one might have a good understanding of what truth is without mastering but a small fraction of the concepts involved in all the instances of the Truth-Schema. But (T_{def}) says much less. It does not, in particular try to define truth but only the semantic value of an arbitrary sentence asserting the truth of another one.

[26] See especially Yaqub (1993, pp. 27–36), where he distinguishes this thesis from neighbouring ones and attributes them to important historical figures.

It says, more specifically, that, if 's' is substituted with a sentence, the semantic value of the 's is true' is defined by *the particular instance* of the Truth-Schema for the sentence s:

$$V(\text{“s' is true'}) : V(\text{“s' is true'}) =_{def} (V(\text{‘s'}))$$

It thus only implies that our understanding of 'V("s' is true')' (where 's' is substituted with a sentence) is explained by *this particular instance* instance of the Truth-Schema. This is weaker and immune to the above objection against (TS). In fact, the schema (T_{def}) should be accepted by anyone who grants that part of what it is to understand 'V("s' is true')' is to assent to the identity, 'V("s' is true') = (V('s'))'. I surmise that any plausible theory of truth and meaning should grant that much, and accept (T_{def}).[27]

We should accordingly conclude that even if the argument for the paradoxical character of the Truth-Teller is a bit more complex and less obvious than that for the paradoxical character of the Liar, it relies on premises which are as plausible or (considering the discussion of (T_{def}) above) roughly as plausible, which makes it as paradoxical, or nearly as paradoxical as the latter.

The paradoxical nature of the Liar paradox's dual—the Truth-Teller—has some interesting implications.

First, as they do not consider the Truth-Teller paradoxical, philosophers and logicians do not, as we have already emphasized, usually consider that it needs to be solved. At best, they try to find out its semantic status and consider their job done. Some have argued that it is true (Smith 1984), others that it is neither true nor false (Read 2008b), and yet many others that it is simply false (Priest 2006a, pp. 64–66), (Yablo 1993, p. 387) and (Read 2008a, p. 213). Now, whatever the merits of their respective arguments, they do absolutely nothing to rebutt the absurd claim to the effect that g_{true} is the identity fuction and yet must have at most one fixed point. They cannot accordingly solve what we may now call the Truth-Teller paradox. Here again, we have a new, unsolved paradox.

Given the symmetry between the Liar and the Truth-Teller, it is natural to check whether the (purported) solutions to the former can generalize to the latter. A natural approach to solving the Liar paradox involves going trivalent, and claiming that the Liar sentence $Ftr =$'Ftr is false' does not have the semantic value 'true' (1) or false (0) but a third semantic value or status (call it $\frac{1}{2}$). It is impossible to do justice, here, to the wide variety of such trivalent approaches to the Liar.[28] Yet, we should emphasize

[27] Another classical objection against views such as (TS) that define truth in terms of the Truth-Schema [an objection which can be traced back at least to Lewy (1947)] claims that the former entail that instances ($T_{=}$) are necessary and analytic but (roughly) that they cannot be either because a given sentence 's' could mean something different from what it happens to mean and thus be true while it is not the case that s. This objection might be thought to generalize to instances of (T_{def}) as the latter entail that instances of ($T_{=}$) are indeed analytic and necessary. However, as noticed by many [see, for instance, the very clear exposition of Raatikainen (2003)] the classical argument to the effect that instances of ($T_{=}$) are not analytic or necessary rests on the assumption that the sentences under scrutiny are *uninterpreted* sentences. But the sentences we have considered so far are all interpreted sentences.

