Deflationary truth is a logical notion

Denis Bonnay, Henri Galinon

Deflationary truth is a logical notion
Penultimate version
- do not cite -

*

D. Bonnay & H. Galinon †

Abstract

The thesis that truth is a logical notion has been stated repeatedly by deflationists in philosophical discussions on the nature of truth. But what is a logical notion? Building on Tarski (1986) several authors have discussed and developed the idea that the characteristic generality of logic can be turned into a criterion of logicality based on the invariance of the symbols' interpretations under large classes of transformations on the domain. In this paper, we show how the deflationist can use invariance criteria in support of her claim that deflationary truth is a logical notion.

*Published version in Pulcini & Piazza (eds.)*Philosophy of mathematics: truth, existence and explanation*, Springer Verlag

†Many thanks to Julien Boyer for helpful discussion and criticisms.

1 Deflationism and the logicality thesis

The claim that the deflationist's notion of truth is a logical notion is found already in Quine (1971) and later in the work of contemporary deflationists such as Horwich (1998), or Field (1986).[1] These authors have typically meant the claim to be suggestive rather than a key formulation of their philosophical stance, and they did not provide arguments for it. Of course, some quick remarks naturally come to mind. For instance that the truth predicate[2] enjoys a form of topic-neutrality that is typical of logical expressions ; or that deflationary truth works more or less as a device for expressing infinite conjunctions; or that it can be systematically eliminated in context with the help of substitutional quantification. These remarks are typically unconclusive. As an informal take on logicality, topic-neutrality is not precise enough (how shall we know that something is topic neutral? are mathematical notions topic-neutral ? etc.), and the claims that infinite conjunction and susbtitutional quantification are logical are by no means self-evident. On the other hand, the logicality thesis is certainly an interesting thesis for the deflationist to build upon, as it articulates the idea of an expressive but non explanatory property in a way compatible with various logical facts that philosophers and logicians have drawn our attention to.[3] We would make some progress in the discussion of deflationism if we could deepen our understanding of the relationships between the notion of deflationary truth and logical notions.

A favorable circumstance is that, despite *prima facie* elusiveness, the boundary between logical and non logical notions is a well-trodden philosophical subject. Among various frameworks developed in the literature in order to test expressions for logicality, the major contender is possibly the 'invariance' framework, which makes it a reasonable option to inquire about the logicality of truth. Since some philosophers have doubted that the deflationist's thesis about the logical nature of truth could survive a test based on a precise criterion of logicality such as invariance,[4] we hope that the results reported in this essay will be of interest to the sceptics as well as to professed deflationists. Indeed what we show in the following is that notion of truth, as the deflationist understands it, *is* a logical notion.

To achieve our goal, we will proceed as follows. The remainder of this sec-

[1] Among others. See Horwich (1998) pp. 2–5,Field (1986), p.76. McGinn (2000) takes truth as one of the logical properties. See e.g. Wyatt (2015) f.n. 28 for further references.

[2] We allow ourselves to move freely from talk about predicates and expressions to talk about notions. We do not think that much turns on this: logical expressions express logical notions and conversely if a notion is logical then an expression that expresses it is a logical expression.

[3] In particular Shapiro (1999) and Ketland (1998). See also the essays in Achourioti & al. (2011). On the philosophical significance of the thesis of logicality of truth in this context, we may refer the reader to Galinon (2015).

[4] See e.g. Wyatt 2016, pp.15-17.

tion will set the stage: we first introduce the basics of the invariance approach to logicality and then decide on what it would mean for truth – "truth as understood by deflationist", as we said – to be logical in that setting. The main uptake is that truth is logical in so far as adding a truth predicate adds to the expressive power of a language only when the underlying logic is not as expressive as the bounds of logic allow. This is a thesis about deflationary truth, because the focus is on the expressive power of a language equipped with the truth predicate, and the approach is invariance based because the bounds of logic are to be understood as given by some invariance criterion. The second section defines what it means for a logic (not) to be as expressive as invariance allows, by putting forward the concept of a *generated logic*. Some sets of logical expressions are, as we shall argue, somehow "incomplete" and the notion of a generated logic is the notion of a logic that do not suffer this special kind of expressive defectiveness. In the third section, we develop a small formal apparatus that allows us to get a grip on the expressive power of the deflationist's notion of truth, in particular regarding the increase in expressiveness that results of adding a truth predicate. This leads to defining the concept of a *truth-complete logic*. In a nutshell, a truth-complete logic is a logic in which what can be expressed by means a deflationary truth predicate can be expressed without it – more precisely, whatever class of structures can be defined using a deflationary truth predicate should be definable without it. The crucial step in Section 4, consists in showing that generated logics are exactly the truth-complete ones. We argue that the truth-completeness of generated logics supports the claim that deflationary truth is a logical notion. In the last section, we briefly compare our methodology to the proof-theoretic one developed in Shapiro (1999) and Ketland (1998).

1.1 Logicality as invariance

What does it mean to say that an expression is logical ? In the sense we are interested in, what makes an expression logical is its denotation - that is, a semantical feature of a linguistic expression, not a feature of the way it is used, or of the specific discursive purposes it serves, though there is of course no denying that the denotation of an expression, the ways that are appropriate to use it and its role in discursive endeavor are somehow tied together. With this preamble in mind, the philosophical starting point to the invariance approach to logicality is a shared intuition that logic is a maximally general science, a science not about any specific domain of the world in particular. Then given this basic idea, one tries to characterize the logical expressions as those whose semantic value comes out *invariant* under an appropriate class of transformations. As Tarski (1986) puts it, this idea is a natural extension of Klein's Erlangen program. Klein had showed that different geometries can be respectively characterized as the study of those notions that

are invariant under such and such classes of transformations - Euclidian geometry by isometric transformations, topology by continuous transformations etc. In the same spirit, one may try to capture the generality of logic by thinking of it as the science of those notions that are invariant under the most general or *biggest* class of transformations. And indeed Tarski has argued along these lines that logical notions were those notions invariant under all permutations (over the fixed universal domain).

