Philosophia (2018) 46:817–830
https://doi.org/10.1007/s11406-018-9948-y

Pinocchio against the Semantic Hierarchies

Peter Eldridge-Smith[1]

Received: 15 October 2017 / Revised: 9 January 2018 / Accepted: 16 January 2018 /
Published online: 24 January 2018

Abstract The Liar paradox is an obstacle to a theory of truth, but a Liar sentence need not contain a semantic predicate. The Pinocchio paradox, devised by Veronique Eldridge-Smith, was the first published paradox to show this. Pinocchio's nose grows if, and only if, what Pinocchio is saying is untrue (the Pinocchio principle). What happens if Pinocchio says that his nose is growing? Eldridge-Smith and Eldridge-Smith (*Analysis, 70*(2): 212-5, 2010) posed the Pinocchio paradox against the Tarskian-Kripkean solutions to the Liar paradox that use language hierarchies. Eldridge-Smith (*Analysis, 71*(2): 306-8, 2011) also set the Pinocchio paradox against semantic dialetheic solutions to the Liar. Beall (2011) argued the Pinocchio story was just an impossible story. Eldridge-Smith (*Analysis, 72*(3): 749-752, 2012b) responded that unless the T-schema is a necessary truth of some sort (logical, metaphysical or analytic), the Pinocchio principle is possible. Luna (*Mind & Matter 14*(1): 77–86, 2016) argues that the Pinocchio contradiction proves the principle is false. D'Agostini & Ficara (2016) discuss a more plausible physical truth-tracking trait, the Blushing Liar, and argue that the Pinocchio contradiction is not a metaphysical dialetheia. I respond to Luna, and D'Agostini & Ficara, and prove that the Pinocchio paradox is a counterexample to hierarchical solutions to the Liar.

Keywords Pinocchio paradox · Liar paradox · Principles of truth · T-schema · Tarski's theory of truth · Kripke's theory of truth · Semantic validity

1 Introduction

This work concerns whether there are necessary principles of truth in relation to semantic validity, the Pinocchio paradox and some other conundrums. I argue the

The paper was presented at the International Conference on Truth, Logic and Philosophy at the University of Peking in Beijing in September 2017. I would like to thank the attendees for their discussion.

✉ Peter Eldridge-Smith
peter.eldridge-smith@anu.edu.au; peter.eldridgesmith@gmail.com

[1] School of Philosophy, Research School of Social Sciences, Australian National University, Canberra, Australia

stratification of the T-schema fails to avoid all Liar paradoxes, particularly the Pinocchio paradox. The article proceeds in three sections. The first introduces some of the issues to be discussed and begins some analysis. The second recounts some logical adventures of Pinocchio across the logical pluriverse. The third uses these stories to progress the analysis a little more.

2 The T-Schema, its Hierarchical Variant, the Liar and Pinocchio Paradoxes, and some Comments on Truth in Relation to Validity

Characteristic of correspondence and deflationary conceptions of truth are a number of principles, in particular a truth schema (Tarski 1933).

> (T-schema) $T\langle\Phi\rangle$ if and only if (iff) Φ
>
> where T is the truth predicate, Φ is replaced by any sentence and $\langle\Phi\rangle$ by a canonical name of that sentence.

Rules for introducing and eliminating the truth predicate.

> Truth Introduction (TI): $\Phi \quad T\langle\Phi\rangle$
>
> Truth Elimination (TE): $T\langle\Phi\rangle \quad \Phi$

The intersubstitutability of truth, also known as *transparency* (Field 2008:12):

> Sentences Φ and $T\langle\Phi\rangle$ can be substituted for each other in any non-opaque context without changing the semantic value of a sentence of which they are part (or the whole).

I note that I prefer to restrict these principles to canonical names because they are more formal. In the context of the current discussion, this gives these principles their best chance of proving necessary.