[28] The most developed contemporary versions of such a trivalent approach stem from the works of Kripke (1975) who showed that it could countenance a strong enough version of the Truth-Schema. Some have

that even though such approaches provide a promising solution to the present version of the Liar, they do not generalize in any obvious way to the corresponding Truth-Teller. Indeed, one can easily claim that even if "'s' is false' is true when 's' is false, and even if it is false when 's' is true, it has the semantic $\frac{1}{2}$ when 's' has it too, which ensures that g_{false} is not a permutation anymore and has exactly one fixed point, namely $\frac{1}{2}$. This would be consistent with claim (i) (to the effect that $g_{\mathcal{P}}$ has at least one fixed point), which would solve the present version of the Liar paradox. The trivalent approach seems however helpless when it comes to the Truth-Teller: even in a trivalent framework, g_{true} will have at least two fixed points (truth and falsity), which will still contradict claim (ii) (to the effect that it has at most one fixed point).[29] It seems that if one is to solve the Truth-Teller by modifying the set of truth-values, one should rather go 'monovalent', claiming that there is only one truth-value: the same for all sentences. This, however, means triviality, and it is a non-starter.

Another popular strategy for solving the Liar consists in claiming that the Liar sentence (or some of its tokens) fail to possess a semantic value and that our initial definition of the Liar sentence fails to define an interpreted sentence. This means denying *Existence* and accordingly, rejecting (i) (s.c. claiming that $g_{\mathcal{P}} = g_{\text{false}}$ can have no fixed point). This solution, which we can call, with one of its most influential contemporary partisans, 'cassationist' (Goldstein 2000) is akin to Richard's own solution to the eponymous paradox. Like in the latter case, however, it does not generalize to the Truth-Teller in any obvious way. Claiming that our initial definition of the Truth-Teller fails to pick an interpreted sentence allows one to reject (i) (s.c. claim that $g_{\mathcal{P}} = g_{\text{true}}$ can have 0 fixed point), *not* to reject (ii) (s.c. claim that $g_{\mathcal{P}} = g_{\text{true}}$ can have two or more fixed points). Yet in order to hinder the paradoxical reasoning associated with the Truth-Teller, we need to reject (ii). In order to do so, one should rather deny *Uniqueness*, claiming that our initial definition of the Truth-Teller sentence picks more than one interpreted sentence and is associated with more than one semantic value[30]—a

Footnote 28 continued
however voiced dissatisfaction with Kripke's framework, most notably with its the expressive power, its difficulties to meet so-called revenge problems or with the weakness of its underlying logic. In response, many sophistications of Kripke's approach have been put forward [see, among many others Burge (1979), Priest (2006a), McGee (1991), Soames (1999), Horsten (2011), Ripley (2013) and Cobreros et al. (2013)]. Although it can be considered as a trivalent approach, in which the valuation function attributes $\frac{1}{2}$ to sentences that are both true and false, Priest (2006a)'s dialetheist solution is more naturally considered as a 'plurivalent' approach, in which the valuation *function* is replaced by a non-functional *relation* that can relate a sentence to true (1), false (0) or to both truth-values at the same time [see Priest (2006a, p. 288, 2014)].

29 This criticism applies to Kripke (1975)'s approach. According to the latter, equally acceptable interpretations of the truth predicate classify the Liar as true, false and neither true nor false. Now, either we pick one interpretation as *the* correct interpretation or we insist that there no such thing as *the* correct interpretation [the latter is Kripke's preferred option, see Kremer (2009, fn. 8)]. In both cases however, g_{true} will have at least two semantic statuses as fixed points—true, false—whereas it should have only one.

One must watch out, by the way, not to confuse the fixed points of Kripke's construction with the ones I have been referring to throughout this article. They are all fixed points all right, but of very different functions, and they have little in common.

30 Notice that this is not the same as claiming that our definition of the Truth-Teller picks just one interpreted sentence, with *one complex semantic value*, say a couple of truth values or a function from contexts to truth-values. The latter claim will not, it seems, help us solve the Truth-Teller because it just consists in expanding

position we might call 'pluralism'[31] and that contrasts with cassationism. Interestingly then, even if it helped solving the Liar paradox, cassationism would not help solving the present Truth-Teller paradox. The partisan of cassationist solutions to the Liar should, it seems, endorse a pluralist solution to the Truth-Teller.