To get an intuitive grasp of what is going on under this invariance criterion, we shall illustrate how it works on two examples, the existential quantifier and an empirical property, redness. Start with the existential quantifier. What is its extension, that we want to test for invariance ? In keeping with the literature on logicality, quantifiers are handled in a Fregean way: They are to be thought as second-order predicates, which, in the case of a unary quantifier, are interpreted by a class of structures of the form $\langle M, P\rangle$, $P \subseteq M$. Writing $Q_\exists$ for the interpretation of the existential quantifier we thus get: $Q_\exists$ = the class of structures $\langle M, P\rangle$ with $P \neq \emptyset$. Now Q being a possible interpretation for a unary quantifier (ie a class of structures of the form $\langle M, P\rangle$), Q is *permutation invariant* if and only if, for all M and all permutations π on M, for all $P \subseteq M$,

$$\langle M, P\rangle \in Q \text{ if and only if } \langle M, \pi(P)\rangle \in Q$$

Consider a universe with domain $M = \{a, b, c\}$, and assume that two objects in the universe are red, say $\text{Red}_M = \{a, b\}$. There are 6 permutations on M:

M	π_1	π_2	π_3	π_4	π_5	π_6
a	a	a	b	b	c	c
b	b	c	a	c	b	a
c	c	b	c	a	a	b

Compare, under Tarski's invariance criterion, the behavior of the intuitively logical notion expressed by the existential quantifier and the behavior of the notion expressed by "Red" . Since $\text{Red}_M(a) \neq \text{Red}_M(\pi_4(a))$, redness in not invariant under permutation. On the other hand $\exists_M = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$ and $\exists_M(\{a, b\}) = \exists_M(\pi_4(\{a, b\}))$. And similarly it is easily seen that, unlike redness, $\exists_M$ is invariant under all permutations. As another way to put it, we may say that "Red" allows one to make some distinctions between structures ($\langle M, Red\rangle$ and $\langle M, (\pi_4(Red)\rangle$ for instance) that a purely logical notion like the existential quantifier cannot distinguish.[5]

[5]One might worry that *Red* and $\exists_M$ are not of the same type, but nothing important hinges on this. We invite the reader to check that the first-order relation $=_M$ (the identity on M) comes out as logical under the proposed criterion, and that the second-order predicate $tobeacoulor$ does not.

The generality of Tarski's specific approach, however, is questionable: paying attention to *transformations on a fixed domain* only may be seen as an unacceptable restriction of the generality intuition - why don't we consider more general class of transformations that allow shrinking and expanding the domain for instance ? And indeed one of the outcomes of Tarski's approach was an expectable overgeneration problem.[6] To generalize Tarski's fixed-domain approach and escape its framing effects, one is led to consider all possible structures of interpretation of the language and seek out to define an appropriate relation of "logical similarity" between them: the logical notions then come out as the notions that are invariant under the chosen similarity relation. The question "what are the logical notions" then amounts to asking what is the "good" similarity relation to consider.

Various answers have been proposed to this question. Taking the existence of an isomorphism as the similarity relation on structures, we get Tarski-Sher's criterion,[7] that is to say Tarski's criterion extended so that isomorphic structures not sharing their domain may now count as "similar".[8] However, entirely new possibilities open up in this setting. Feferman (1999), for instance, has considered the possibility of taking the existence of a surjective homomorphism between two structures for the first one to be similar to the other; and Bonnay (2008) has argued at length that potential isomorphism should be considered a better candidate, achieving a good balance between the absence of mathematical and empirical content on the one hand, and triviality on the other hand.[9]

As regards the constraints that are appropriate to define a relation of 'similarity' between structures which would be universally regarded as purely *logical* similarity, various sets of conceptually motivated constraints are possible, each package giving rise to different proposals such as those just mentioned. It is fair to say that there is still place for disagreements on the precise implementation of this invariance idea. Fortunately, and this an important point to make in regard to our project here, the fine-tuning of the invariance criteria does not matter for our purpose. What we will show below is that it is sufficient that one natural constraint, "closure under definability", be in the package for the corresponding logic to make the addition of an interpreted truth predicate idle. Thus at the end of the day, the validity of our conclusion regarding the logicality of deflationary truth will certainly depend (among other things) upon the validity of the invari-

[6]Intuitively: the smaller the transformation class, the bigger the resulting class of logical operators ; with Tarski's criterion, the scope of logic is identified with the scope of mathematics.

[7]Sher (1991).

[8]This in line with the idea that logical notion should not be sensitive to the identity of objects

[9]Note that just considering the biggest similarity relation over structures (the universal relation) would lead nowhere and that consequently the class of candidates for a good "similarity relation" , that is a good notion of what it is for two structures to be logically indistinguishable, has to be somehow constrained.

ance approach to logicality, but it will *not* depend on the validity of the choice of any specific 'logical similarity' relation between structures.

1.2 How can truth be logical ?

How then should we handle truth in the invariance framework? Obviously, we would run into problems if we were to apply *directly* an invariance criterion of logicality to the truth predicate. Logical notions are assumed to be maximally general in the sense that they do not distinguish among objects. But truth applies only to a very special kind of objects (sentences, say), and it does make differences among them.

However, we think that it would be superficial to take this point as a refutation of our project right from the start. The reason is that the truth predicate, as understood by the deflationist, is not really meant as a way to talk about sentences and the language – it is rather employed to cancel talk about language. Thus, when it is stated that all theorems of arithmetics are true, the statement is not really about sentences with such and such properties, but is really about numbers that are the subjects of the theorems themselves. Recall Quine's deflationist insight that by attributing truth to the sentence "snow is white" we basically just attribute whiteness to snow. The point of the truth predicate, Quine insists, is that it allows one to talk about the world even when, for some technical reasons (epistemic reasons), one has been forced to semantic ascent :

> Where the truth predicate has its utility is in those places where, though still concerned with reality, we are impelled by certain technical complications to mention sentences. Here the truth predicate serves, as it were, to point through the sentence to the reality; it serves as a reminder that though sentences are mentioned, reality is still the whole point. (Quine 1970, p.11)

Thus to vindicate the logicality of the deflationist's notion of truth one should analyze the behavior of the truth predicate in combination with devices expressing generality and naming sets of sentences. The question is what happens when they are used together as tools to talk about the (possibly non-linguistic) world, and one should thus abstract away from the possible role of the truth predicate as a tool to talk about sentences. From that perspective, it becomes clearer that applying invariance criteria to the denotation of the truth predicate would be at odd with the deflationist's semantic stance.