There is a convincing argument that the T-schema is not a logical truth (Cook 2012).[1] There may be a way around Cook's argument; one might treat the truth predicate as a logical constant. We will have cause to discuss whether any of these principles, particularly the T-schema and its hierarchical variants, are logically, analytically or metaphysically necessary.

The Liar paradox is an obstacle for a theory of truth that entails each instance of the T-schema. Such a theory would in effect take the T-schema to be necessarily true (i.e.

[1] Cook gives the following succinct proof that the intersubstitutability of truth, which he labels 'T-substitutivity', is not a logical necessity.

> [L]et Φ_4 be a logical truth named by t_4, and Φ_5 a logical falsehood named by t_5. Since Φ_4 is a logical truth, by T-substitutivity, $T(t_4)$ is a logical truth, but then by Logical substitutivity [that is, the substitutivity of non-logical constants in a logical truth], $T(t_5)$ is a logical truth, but then by T-substitutivity Φ_5 is a logical truth. Contradiction. (Cook2012:236)

true for any well-formed sentence). However, given classical logic, paradox would ensue. For example, where '*a*' is used as a singular term:

(i) $a = \langle{\sim}Ta\rangle$ Premise, contingent or necessary,[2] and that mentions an embedded Liar sentence, i.e. '$\sim Ta$', to the effect of 'This sentence is not true'.
(ii) $T\langle{\sim}Ta\rangle \leftrightarrow {\sim}Ta$ T-schema
(iii) $Ta \leftrightarrow {\sim}Ta$ (i), (ii) = E
(iv) $Ta \,\&\, {\sim}Ta$ (iii) sentential logic (SL)
(v) Q (iv) SL, viz., *ex contradictione quodlibet* (ECQ), also known as *Explosion*

There are various Liar scenarios and syntactic functions that give rise to a Liar sentence, like the sentence mentioned in (i). If the argument to (v) is sound, the language in which it is expressed is trivial. Imagine someone reasons in the above way to any and every sentence Q. The Liar argument shows that either not all instances of the T-schema can be true (for a language in which such a Liar sentence is well-formed as in a natural language like English) or the derivation of (iii), (iv) or (v) is fallacious. Moreover, the Liar paradox is two-faced, in that it poses both an issue about the semantic value of a Liar sentence as well as the above issue about what is wrong with the derivation of the paradox.

It has been popular to index truth in some way, relating it to a hierarchy of some sort or another. In particular, Tarski's response to the Liar was to define the truth predicate for a language in a meta-language (Tarski 1933). Then an indexed T-schema becomes characteristic of a hierarchy of truth predicates.

(Indexed T-schema): $T_i(t_\Phi)$ iff Φ

where each instance of the biconditional is an $(i+1)$th level sentence, T_i is a truth predicate, t_Φ is a canonical name for a sentence of the i th level language whose translation into the $(i+1)$th level language is Φ.[3]

A related strategy is to restrict use of the truth predicate, so that it only takes a semantic value when predicated of germane expressions, such as *grounded* sentences. In particular, one may distinguish an extension of truth from its anti-extension, allowing for a gap for Liar sentences (Kripke 1975). Then, one might expect that an empirical sentence predicated by truth would have a non-paradoxical evaluation. However, a Liar sentence need not contain a semantic predicate, it may be an empirical sentence. The Pinocchio paradox, devised by Veronique Eldridge-Smith, was the first published paradox to show this.

Pinocchio was the hero of a popular Italian children's novel. He was a puppet who grew into a little boy when (at the end of the novel) he learnt to pull his own strings. On principle, Pinocchio's nose grows when he tells a lie. The paradox ensues from

[2] Quine (1995) shows how such an identity can be derived using a syntactic function, in particular a self-predication function.

[3] Adequate definitions of truth will entail each instance of this schema, or each instance of a more general schema using the truth relation or the converse satisfaction relation. My presentation assumes for convenience that the meta-language contains (and extends) the object-language. This will make proofs in section 3 more perspicuous. The same results can readily be obtained without the use of canonical naming. Nevertheless, as I stated previously, canonical naming, being more formal, gives the T-schema its best chance of being necessary.