There is a final solution to the Liar that should be considered here. Partly in order to treat the Liar paradox and to provide a proper semantic characterization of the Truth-Teller, Gupta and Belnap (1993) have developped a theory of circular definitions and of truth (considered as a cicularly defined concept) in terms of revision rules. Their idea, very roughly, is that circular definitions provide us not generally with an extension for the defined term, but with a way to revise the extension we can attribute to it on different hypotheses. From a given initial hypothesis, iterated application of the rule form *revision sequences* which models the way we can improve an hypothesis concerning the extension of a circularly defined term. A full presentation and discussion of their theory would far outreach the scope of this essay. What needs to be said, however, is this. First, the revision theory allows one to defuse most versions of the Liar paradox as well as other paradoxes that seem to be grounded on viciously circular definitions.[32] Second, Gupta and Belnap (1993) do not claim that some duals are as paradoxical as the classical paradoxes of which they are the duals, and they do not put forward arguments similar to those presented here, that simultaneously show that the classical paradoxes and their duals lead to a contradiction. Interestingly, however, the revision theory of definitions can block our argument and solve, by the same token, both paradoxes and their paradoxical duals. For it weakens a rule that have been crucial so far, namely, Definiendum-Elimination. Simplfying somewhat, Gupta and Belnap (1993, pp. IV–§3,V) associate an index with each step in a derivation and urge that, generally, we can replace an occurrence of a definiendum at i by its definiens, but only if we assign it the index $i - 1$. With this weakened version, we cannot, for example, derive (7_{def}) (to the effect that Ptr is defined by its identity with $f_{\mathcal{P}}(Ptr)$) and conclude, paradoxically, that $g_{\mathcal{P}}$ (and $f_{\mathcal{P}}$) must have a unique fixed point. We can only conclude, rather, that Ptr is defined by the fact that *at a given step* i, it is identical with $f_{\mathcal{P}}(Ptr)$ *at step* $i - 1$, which does not imply that $g_{\mathcal{P}}$ (and $f_{\mathcal{P}}$) must have at least nor at most one fixed point. The very same could be said about (4_{def}) and our derivation of a contradiction in the case of Richard's paradox and of its paradoxical dual.

Footnote 30 continued
the set of semantic values. It will thus run into a problem which is similar to the one posed by the trivalent approach. Indeed, when 'P' is substituted with truth, $g_{\mathcal{P}} = g_{\text{true}}$ will still arguably have more than one fixed point in this extended set of semantic values, whereas it should have at most one. If, for example, the semantic values are functions from contexts to truth-values, it seems that we will be bound to acknowledge that if a sentence q is true (res. false) in all contexts, so is the sentence asserting that q is true. Yet, this implies that all true acontextual sentences are a fixed point of $g_{\mathcal{P}} = g_{\text{true}}$, a function which has accordingly more than one fixed-point.

31 The word 'pluralism' has been used by Read (2008b) to designate a solution to the Liar that is quite different from what I call 'pluralism' here. What Priest (2014) calls a 'plurivalent' semantic (that is the claim that sentences can be associated with no, or with many semantic-values) is, on the other hand, a view that fits both cassationist solutions to the Liar and what we have called '*pluralist*' solutions to the Truth-Teller.

32 See Gupta (1997, 439ff). for some caveats concerning so-called revenge problems and their significance.

8 Taking stock and generalizing

Let us take stock. Both in the case if Richard's paradox, the version of the Liar presented here and in the case of their duals, we seem to be defining an object (be it a number or a sentence) that cannot be defined. More precisely, we have three conflicting claims that seem equally true:

1. *Successful Definition* The definition D of an object X in a domain U is successful
   (i) in picking out a least one object in U (Existence).
   (ii) and at most one object in U (Uniqueness).
2. *Implict Circularity* D is implicitly circular, yielding a definition of the form X: $X =_{def} f(X)$ (where f is a function from U to U) through the elimination of non-primitive terms in its definiens. Accordingly,
   (i) D succeeds in picking at least one object iff f has at least one fixed point in U.
   (ii) D succeeds in picking at most one object iff f has at most one fixed point in U.
3. *No Unique Fixed Point* f, however,
   (i) has either no fixed-point.
   (ii) or more than one fixed-point.