We then have to adopt an *indirect* strategy. The question we wish to ask is whether what the truth predicate enables us to say about the world is nothing more than what logic contributes to what we say about the world. If this is so,

deflationary truth will rightly count as logical: its specific expressive power remains within the bounds of logic. If this is not so, deflationary truth should not be regarded as logical: we may say with a truth predicate things about the world which do not possess the hallmark of logical generality. In order get a formal grip on this question, we shall proceed in the following way. First, we will assume that a particular invariance criterion for logicality is given, in the form of a relation of logical similarity between structures which logical notions should be invariant under (eg, Tarski's criterion corresponds to picking up invariance under automorphisms). Then consider the set of logical constants that are invariant for the given similarity relation – maybe these are the logical constants of an extension of first-order logic.[10] Then again, add to the logical constants the expressive possibilities made available by the use of a deflationary truth predicate and its companion devices. Finally, ask whether the classes of structures that are now definable with these extended alethic means were already definable in the logic (this is our property of truth-completeness). Again, if the answer is yes, it follows that deflationary truth talk just is logical talk by the light of invariance. We shall show that the answer is indeed positive, not just for one specific choice of invariance criterion, but for any such criterion which enjoys a natural property of closure under definability (quantifiers, as class of structures, that can be defined by sentences built from invariant quantifiers are still invariant). This will show up as our main result to the effect that a logic is truth-complete if and only if it is exactly generated, because a logic is exactly generated when it may be viewed as the logic one gets on account of an invariance condition closed under definability.

2 From invariance to logic and back

In Section 1.1, we introduced invariance under permutation as a logicality criterion. The present Section generalizes invariance under permutation to invariance under an arbitrary similarity relation, to allow for stronger invariance conditions such as those favored by Feferman or Bonnay. We then make implicit the connection between invariance criteria and first-order logics with generalized quantifiers. But not every similarity relation yields such a logic: sometimes, adding invariant quantifiers may allow one to define on a purely logical basis operations which are not invariant. This does not seem right if invariance is meant to capture what it means for an operation to be logical, so we accordingly restrict our attention to invariance criteria for which this does not happen (Principle of closure of definability).

[10]Remember, quantifiers are interpreted as class of structures, and conversely any class of structures can be seen as the interpretation of a putative quantifier.

2.1 Similarity relations, invariance and logics

A **similarity relation** S is a relation between structures respecting signatures (*i.e.* S is a family of relations S_σ between σ-structures for all signatures σ), the notation is $\mathscr{M}\ S\ \mathscr{M}'$. S is meant to capture what it means for two structures to be indistinguishable from a logical point of view (eg: being identical up to an automorphism).

We are then interested in the **invariance** of operators under similarity relations. An operator is just a class of structures. Recall the example in section 1: the operator associated with $\exists$ is the class $Q_\exists$ of structures $\langle M,P\rangle$ such that P is not empty. We write $Q_\exists(\langle M,P\rangle)$ for $\langle M,P\rangle \in Q_\exists$. The connection with the standard satisfaction clause for $\exists x$ is the following (where $\mathscr{M}$ is an arbitrary model and τ an arbitrary assignment over $\mathscr{M}$):

$\mathscr{M} \vDash \exists x\ \phi(x)\ \tau$

iff

there is an $a \in M$ such that $\mathscr{M} \vDash\ \phi(x)\ \tau[x := a]$

iff

$Q_\exists(\langle M, \{a \in M\ /\ \mathscr{M} \vDash\ \phi(x)\ \tau[x := a]\}\rangle)$

We say that an operator Q is S**-invariant** iff, for any structures $\mathscr{M}$, $\mathscr{M}'$, if $\mathscr{M}\ S\ \mathscr{M}'$, then $Q(\mathscr{M})$ iff $Q(\mathscr{M}')$. We denote by $Inv(S)$ the class of operators which are S-invariant.

By a *logic* L, we shall mean whatever class of interpreted logical symbols of first-order type (quantifiers of type level at most 2 and propositional connectives), plus the required syntactic sugar (first-order variables $x, y, ...$ and parentheses). [11]

Let K be a class of operators. The **logic** L_K **associated with K** consists in first-order variables and logical constants interpreted by operators in K. For any signature σ, we thus obtain a language $L_K(\sigma)$ with extra-logical symbols corresponding to σ, whose interpretation varies freely, and logical symbols whose interpretation is taken from K and is kept fixed. For example, let $Q \in K$ be a class of structures of the form $\langle M,R\rangle$ where $R \subseteq M \times M$, L_K contains a logical symbol $\overline{Q}$ which is interpreted by Q. This means that, in the recursive definition of satisfaction for L_K, the clause for Q is the following one:

$$\mathscr{M} \vDash \overline{Q}x,y\ \phi(x,y)\ [\sigma] \text{ iff } Q(M, ||\phi(x,y)||_{\mathscr{M},\sigma})$$

where $||\phi(x,y)||_{\mathscr{M},\sigma}$ is the interpretation of ϕ over $\mathscr{M}$ according to σ, that is the set of pairs $\langle a,b\rangle$ of elements of $\mathscr{M}$ such that $\mathscr{M} \vDash \phi(x,y)\ [\sigma][x := a][y := b]$.

[11] The definition is very (maximally !) liberal about what logic should be: at this point we do not impose any special semantic properties that 'truly logical' symbol should have.

We just gave the example for Q_{exists} of type $\langle 1 \rangle$. As an example for a quantifier of type $\langle 2 \rangle$, consider Q_{WF} the class of relational structures $\langle M, R \rangle$ where R is a well-ordering. $\overline{Q_{WF}}$ is the well-foundedness quantifier, $\overline{Q_{WF}} x, y\ \phi(x, y)$ being true iff $\phi(x, y)$ defines a well-ordered relation.

Conversely, given a logic L and an operator Q – for simplicity, we assume the type of Q is the same as before – we shall say that Q **is definable in** L iff there is a sentence ϕ_Q of $L(\overline{R})$ such that :

$$Q(\langle M, R \rangle) \text{ iff } \langle M, R \rangle \models_L \phi_Q$$

2.2 Closure under definability and generated logics

We shall now look at the conditions under which a logic may be construed as the logic generated by an invariance criterion. The invariants of a similarity relation should be *closed under definability*, this is the constraint we have alluded to in the introductory section. Let us now explain what it is about. Given a similarity relation S, $Inv(S)$ is the class of S-invariant operators. We are interested in the class $Inv(S)$ as the putative class of logical operators. This means that we want to use the operators in $Inv(S)$ as building blocks for the logical part of a language. Given a language, it is possible to define in it certain operators in a purely logical way. For example, let us consider the class K of operators containing just the existential and universal quantifiers, the operator for equality and the boolean operators. The logic associated with K is just FOL. Now, in FOL, it is possible to define new logical operators. For example, the purely logical formula "$\exists x, y, z\ ((Px \wedge Py \wedge Pz) \wedge (x \neq y \wedge x \neq z \wedge y \neq z))$" defines the operator $Q_{\geq 3}$ ("there are at least three"), which is the class of all structures of the form $\langle M, P \rangle$ where P is a subset of M containing at least three elements.[12] Even if Q was not in K, it was "implicitly" there, because it is definable in a language based on K.