Pinocchio saying his nose is growing (or will grow in some versions). Here is a version of the Pinocchio paradox[4]:

(Pinocchio principle): Pinocchio's nose grows iff what he is saying is untrue (is a lie)

(1) What Pinocchio is saying = 'Pinocchio's nose is growing' Premise

(2) 'Pinocchio's nose is growing' is true iff Pinocchio's nose is growing (T-schema)

(3) Pinocchio's nose is growing iff 'Pinocchio's nose is growing' is untrue (1), (Pinocchio principle) =E

(4) 'Pinocchio's nose is growing' is true iff 'Pinocchio's nose is growing' is untrue (2), (3) SL

Alternatively, (5) Pinocchio's nose is growing iff Pinocchio's nose is not growing (2), (3) SL

(6) Pinocchio's nose is growing and Pinocchio's nose is not growing, (5) SL

From (4) or (5) one can derive an explicit contradiction, such as (6), and Explosion will result in triviality. It will be useful to use the following labels in some discussion:

(the Pinocchio statement): "My nose is growing" or "Pinocchio's nose is growing"

(Pinocchio scenario): Pinocchio principle and Premise (1), which gives the identity of what Pinocchio is saying

The Pinocchio paradox uses an empirical predicate 'is growing', which is not a synonym for the truth predicate, so the Pinocchio statement can appear in the object language and poses the threat of a Liar paradox re-appearing, even for a hierarchical solution. Nor is the Pinocchio paradox alone anymore. These days Pinocchio has a sister, Victoria, who blushes just when she lies: D'Agostini and Ficara (2016) point out that blushing is a more plausible tell-tale for lying than having one's nose grow.[5] It is indeed; now Pinocchio and Victoria have many cousins; for example, Heimlich, the coughing liar,

[4] While the Pinocchio paradox and the Liar pose obstacles for a theory of truth, whether these paradoxes need to use a principle of truth depends on one's theory of truth, truth bearers and logical systems. For example, Prior (1958) did not use a truth predicate, but still found it necessary to devise some way of avoiding a version of the Epimenides paradox. In the case of the Pinocchio paradox, an alternative principle would be 'Pinocchio's nose grows iff what he is saying is not the case', but this would only avoid having to use the T-schema if one's theory and logical system was such that 'is not the case' was not just a synonym for the truth predicate. Most people think it is; in any case, my concern here is to pose a challenge for theories of truth, not to avoid the T-schema or similar principles of truth.

[5] D'Agostini and Ficara (2016) credit the idea for the Blushing Liar paradox to Victoria Schneider.

who always and only gives a distinctive cough after lying and before he says anything else. While Victoria and Heimlich gain in plausibility, they may lose a little in generality. We may suppose their tell-tale principles are psychological or causal. The nature of the Pinocchio principle is less specific; it is supposed to be a principle, whether it is a causal truth or even say a metaphysical necessity is usually unspecified. Ultimately, an analytic principle of truth, if there is one, ought to show why one cannot have a metaphysical, analytic or causal principle like the Pinocchio principle. This would require an argument for that truth principle being necessary, not merely an appeal to intuitions.

Eldridge-Smith and Eldridge-Smith (2010) posed the Pinocchio paradox against Tarski's hierarchical solution to the Liar, but did not prove the contradiction is derivable in systems implementing that solution. To be proven against Tarski's theory of truth, one needs to use a hierarchical variation of the T-schema, such as the Indexed T-schema.[6] Eldridge-Smith and Eldridge-Smith (2010) do state that the Pinocchio paradox is provable in Kripke's object language in a world in which the Pinocchio principle holds. Indeed, the Pinocchio contradiction is provable as above in a Kripkean object language, but whether that argument is really a paradox or a *reductio ad absurdum* of the Pinocchio principle is then in question. I return to this matter, which is raised in Luna (2016), shortly.