In the case of the Richard's paradox and dual the object X is a real number, in the case of the Liar and its dual, it is an interpreted sentence.

The classical paradoxes hinge on the confict between (1i), (2i) and (3i), the paradoxical hypodoxes on that between (1ii), (2ii) and (3ii). The role of *Successful Definition* (1) should be emphasized. What distinguishes these paradoxes from mere 'bad definitions', like the circular definitions of real numbers mentioned in Sect. 2 ($b : b =_{def} b + b^2 + 1$ which picks no real number and $c : c =_{def} c$ which picks too many), is that while we have no good reason to believe that these definitions succeed, we have good reason to believe that the definitions involved in the classical paradoxes and their hypodoxes are indeed successful. In the case of the Liar paradox, this reason is the usual claim that Ptr = 'Ptr is P' is a well-formed grammatical expression that we seem to understand very well and that should accordingly pick a unique interpreted sentence (that is, the couple of a sequence of marks and a semantic value). This is indeed the reason classically invoked against cassationism [see. e.g. Read (2008a)]. In the case of Richard's paradox, the reason is, likewise, that the definite expression defining α is well-formed, clearly understandable, and that it clearly seems to pick a real number between 0 and 1.

What should we answer, then, to someone who wants to discount the paradoxical character of the paradoxical hypodoxes put forward here by invoking some constraint $\mathcal{C}$ that his favourite account of good definitions prescribes but that our definitions of α or Ptr does not satisfy?—being non-circular is one example of such a constraint, but the objector could invoke others, such as being in so-called "normal form", see Gupta (2015). To such an objector we should first remind that our definitions are of the same form as that through which the classical Richard's paradox and a quite common version of the Liar using quotations (L = 'L is false') are introduced, and that if the former infringe the constraint $\mathcal{C}$, the latter do it as well. This means that our hypodoxes are not less paradoxical than the above mentioned classical paradoxes for infringing this constraint. We might also argue, in some cases at least, that the constraint $\mathcal{C}$ is too strong, forbidding some definitions that are integral to our practice of defining terms, say in mathematics (this seems to be the case, as we have seen in Sect. 2, for the circularity constraint). But overall, we should insist that by (1), we have a good argument to the effect that, whether they satisfy the invoked constraint $\mathcal{C}$ or not, our definitions of α and Ptr should, *be successful enough to pick a single object*, which contradicts (2–3). This strongly suggests that the paradoxical hypodoxes we have put forward are not rooted in a failure to meet $\mathcal{C}$, but rather on their *simultaneous failure and success* to meet the constraint of picking exactly one object (as well as the related constraint of non-creativity).[33]

An interesting question, at this stage, concerns the scope of this schema: does it apply to all the classical paradoxes of self-reference or to only one portion of them? My conjecture is that it applies to all, or almost all of them. Unfortunately, articulating and motivating this conjecture precisely (let alone showing that it is true) would far outreach the scope of this paper.[34]

9 Conclusion

I have said in the introduction that whereas the classical paradoxes hinge on an object that seems overdetermined (there is a proof that it has two inconsistent properties), duals hinge on an object that seems underdtermined because, for some property, it

[33] A referee for this journal has put forward the following definition of a natural number:

$$a : (x = a) \quad =_{def} (x = 1) \vee (x = 2)$$

which picks two objects (1 and 2) for a and accordingly allows to derive the absurd claim that $1 = 2$. He or she has also suggested that (a) there is some $\mathcal{D}$ constraint that this definition does not meet (b) that our definitions of α and Ptr do not meet $\mathcal{D}$ either, and (c) that this allows one to discount the paradoxical character of our hypodoxes. Finally, he or she also has added that $\mathcal{D}$ might be a constraint of homogeneity, which imposes the defined term to be of the same logical category as the definiendum (and the definiens).