We claim that operators which are definable in a purely logical manner are logical. We just do not see how a non-logical element could creep in the logical elements of the definition and make the defined operator non logical. This is what we might call the principle of closure under definability:

Principle of closure under definability. *An interpreted symbol definable only by means of logical constants is a logical constant.*

A similarity relation S is **closed under definability** if and only if every operator which is definable in a language L whose logical constants are interpreted by S-invariant operators is S-invariant. If we accept the principle of closure under

[12] The formula is purely logical since the so-called *non logical constants* are left uninterpreted here

definability, every similarity relation which can be used to characterize logical operations should be closed under definability.

Keeping in mind this closure principle, we shall say that a logic L is **exactly generated by a similarity relation** S when an operator Q is in $Inv(S)$ iff Q is definable in L. A logic is exactly generated *tout court* iff it is exactly generated by some S.[13]

The similarity relations which can be used to generate a logical language are precisely those that are closed under definability. If it were not the case, there would be a discrepancy between what the logic can express and the invariants of the similarity relations. This intuition is made precise by the following fact:

Fact 1. *S is closed under definability iff there is a logic L which is exactely generated by S.*

Proof. This straightforwardly follows from the definition of closure under definability. If S is closed under definability, taking as logical operators all the S-invariants yields a logic L whose elementary classes are S-invariant,[14] since S is closed under definability. Conversely, If S generates a logic L, $El_L \subseteq Inv(S)$, hence S is closed under definability by definition. □

Finally, note that not all familiar logics are exactly generated and not all familiar similarity relations are closed under definability. As a case in point, $S = Iso_\omega$, partial isomorphisms of finite length is not closed under definability (see the proof in the Appendix). However, Iso_ω was the natural candidate to generate first-order logic, since there is a partial isomorphism of finite length between two structures $\mathscr{A}$ and $\mathscr{B}$, $\mathscr{A}\ Iso_\omega\ \mathscr{B}$, iff $\mathscr{A}$ and $\mathscr{B}$ are elementary equivalent structures in ordinary first-order logic, but one may prove that actually no similarity relation exactly generates standard first-order logic. On the other hand, some of the natural similarity relations in this context happen to be closed under definability. This is the case for the relations chosen as a basis for the distinction between logical and non logical constants by Tarski and Sher ('being isomorphic'), by Feferman ('being strictly homomorphic') and by Bonnay ('being potentially isomorphic').

[13] Writing El_L for the class of elementary classes of L (following the usual terminology, an elementary class of structure is a class of structures which is definable by a single sentence), we have that L is exactly generated iff there is an S such that $El_L = Inv(S)$.

[14] Of course, in real life, we are interested in finding simpler logics generated by S, that is, the game consists in finding some relevant subset of invariant operators that can be used to define all the other invariants.

3 A framework for deflationary truth

We shall now turn to the formal setting for the truth predicate, which we will use to test whether the extra-expressive power it provides remains within the bounds of logic, as given by means of an invariance criterion. We are thus interested in the *expressive power* of the truth-predicate. The truth predicate, as noted above, does not come alone. It applies to naming devices describing set of sentences and those devices are called for to display the expressive abilities of truth. A framework aimed at the mathematical study of the deflationary concept of truth must thus provide a way to deal with sentences, set of sentences or whatever truth applies to. Recall also that, under deflationist interpretation, talk of sentences in truth ascriptions is primarily understood as a mere technical detour and is essentially parasitic on the intended domain of discourse. To meet those requirements, we will consider some *sorted languages* and related *sorted structures*, which in some sense mimic a familiar distinction between object-language and metalanguage where the metalanguage is understood as a tool to state in a 'formal mode' (as a Carnapian would have it) what could be said (or not!) in a material mode.

Grammatically, we associate to any langage L in signature σ a sorted language which is built out of it by adding a vocabulary of a distinguished sort, including a truth predicate, new distinguished variables, and perhaps other predicates and relation symbols of the new sort.[15] The set of sentences of this extended langage is simply the union of two sets: the set of sentences made out of the vocabulary of the first sort[16] and the set of sentences made out of vocabulary of the second sort.[17] As for the semantics of such a language, the domain of discourse over which the variables of the metalinguistic sort range and the predicates of the metalinguistic sort - including the truth predicate - are interpreted is simply a set assumed to contain at least the sentences of the object language. We call such a particular choice of a metalangage an "alethic extension" (of a logic), it comes with fixed and interpreted metalinguistic vocabulary.

Thus, let L_σ be a language of signature σ in an arbitrary logic L. An *alethic extension* $L_{\sigma,\mathscr{A}}$ for L_σ consists in:

1. adding to the basic vocabulary of L_σ variables of a new sort (printed in boldface) $\mathbf{x}, \mathbf{y}, ...$, new sort predicates $\mathbf{P_1}, ..., \mathbf{P}_\alpha$ and a truth-predicate $\mathbf{Tr}$. The set of formula and sentences in the sorted language are obtained as the union of the standardly defined sets of formula and sentences in the respective vocabularies of the two sorts.[18]

[15] The metalanguage sort of vocabulary will thereafter be printed in bold face.

[16] Using the usual building rules on the vocabulary of the first sort.

[17] We conventionally take the logical constants as being available for constructing sentences in each of the two langages - nothing hinges on that.

[18] Remark that the definition does not allow for 'mixed' formula containing vocabulary of the

2. selecting a structure $\mathscr{A} = \langle A, \overline{\mathbf{P_1}}, ..., \overline{\mathbf{P}_\alpha} \rangle$ where $\mathscr{A}$ is a set containing all the sentences of L_σ,

3. associating with each σ-structure $\mathscr{M}$ a *sorted* expansion $\mathscr{M}^{\mathscr{A}}$ of the form $\langle \mathscr{M}, \mathscr{A}, \overline{\mathbf{Tr}} \rangle$ where $\mathscr{A}$ is as above and $\phi \in \overline{\mathbf{Tr}}$ iff ϕ is a $\sigma-$sentence such that $\mathscr{M} \vDash \phi$.

Note that the perspective on truth we adopt is entirely semantic. In alethic extensions, the metalinguistic predicates and the truth predicate are added as *interpreted* predicates. Their extension is assumed to be uniquely defined: the metalinguistic predicates are interpreted on a fixed model $\mathscr{A}$ of the metalanguage and the extension of the truth predicate is uniquely determined in $\mathscr{M}^{\mathscr{A}}$ for each σ-structure $\mathscr{M}$.