Another strategy to deal with the liar paradox is to modify sentential logic to avoid either (iv) or (v). The dialetheic strategy accepts (iv), but rejects Explosion. Among the dialetheic theories are ones that restrict dialetheia to semantic contradictions, such as Beall (2009). Eldridge-Smith (2011) also sets the Pinocchio paradox against semantic dialetheic solutions to the Liar. Beall (2011) argued the Pinocchio story was just an impossible story. Eldridge-Smith (2012b) responded that unless the T-schema is a necessary truth of some sort (logical, metaphysical or analytic), the Pinocchio principle is possible. Basically, the Pinocchio scenario on its own is not a contradiction; a contradiction is derivable from the scenario given the T-schema. Unless the T-schema is some kind of necessary truth, the Pinocchio scenario is not impossible.

(6), the contradiction that follows from (5) above, is a gross physical contradiction, an undesirable dialetheia for most dialetheists, and a counterexample to semantic dialetheism and perhaps other forms of dialetheism. Nevertheless, D'Agostini and Ficara (2016) argue that the Pinocchio contradiction is not a metaphysical dialetheia because the T-schema is a metaphysical necessity. And I will try to argue against this in parts of the next two sections.

For present discussion, *the Pinocchio principle* is supposed to be just that – a principle – not a merely contingent statement that happens to have been true up till Pinocchio makes his statement. Supposing logical worlds may have non-

[6] Instead, Eldridge-Smith and Eldridge-Smith (2010) use the Pinocchio paradox to criticise part of Tarski's analysis of the Liar.

> The Pinocchio paradox is, in a way, a counter-example to solutions to the Liar that would exclude semantic predicates from an object-language, because 'is growing' is not a semantic predicate. Tarski's analysis of the source of pathology of which the Liar is symptomatic led him to conclude it arose from free use of semantic predicates in the object-language. Tarski's solution was to restrict such predicates strictly to the metalanguage. Intuitively, predicates like 'is growing' are typical of just the sorts of predicates one wants in a useful object-language. If empirical predicates like 'is growing' need to be restricted in the object-language to avoid versions of the Liar, the intuitive bounds on which predicates need to be restricted in the object-language to avoid Liar-like paradoxes have been breached. [Eldridge-Smith and Eldridge-Smith 2010: 213]

logical principles, as causal, metaphysical or analytic necessities, then the case is that the Pinocchio principle is supposed to be one of these.

In partial contrast, Luna (2016) uses the Pinocchio paradox as a *reductio ad absurdum* in an argument against physicalism, particularly mind-body physicalism. (For details of Luna's argument, please see the work cited.) Luna merely assumes the Pinocchio principle towards a reductio and takes the T-schema (or a hierarchical variant of it) as a necessary truth. However, as above, the paradox is not a straight forward reductio of the Pinocchio scenario unless it can be somehow proven that the T-schema is a necessary truth (logical, analytic or metaphysical). The Liar paradox already requires either a modification of classical logic or a restricting of the T-schema. Apart from regularity with respect to other non-problematic instances, what do we have to support the T-schema beyond strong belief?[7]

In short, the Pinocchio scenario is possible unless the T-schema is a necessary truth. Without a proof that the T-schema holds in this case, rejection of the Pinocchio scenario as impossible seems dogmatic. Luna may, and many others would presumably claim that the T-schema is an analytic truth, and therefore a necessary truth. Cook (2012: 237), for example, seems to make such a claim. I think that the analysis of truth is slightly more complex; while the T-schema provides the normal relation between being the case and being true, I will argue that the T-schema obfuscates a distinction required to address the semantic value of Liar sentences, a distinction foreshadowed in works on the Liar by the scholastic logicians, Thomas Bradwardine and John Buridan.

There are other concepts related to purportedly analytic necessities involving truth. In particular, the semantic conception of validity is that validity preserves truth.