However, heterogeneous definitions seem ubiquitous and generally unproblematic, and they are sometimes indispensable [see for example Gupta (2015) on implicit definitions]. Morover, we only use heterogeneous for clarificatory purposes, on p. 14, rehearsing an argument that originally relied on homogeneous definitions with heterogeneous, sentential, definitions of numbers instead. More deeply, as we have seen above, if there is a constraint on definitions whose infringement is directly responsible for the paradoxical hypodoxes it is the the constraint of picking exactly one object. But if $\mathcal{D}$ was taken to be the constraint of picking exactly one object (c) would not hold (ie. one could not discount the paradox) because, again and as stated by (1), we have an argument to the effect that this constraint *is* in fact satisfied, and this precisely is where the paradox comes from.

[34] The only obvious exception I see is what is sometimes called ‘Routley’s paradox’ [see Routley (1980)]. The following definition of Ro leads to a paradox, at least if we assume a very liberal form of the Naïve Comprehension Schema which guarantees that this definition picks at least one set.

$$Ro : Ro =_{def} \{x : x \notin Ro\}$$

The related hypodox, which hinges on $DRo : DRo =_{def} \{x : x \in DRo\}$ will not be paradoxical however, because neither Extensionality, nor any plausible axiom of Naïve Set Theory give us reason to believe that this definition should pick at most one set.

seems impossible to decide whether it has it or not. This is the common take on the difference between paradoxes and their duals. If the present paper is right, however, this is not quite the full story. The problem with the "dual object" is not so much that there seems to be no proof that it has or has not this property. It is not even that there seems to be a proof that there is no such proof. It is, at least in the cases of Richard's dual and of the Truth-Teller, that the dual object's definition seems to pick, as it should, at most one object (and there is a proof that it does), but cannot pick at most one object (and there is a proof that it cannot). More broadly, the problem with the dual object, just like that with the paradoxical object, is that we seem to be able to define it even though it cannot be defined. The difference between the two is that while the paradoxical object cannot be defined because its purported definition picks too few objects, the dual object cannot be defined because its purported definition picks too many objects.

The duals or hypodoxes of Richard's paradox and of the Liar are accordingly paradoxical too. Against common wisdom, duals or hypodoxes can thus be paradoxical. Classical paradoxes involve a paradoxical reasoning leading to a contradiction from acceptable premises. Hypodoxes likewise involve a paradoxical reasoning leading to a contradiction from equally acceptable premises. They should be treated, just like the classical paradoxes, by showing that there is something wrong with such a paradoxical reasoning. This means that their common treatment, which just consists in puting forward a characterization of the dual object, is insufficient: duals are still untreated paradoxes. Moreover, as many solutions to the classical paradoxes do not generalize in any obvious way to their duals, the treatment of paradoxical duals calls for new solutions (like the pluralism) or favours some old ones (like the revision-theoretic one) against others.[35]

References

Armour-Garb, B., & Woodbridge, J. A. (2012). Liars, truthtellers and naysayers: A broader view of semantic pathology i. *Language and Communication*, *32*(4), 293–311.
Austin, J. L. (1950). Truth. *Aristotelian Society Supplements*, *24*(1), 111–29.
Beall, J. (2009). *Spandrels of truth*. Oxford: Oxford University Press.
Belnap, N. (1993). On rigorous definitions. *Philosophical Studies*, *72*(2), 115–146.
Billon, A. (2011). My own truth, relative truth and the semantic pathologies of self-reference. In S. Rahman & G. Primiero (Eds.), *Antirealism: The realism/antirealism debate in the age of alternative logics*,., Logic, epistemology, and the unity of science Dordrecht: Springer.
Billon, A. (2013). The truth-tellers paradox. *Logique Et Analyse*, *224*, 371–389.
Burge, T. (1979). Semantical paradox. *Journal of Philosophy*, *76*, 169–198.
Cobreros, P., Égré, P., Ripley, D., & Van Rooij, R. (2013). Reaching transparent truth. *Mind*, *122*(488), 841–866.