Given a logic and a signature, defining an alethic extension yields a standard notion of validity. A formula ϕ in the extended language $L_{\sigma,\mathscr{A}}$ is valid (notation: $\vDash \phi$) iff for all σ-structures $\mathscr{M}$, $\mathscr{M}^{\mathscr{A}} \vDash \phi$. Our extensions are *alethic* extensions in the sense that the T-equivalences are valid, no matter how we defined our alethic extensions. First, $\vDash \phi \leftrightarrow \mathbf{Tr}(\mathbf{x})\ [\phi]$ by clause 3. In case there is a name 'ϕ' for ϕ in $L_{\sigma,\mathscr{A}}$, we get as a valid formulas the full-fledged T-equivalence: $\vDash \phi \leftrightarrow \mathbf{Tr}(\text{`}\phi'\text{)}$.

We leave much freedom regarding what is in an alethic extension and what is not. In a given alethic extension, we might get much more than the T-equivalences. For example, assume $L_{\sigma,\mathscr{A}}$ has a predicate $Sent(\mathbf{x})$ which is such that for all $a \in A$, $a \in \overline{Sent}$ iff there is a sentence ϕ of L_σ with $a = \ulcorner \phi \urcorner$, and a function $not(\mathbf{x})$ such that for all $a, b \in A$, $\overline{not(a)} = b$ iff b represents a sentence which is the negation of a sentence represented by a, then $\vDash \forall \mathbf{x}\ (\mathbf{Sent}(\mathbf{x}) \rightarrow \mathbf{Tr}(not(\mathbf{x})) \leftrightarrow \neg \mathbf{Tr}(\mathbf{x}))$. Similarly, assume that the syntactic operation of conjoining two sentences is encoded by a syntactic function $\mathbf{and}(\mathbf{x}, \mathbf{y})$ such that for all $a, b, c \in A$, $\overline{\mathbf{and}(a, b)} = c$ iff c represents a sentence which is the conjunction of two sentences represented by a and b. We will get $\vDash \forall \mathbf{x}, \mathbf{y}\ [\mathbf{Tr}(\mathbf{and}(\mathbf{x}, \mathbf{y})) \leftrightarrow \mathbf{Tr}(\mathbf{x}) \wedge \mathbf{Tr}(\mathbf{y})]$. So, as we leave the notion of alethic extension underspecified, the theory of truth associated with alethic extensions is underspecified too. In the minimal case, we get just the T-equivalences as truth-theoretic truths. Using richer alethic extensions, we can get more, *e.g.* the previous recursive clause for $\wedge$. As it will turn out, leaving the notion of alethic extension underspecified is fine, because the result we eventually prove will hold for all alethic extensions, that is, no matter how weak or strong the induced theory of truth is[19]

Here is an important fact which holds no matter the alethic extensions we choose to consider:

two sorts. This is a simplification, and we could have allowed for a richer syntax. For instance, boolean combinations of sentences of the two sorts would not harm. See footnote 19.

[19]It has to be kept in mind, however, that in any case our theory of truth is strong in the sense that we use an interpreted truth-predicate.

Fact 2. *For any L and any alethic extension $L_{\sigma,\mathscr{A}}$ for L, if $\mathscr{M} \equiv \mathscr{M}'$ then $\mathscr{M}^{\mathscr{A}} \equiv \mathscr{M}'^{\mathscr{A}}$.*

Fact 2 seems a reasonable property[20] for extensions by an interpreted truth-predicate: since the truth predicate only speaks of the truth or the falsity of the sentences of L_σ, one should not get non-elementary equivalent models from elementary equivalent ones. Note however that Fact 2 does *not* tell us that adding a truth predicate for L_σ does not add some expressive power. To see this, take L to be standard first-order logic with equality and σ the empty signature. Consider the sentences ϕ_n of the form $\exists x_1,...,x_n \forall y\ (y = x_1 \vee ... \vee y = x_n)$ and pick an alethic extension $L_{\sigma,\mathscr{A}}$ such that there is a formula $\phi(\mathbf{x})$ satisfying, for every $\mathscr{M}^{\mathscr{A}}$, $\mathscr{M}^{\mathscr{A}} \vDash \phi(\mathbf{x})\ [a]$ iff $a = \ulcorner \phi_n \urcorner$ for some n. If needed, such an alethic extension can be created by fiat: take a predicate $\mathbf{P_1}$ interpreted by $\{a \in A \ / \ a = \phi_n \text{ for some } n\}$. Then $\exists \mathbf{x}\ (\phi(\mathbf{x}) \wedge \mathbf{Tr}(\mathbf{x}))$ is a sentence of $L_{\sigma,\mathscr{A}}$ which defines the class of finite structures. By compactness, there is not even an infinite set of sentences of L_σ defining that class. This is perfectly compatible with Fact 2 being true: a finite structure and an infinite structure are not elementary equivalent in $L_{\sigma,\mathscr{A}}$, but they are not elementary equivalent in pure first order logic either.[21]

4 Truth as a logical notion

4.1 Truth completeness and closure under definability

Recall our main purpose : to prove that *if* our underlying logic is well behaved, adding a truth predicate will not gain us any expressive power. For such a 'well behaved logic', what does it mean to say that adding an interpreted truth predicate to it does not increase its expressive power ? We understand this as saying that truth extensions would provide us with new tools to express things which were already expressible in principle by purely logical means. In an abstract model-theoretic fashion, we take the expressive power of a logic to be given by the elementary

[20] The fact is obvious since $\mathscr{M} \equiv \mathscr{M}'$ just means that the same sentences of L_σ are true in $\mathscr{M}$ and $\mathscr{M}'$ and by our definition of alethic extensions.

[21] Note that for Fact 2 to hold it is crucial that a sorted language is used. Assume we are dealing with a very simple language which has only one sentence which says that there are no more than n objects. Consider two structures $\mathscr{M}$ and $\mathscr{M}'$ with $|M| = n-1$ and $|M'| = n$. In our toy language, these two structures are elementary equivalent. Now consider an alethic expansion in which the syntactic domain A has only one object representing the only sentence in the language. If one were to use a non-sorted language, $\mathscr{M}^{\mathscr{A}}$ and $\mathscr{M}'^{\mathscr{A}}$ would not be elementary equivalent anymore, just because now $|\mathscr{M}'^{\mathscr{A}}| > n$ whereas $|\mathscr{M}^{\mathscr{A}}| = n$. However, the constraints we have imposed on the syntax of $L_{\sigma,\mathscr{A}}$ could be somewhat relaxed so as to allow sentences combining variables and predicates of the two sorts. If the combinations allowed are reasonable, then expect Fact 2 to remain true.

classes associated with it. The following notion of truth-completeness of a logic then captures what we are after:

Definition 1. *A logic L is truth-complete iff for every signature σ, for every alethic extension $L_{\sigma,\mathscr{A}}$, for every sentence ϕ of $L_{\sigma,\mathscr{A}}$, there is a sentence $\hat{\phi}$ of L_σ such that for all L_σ-structures $\mathscr{M}$:*
$\mathscr{M}^{\mathscr{A}} \vDash \phi$ *iff* $\mathscr{M} \vDash \hat{\phi}$.