[35] I have already put forward arguments for the paradoxical character of the Truth-Teller elsewhere (Billon 2011, 2013). The present argument supersedes these. I am grateful to Peter Eldridge-Smith, Graham Priest, Lucas Rosenblatt and four referees of this journal for their useful comments on earlier drafts of this paper. Primitive versions of this paper have been presented at the Logic CNRAP Networkshop in Princeton in 2013, at the IEC/MCMP Paris-Munich Networkshop in Munich in 2013 and at the University of Lille in 2016. I have benefited from founds of the program New Ideas in Mathematical Philosophy from the ENS "département d'études cognitives".

Eldridge-Smith, P. (2007). Paradoxes and hypodoxes of time travel. In J. L. Jones, P. Campbell, & P. Wylie (Eds.), *Art and time* (pp. 172–189). Melbourne: Australian Scholarly Publishing.
Eldridge-Smith, P. (2008). *The liar paradox and its relatives*. Melbourne: The Australian National University.
Eldridge-Smith, P. (2012). A hypodox! a hypodox! a disingeneous hypodox!. *The Reasoner*, *6*(7), 8–10.
Eldridge-Smith, P. (2015). Two paradoxes of satisfaction. *Mind*, *124*(493), 85–119.
Feferman, S. (2005). Predicativity. In S. Shapiro (Ed.), *The Oxford handbook of philosophy of mathematics and logic* (pp. 590–624).
Gaifman, H. (2004). The easy way to gödel's proof and related matters. http://www.columbia.edu/~hg17/Diagonal-Cantor-Goedel-05.pdf.
Gaifman, H. (2006). Naming and diagonalization, from cantor to gödel to kleene. *Logic Journal of IGPL*, *14*(5), 709–728.
Goldstein, L. (2000). A unified solution to some paradoxes. In *Proceedings of the Aristotelian Society, New Series*. The Aristotelian Society, Blackwell Publishing.
Goldstein, L. (2006). Fibonacci, yablo, and the cassationist approach to paradox. *Mind*, *115*(460), 867–890.
Gupta, A. (1993). Minimalism. *Philosophical Perspectives*, *7*, 359–369.
Gupta, A. (1997). Definition and revision: A response to McGee and Martin. *Philosophical Issues*, *8*, 419–443.
Gupta, A. (2015). Definitions. In Zalta, E. N. (Eds.), *The stanford encyclopedia of philosophy*, Summer 2015 edn.
Gupta, A., & Belnap, N. (1993). *The revision theory of truth*. Cambridge: MIT Press.
Halbach, V., & Visser, A. (2014). The henkin sentence. In E. Alonso., M. Manzano & I. Sain (Eds.), *The life and work of Leon Henkin* (pp. 249–263). Springer.
Herzberger, H. G. (1970). Paradoxes of grounding in semantics. *The Journal of philosophy*, *67*(6), 145–167.
Horsten, L. (2011). *The Tarskian turn: Deflationism and axiomatic truth*. Cambridge: MIT Press.
Kremer, M. (1988). Kripke and the logic of truth. *Journal of Philosophical Logic*, *17*(3), 225–278.
Kremer, P. (2009). The revision theory of truth. In Zalta, E. N. (Eds.), *The stanford encyclopedia of philosophy*, Spring 2009 edn.
Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, *72*(19), 690–716.
Lewy, C. (1947). Truth and significance. *Analysis*, *8*(2), 24–27.
Mackie, J. L. (1973). *Truth, probability and paradox: Studies in philosophical logic*. Oxford: Oxford University Press.
McGee, V. (1991). *Truth*. Hackett: Vagueness and Paradox.
Poincaré, H. (1906). Les mathématiques et la logique. *Revue de métaphysique et de Morale*, *14*(3), 294–317.
Priest, G. (1994). The structure of the paradoxes of self-reference. *Mind*, *103*(409), 25–34.
Priest, G. (2006a). *In contradiction* (2nd ed.). New York: Oxford University Press.
Priest, G. (2006b). The paradoxes of denotation. In T. Bolander, V. F. Hendricks, & S. A. Pedersen (Eds.), *Self-reference* (pp. 137–150). Stanford: CSLI Publications.
Priest, G. (2014). Plurivalent logics. *The Australasian Journal of Logic*, *11*(1). https://ojs.victoria.ac.nz/ajl/article/view/1830.
Priest, G., & Mortensen, C. (1981). The truth-teller paradox. *Logique et Analyse*, *24*, 381–388.
Quine, W. V. O. (1953). Notes on the theory of reference. In *From a logical point of view* (pp. 130–138). Harper & Row.
Raatikainen, P. (2003). More on Putnam and Tarski. *Synthese*, *135*(1), 37–47.
Read, S. (2008a). Further thoughts on Tarski's T-scheme and the liar. In S. Rahman, T. Tulenheimo, & E. Genot (Eds.), *Unity, truth and the liar* (Vol. 8, pp. 205–225)., Logic, epistemology, and the unity of science Dordrecht: Springer.
Read, S. (2008b). The truth schema and the liar. In S. Rahman, T. Tulenheimo, & E. Genot (Eds.), *Unity, truth and the liar, logic, epistemology, and the unity of science* (Vol. 8, pp. 3–18). Dordrecht: Springer.
Richard, J. (1905). Les principes des mathématiques et le problème des ensembles. *Revue générale des Sciences Pures et Appliquées*, *16*(541), 142–144.
Ripley, D. (2013). Paradoxes and failures of cut. *Australasian Journal of Philosophy*, *91*(1), 139–164.
Routley, R. (1980). *Exploring meinong's jungle and beyond*. Camberra: Australian National University.
Russell, B. (1906). Les paradoxes de la logique. *Revue de métaphysique et de morale*, *14*(5), 627–650.
Sainsbury, R. M. (1995). *Paradoxes*. Cambridge: Cambridge University Press.
Scharp, K. (2013). *Replacing truth*. Oxford: OUP Oxford.

Simmons, K. (1993). *Universality and the liar: an essay on truth and the diagonal argument*. Cambridge: Cambridge University Press.
Smith, J. W. (1984). A simple solution to Mortensen and Priest's truth teller paradox. *Logique et Analyse*, *27*, 217–220.
Soames, S. (1999). *Understanding truth*. Oxford: Oxford University Press.
Sorensen, R. (2001). *Vagueness and contradiction*. Oxford: Oxford University Press.
Suppes, P. (1957). *Introduction to logic*. New York: Dover Publications.
Tarski, A. (1943). The semantic conception of truth: And the foundations of semantics. *Philosophy and Phenomenological Research*, *4*(3), 341–376.
Wagner, P. (2017). Definition. In Kristanek, M. (Eds.), *L'encyclopédie Philosophique*.
Woodbridge, J. A. (2004). A neglected dimension of semantic pathology. In *Logica yearbook* (pp. 277–292). Prague. Filosofia: Institute of Philosophy, Academy of Sciences of the Czech Republic.
Woodbridge, J., & Armour-Garb, B. (2005). Semantic pathology and the open pair. *Philosophy and Phenomenological Research*, *71*(3), 695–703.
Yablo, S. (1993). Hop, skip, and jump: The agnostic conception of truth (Corrections in Philosophical Perspectives, 9, 509–506). *Philosophical Perspectives*, *7*, 371–393.
Yaqub, A. M. (1993). *The liar speaks the truth: A defense of the revision theory of truth*. Oxford: Oxford University Press.