By quantifying universally over alethic extensions, we require this no matter how powerful the alethic extensions (intuitively, an alethic extension is more powerful if it has more definable sets of sentences to which the truth predicate can be applied).

We can now state the following:

Theorem. *L is exactly generated iff it is truth-complete*

When L is exactly generated, it will be truth-complete hence truth will be implicitly definable[22] *no matter how strong the the underlying alethic extension is*. Therefore the condition that L is exactly generated is quite powerful. Even if we hardwire a lot of definable sets over alethic extensions through interpreted predicates over the set of sentences of L_σ, so that the truth predicate has a lot to speak about, L will still be powerful enough to say all that can be said with that truth predicate. On the other hand, note that the implicit definability of truth we get for a truth-complete logic is not uniform: for every sentence containing the truth predicate, there is an equivalent sentence which does not contain it, but this does not imply that a translation procedure exists.

The right-to-left part of the theorem hinges on the fact that our notion of alethic extension is very liberal. All invariant operators we need to define can be defined with the truth predicate only because we make the very strong assumption that all what we might need to make definition through the truth predicate can be made available by the syntactic predicates. We remained silent on the strength of the metatheory we need to fix the interpretation of all those interpreted syntactic predicates, which of course will depend itself on how fine-grained the underlying similarity relation is.

4.2 Interpreting the result

The Theorem establishes an equivalence between two properties of logics: being truth-complete and being generated by a similarity relation. The connection

[22] By *implicitely definable* we mean that any class of structures which is definable (in the sense given on page 2) with the truth predicate, is definable without it.

with closure under definability, which we take to be well-motivated constraint on similarity relations, is provided by Fact 1.

Now consider a particular logic L. Is L truth-complete ? If it is, there's no reason not to count the deflationist truth-predicate as logical. But what if it is not? By Theorem 6, L is not exactly generated. Whatever the notion of logical similarity S we think is appropriate, if the logical operations of L are S-invariant, then there are some S-invariant operations which are not definable in L. So the failure of truth-completeness for L should not count as an argument against the logicality of the truth-predicate, because it may result from L's unduly restricted range of available logical operations. And indeed, it does, by Theorem 6, if we were to move from L to an extension L' strong enough to define all S-invariants, then L' would indeed be truth-complete. To repeat, it might well be the case that a given logic is not truth complete. But if we consider that this logic only partially expresses the S-invariants for some S which is closed under definability, then we know that the extra expressive power provided by the truth predicate should not count as substantial. It stays within the realm of logic, in the sense that if the logic had been powerful enough to express all the invariants, then the truth predicate would have been implicitly definable.

We now turn to examples of generated logics. Here are some examples, with the similarity relations that generate them :

- propositional calculus over finite signatures
- on a fixed domain, $L_{\infty,\omega}$ and potential isomorphisms.
- on a fixed domain, $L_{\infty,\infty}$ and isomorphisms.
- on a fixed domain, $L^{-}_{\infty,\infty}$ and strong homomorphisms

Note that the last three logics are infinitary. In those cases, it was quite expected that adding an (interpred) truth predicate will not add any expressive power: the infinitary combinations of sentences that can be expressed thanks to the truth predicate can already be expressed in the logic.

Finally, it should be emphasized that only one direction of our theorem is required to support our claim that deflationary truth is a logical notion. For the logicality of the deflationary truth predicate is understood as a consequence of the fact that closure under definability implies truth-completeness. And this part of the result holds no matter the class of alethic extension one may find conceptually acceptable. If one is willing to allow only sets of sentences to interpret the metalinguistic predicates in alethic extensions, or only those sets of sentences that are computable, or those that are definable in such and such theory, or whatever, this direction of the result still holds. Although it would certainly be interesting

to find a natural characterization of the accordingly modified notions of truth-completeness in terms of invariance conditions, our not doing so does not count against our main conclusion concerning the concept of truth.

5 Expressive power of the truth predicate or proof-theoretic strength of truth theories ?

Our result supports the deflationist's claim that truth can be thought of as logical and, in this sense, a "non-substantial" notion. It might be of interest to compare our methodology to the one driving the so-called "conservativity argument" against deflationism[23]. Very briefly, the idea behind the conservativity argument is the following : were "true" not to have any explanatory power, the Tarskian truth theoretic extension $T(A)$ for a theory A should be a conservative extension of A. But, the argument continues, this is not the case: for many theories A their Tarskian truth-theoretic extension explains (in the sense of proving) facts statable in non-alethic A-terms which the theory A does not explain. The classical exemple of this phenomenon is the provability of the consistency of *PA* in Tarski's theory of truth for *PA*. Hence deflationism is false, or so it argued.

Conservativity-based approaches to the truth-predicate focus on the strength of *theories supposed to somehow define the truth predicate*. They ask : how strong must such a theory be to adequately constrain the interpretation of the truth predicate over a given base theory B ? Our standpoint is different. To study the logicality of an expression we first take its interpretation as *given* and look out whether it is logical or not. The logicality of an expression, if the invariance approach has got it right, is something about what kind of distinctions an expression can do between different structures of interpretation of the language. There are those expressions, such as "red", that distinguish between structures on account of some empirical features of them, or strong mathematical content of them, while logical expressions should distinguish among structures only on account of what could be argued to be purely non-mathematical, formal, or most general, features of them.

Both approaches to the "substantiality of truth" face difficulties when the time comes to carry out their program. Conservativity-based approaches have to face the fact that non-conservativity results are not robust, in the sense that there are base theories over which the target Tarskian truth-theoretic extensions *are* conservative. For instance, the Tarskian truth-theoretic extension A is conservative over A when A is first-order Peano arithmetic augmented with an omega-rule or, to take a quite different exemple, when A is some specific arithmetic theory weaker

[23] Shapiro (1998), Ketland (1999).

than *PA*.[24] The alleged philosophical conclusion to be drawn from the logical fact of non-conservativity is thus afflicted of unstability: the notion of Tarskian truth appears as substantial or not depending on the choice of the base theory. Yet the choice of a base theory over another as a suitable basis for assessing the substantiality of truth is by no mean settled.[25] On our semantic road, we can do better.

As an interpretation for the truth-predicate in a language L, we have taken, in each structure $\mathscr{M}$, the set of truths-in-$\mathscr{M}$ in the standard model-theoretic sense. That is, given a language *L*, the interpretation of truth-in-L is taken to be a function of L but is not taken to bring out just a "fixed" set of L-sentences. Rather, truth-in-L is interpreted as a set of such sets varying along the extensions of other non-semantical terms appearing in the different models. In our model, the extension of "true" in a structure $\mathscr{M}^*$ is entirely dependent on non-semantical states of affairs as given by the structure $\mathscr{M}$, a feature which fits well with the broadly received idea that truth supervenes on non-semantical facts. So our truth predicate *has* a definite extension over the class of all models, and it is, in each model $\mathscr{M}$, the set of truths-in-$\mathscr{M}$. Why should our above main result be understood as a case for the logicality of this truth predicate ? At first sight our semantic approach faces a problem quite similar to the one we have just mentioned in the case of conservativity-based approaches, one of unstability. For our interpreted truth-predicate increases the expressive power of some logics while it does not increase the expressive power of other, so that the truth-predicate appears as substantial in some cases while it does not appears so in others. But, importantly, we have offered a philosophically significant necessary and sufficient condition under which it appears as "substantial": it is when the underlying logic is not generated. This condition is significant because closure under definability of a similarity relation seems to be by itself a well-motivated constraint on similarity relation, with the consequence in acceptable logic (i.e. generated by a acceptable similarity relation) truth will indeed be deflated as purely logical.

[24] See Halbach (1999) for a finitist proof of this. The theory can be taken to be Robinson Arithmetic *Q* with suitable "unique readability principles".

[25] For instance the following questions do not have universally accepted answers. Is the relevance of the base theory for a philosophical assessment of truth to be judged from the fact that its truth-theoretic extension meets some adequacy conditions, for instance the provability of some reflection principles over the base theory, as Ketland (1999) has it ? (See also Leitgeb (1999), Halbach and Horsten (2015). See Fischer (2015) for another take on the issue.) Or should it be assessed solely from its intrinsic virtues as a syntactical theory - assuming that this 'intrinsic' talk could be made precise ? But then, regarding the notion of truth itself, what is the proper interpretation of the fact that conservativity of Tarskian axioms for truth over *PA* (say), obtains or does not obtain depending on whether induction axioms in the base theory are understood as a list or as a scheme ? On this last issue, see e.g. Field (1999), and more recently, in connection with the logicality of truth, our remarks in Galinon (2015).

6 Conclusion

We do not claim that our results show a certain brand of deflationism to be the best view of truth available on the market, nor that the ordinary notion of truth is a purely logical notion. Our purpose was more modest. It was to lend support to the view that *deflationary truth* is a logical notion. More precisely we have shown that on the background of an invariance-based conception of logicality it is consistent to claim that *true* has the meaning the deflationist says it has and that truth so understood is a logical notion.

Appendix : proofs

Fait. *Iso_ω is not closed under definability.*

The proof is an elementary exercise in model theory.

Proof. We shall consider the operator $Q_{\geq \aleph_0}$. $Q_{\geq \aleph_0}$ is Iso_ω-invariant: let $\langle M,P\rangle$ and $\langle M',P'\rangle$ be two structures such that we have $Q_{\geq \aleph_0}(\langle M,P\rangle)$ but not $Q_{\geq \aleph_0}(\langle M',P'\rangle)$ (thus, $|P| \geq \aleph_0$ where as $|P'| < \aleph_0$). There is an integer n such that $|P'| = n$, but then it is not the case that $\langle M,P\rangle Iso_{n+1}\langle M',P'\rangle$, hence it is not the case that $\langle M,P\rangle Iso_\omega\langle M',P'\rangle$. We shall now consider the operator Q' defined by the sentence "$\overline{R}$ is an equivalence relation $\wedge$ $\exists x\overline{Q_{\geq \aleph_0}y}\, x\overline{R}y$" ($Q'$ picks out the relational structures $\langle M,R\rangle$ such that R is an equivalence relation with an infinite equivalence class). Since $Q_{\geq \aleph_0} \in Inv(Iso_\omega)$, we have that $Q' \in CInv(Iso_\omega)$. It is now sufficient to show that Q' is not Iso_ω-invariant. We construct two $L(\overline{R})$-structures $\mathscr{M} = \langle M,R\rangle$ and $\mathscr{M}' = \langle M,R'\rangle$ such that:

- The interpretation of $\overline{R}$ on both models is an equivalence relation.
- $\mathscr{M}$ contains an infinite number of R-equivalence classes of arbitrary big finite cardinality, but no infinite equivalence class.
- $\mathscr{M}'$ is just as $\mathscr{M}$ but it contains also an infinite equivalence class.

It is clear that $\mathscr{M}$ Iso_ω $\mathscr{M}'$. But we have that $Q'(\mathscr{M}')$, whereas $\mathscr{M}$ is not in Q'.

□

Theorem. *L is exactly generated iff it is truth-complete*

Proof. **If L is exactely generated then it is truth-complete**
Let L be a logic and S a similarity relation such that $El_L = Inv(S)$. Let σ be an arbitrary signature, $L_{\sigma,Tr}$ an arbitrary alethic extension for L_σ and ϕ a sentence

of $L_{\sigma,Tr}$. It is sufficient to show that $Q_\phi = \{\mathscr{M} \ / \ \mathscr{M}^* \vDash \phi\}$ is closed under S, because then $Q_\phi \in El_L$ by $El_L \supseteq Inv(S)$. Assume that $\mathscr{A} \in Q_\phi$ and $\mathscr{A}\ S\ \mathscr{B}$. We want $\mathscr{B} \in Q_\phi$.

By hypothesis, $El_l \subseteq Inv(S)$, hence $\mathscr{A}\ S\ \mathscr{B}$ implies $\mathscr{A} \equiv_{L_\sigma} \mathscr{B}$ (or there would be a sentence ψ of L_σ which is true in $\mathscr{A}$ and not in $\mathscr{B}$, which implies that the class of models of ψ is not closed under S, contradicting $El_l \subseteq Inv(S)$). By Fact 2, $\mathscr{A} \equiv_{L_\sigma} \mathscr{B}$ implies in turn $\mathscr{A}^* \equiv_{L_{\sigma,Tr}} \mathscr{B}^*$. Since $\mathscr{A}^* \vDash \phi$, $\mathscr{B}^* \vDash \phi$ as well, hence $\mathscr{B} \in Q_\phi$ as required.

If L is truth-complete , there is an S such that $El_L = Inv(S)$. We take for S the relation $\equiv_L$ of L-elementary equivalence. First, $El_L \subseteq Inv(S)$. Let Q be an operator in El_L, then $\mathscr{M} \in Q$ and $\mathscr{M}' \equiv_L \mathscr{M}$ straightforwardly implies $\mathscr{M}' \in Q$: the formula ϕ defining Q is true in $\mathscr{M}$, hence in $\mathscr{M}'$ as well.

For $El_L \supseteq Inv(S)$, let $Q \in Inv(\equiv_L)$. We choose a signature σ matching Q's similarity type and we are looking for a formula χ of L_σ such that its models are precisely the structures in Q. Let I be an indexing of the models $\mathscr{B}$ in Q. We note $\mathscr{B}_i$ the structure indexed by i for $i \in I$. We pick a particular alethic extension $L_{\sigma,Tr}$ such that there is a formula $\psi(\mathbf{x},\mathbf{y})$ of L_σ, Tr such that for all $\mathscr{M}*$, for all $i, j \in A$, $\mathscr{M}^* \vDash \psi(\mathbf{x},\mathbf{y})\ [i,j]$ iff $i \in I$ and $j = \ulcorner\phi\urcorner$ for some sentence ϕ of L_σ with $\mathscr{B}_i \vDash \phi$[26]. Let χ be the formula $\exists\mathbf{x}\forall\mathbf{y}\ (\psi(\mathbf{x},\mathbf{y}) \rightarrow Tr(\mathbf{y}))$. By truth-completeness of L, there is a formula $\hat{\chi}$ such that for all $\mathscr{M}$, $\mathscr{M} \vDash \hat{\chi}$ iff $\mathscr{M}^* \vDash \chi$. $\hat{\chi}$ is the sentence we are after. If $\mathscr{M} \in Q$, $\mathscr{M} = \mathscr{B}_i$ for some $i \in I$. So, by definition, $\mathscr{M}^* \vDash \psi(\mathbf{x},\mathbf{y})\ [i,j]$ implies $j = \ulcorner\phi\urcorner$ for some ϕ with $\mathscr{M} \vDash \phi$, hence $\mathscr{M}^* \vDash Tr(\mathbf{y})\ [j]$. Therefore, $\mathscr{M}^* \vDash \chi$, hence, $\mathscr{M} \vDash \hat{\chi}$. Conversely, assume $\mathscr{M} \vDash \hat{\chi}$. Then, $\mathscr{M}^* \vDash \chi$, so there is an $i \in I$ with $\mathscr{B}_i \in Q$ such that for all b with $b = \ulcorner\phi\urcorner$ for some ϕ, $\mathscr{M}^* \vDash \psi(\mathbf{x},\mathbf{y})\ [i,b]$ implies $\mathscr{M}^* \vDash Tr(\mathbf{y})\ [b]$. By definition of ψ and Tr, we have $\mathscr{M} \vDash \phi$ for all ϕ such that $\mathscr{B}_i \vDash \phi$, hence $\mathscr{M}$ and $\mathscr{B}_i$ are L_σ-elementary equivalent. Since $Q \in Inv(\equiv_L)$ and $\mathscr{B}_i \in Q$, this implies that $\mathscr{M} \in Q$ as well.

□

[26]Should we be worried whether such an alethic extension exist? As noted in paragraph 3, alethic extensions are so defined that we are free to use interpreted syntactic predicates, the only limit, so to speak, being that Fact 2 will always hold. So in our case, we could just add a syntactic relation $\overline{R}$ doing what we want, that is, such that for all $i, j \in A$, $\langle i,j\rangle \in R$ iff $i \in I$ and $j = \ulcorner\phi\urcorner$ for some sentence ϕ of L_σ with $\mathscr{B}_i \vDash \phi$. The only thing we have to worry about is that the alethic extension has to be big enough to encode the indexing I. Note that if Q is a proper class, this will have the somewhat awkward consequence that our alethic extension is itself a proper class. In the examples given in the next paragraph, restriction to a fixed domain ensures that proper classes are not needed.

References

Achourioti T., Galinon, H., Martinez, J., Fujimoto, K. (2015), *Unifying the philosophy of truth*, New York, Springer Verlag.

Bonnay, D. (2008), Logicality and Invariance, *The Bulletin of Symbolic Logic*, 14 (1), pp. 29-68.

van Eijck, J. (1996), Quantifiers and Partiality, in J. van der Does & J. van Eijck (ed.) *Quantifiers, logic and language* pp. 105-144, CSLI, Stanford.

Feferman, S. (1999), Logic, Logics, and Logicism, *Notre Dame Journal of Formal Logic*, 40 (1), pp. 31-54.

Field, H. (1999), Deflating the Conservativness Argument, *Journal of Philosophy*, 96, pp. 533-540.

Fischer, M. (2015) Deflationism and Instrumentalism, *in* Achourioti & al. (2015), pp. 293-306.

Galinon, H. (2015) Deflationary Truth: Conservativity or Logicality ?, *Philosophical Quarterly*, 65 (259), pp. 268-274.

Halbach, V. (1999) Deflationism and Infinite Conjonctions , *Mind*, 108, pp. 1-22.

Halbach, V. and Horsten, L. (2015) Norms for theories reflexive truth, *in* Achourioti & al. (2015), pp. 263-280.

Horwich, P. (1998) *Truth* 2nd edn. Oxford, Oxford University Press.

Ketland, J. (1999) Deflationism and Tarski's Paradise, *Mind*, 108, pp. 69-94.

Leitgeb, H. (2007) What theories of truth should be like (but cannot be), *in Blackwell Philosophy Compass 2/2*, Blackwell, pp. 276-290.

McGee, V. (1996) Logical Operations, *Journal of Philosophical Logic*, 25, pp. 567–580.

McGinn, C. (2000) *Logical properties: Identity, Existence, Predication, Necessity, Truth*, Oxford, Oxford University Press.

Shapiro, S. (1998) Truth and Proof: Through thick and thin, *Journal of Philosophy*, 95 (10), pp. 493–521.

Sher, G. (1991), *The Bounds of Logic: A Generalized Viewpoint*, Cambridge, MIT Press.

Tarski, A. (1986), What are logical notions, *History and Philosophy of Logic*, vol. 7, pp. 143-154.

Wyatt, J. (2016) The Many (Yet Few) Faces of Deflationism, *Philosophical Quarterly*, 66 (263), pp. 362-382.