(Semantic validity): It is impossible for the conclusion of a valid argument not to be true when its premises are true (at the same time and in the same way).

Given that validity preserves truth, then:

$$A \vDash A$$

is valid iff

$$T\langle A\rangle \vDash T\langle A\rangle$$

is valid. Does it follow just from our definition of validity and that '$A \vDash A$' is valid, that either

$$T\langle A\rangle \vDash A, \ (or)$$

$$A \vDash T\langle A\rangle$$

is valid? Well, no. The valid deductive inference of A from A is not ampliative; that is an analytic truth about deduction, so this is not a basis for claiming that A entails $T\langle A\rangle$ or the converse.[8] (Of course, if A is a necessary truth, A validly entails $T\langle A\rangle$ on the

[7] Personally, I think this only slightly complicates the argument against physicalism. The complication is that either the instance of the Pinocchio principle or the relevant instance of the T-schema is false by reductio, but either way I think Luna still has an argument against physicalism.

[8] The aporia of deductive logic not being ampliative is ECQ or Explosion.

normal modal definition. But that is just a special case.) So the modal definition of validity, which preserves truth, does not necessitate the purported principles of truth: the T-schema, TI, TE, or *Transparency*.[9]

Many purported principles of truth like the T-schema seem to suffer from exceptions or paradoxes. Even the modal definition of validity seems to have exceptions. Here is an example I have updated from one of Buridan's sophisms (Hughes 1982).

(P1) All truths are sentence tokens
(P2) All truths are affirmative
(C1) No true sentence token is negative (i.e. non-affirmative)

This is a counterexample to the definition of validity based on preserving truth. We can imagine a possible world where the premises are true but the conclusion is not, even though it is the case! What has to happen here, broadly in line with Buridan's own theory, is for truth and what is the case to be distinguished (and not always go together). Eldridge-Smith 2004 traces the thought that what is true and what is the case sometimes come apart back to the scholastic logicians Bradwardine and Buridan. Particularly with respect to Liar sentences, Bradwardine's brilliant idea was that although things are as the sentences says, not everything is as is entailed by a Liar sentence, and therefore it is false. Another way of making this distinction is to distinguish truth in a world from truth at a world, and there are technical ways of representing that. Indeed, I think an argument is valid iff it is impossible for the premises to be true *at* a world and the conclusion to be untrue *at* such a world. That is, the above is a case where the premises may be true *in* and *at* a world, but the conclusion is only true *at* such a world and is untrue *in* such a world. I think it follows from the validity of the above argument and the possible truth of its premises in some possible world, that there can be sentences, such as (C1), in a possible world such that they are the case but are not true (in that world).

However, rather than formalize truth *in* a world and truth *at* a world, for present purposes, I am going to use a very informal, figurative device: for each logically possible world, there will be a metaphysical layer that in some way contains truths in green and falsehoods in red. These are, perhaps, Buridanal worlds with stratospheres of cloudy truths and falsehoods, respectively green and red, where, figuratively speaking now, the inhabitants of some of the planets watch Logistica Borealis in the sky in the evenings (play cricket and sing folk songs). I will use such imaginative scenarios in the next section to illustrate some consequences of the Pinocchio paradox in hierarchical semantics, then in the third section I will provide formal proofs of these consequences.

To further support the view that what is the case and what is true come apart sometimes, here is a version of the Liar paradox where the premises are motivated

[9] There are other principles of truth. There are compositional principles of interpretation, which if they are analytic, assure rules such as:

'A & B' is true iff 'A' is true and 'B' is true.
'~A' is true iff 'A' is not true

On these principles as axioms, see. Volker Halbach (2014). Note, nevertheless, that these principles of interpretation do not assure the T-schema.

by the thoughts that (7) it is not possible for something to be the case and not be true, and (8) it is not possible for something to be true and not be the case.[10]

(7) ~(A & ~T < A>) Premise
(8) ~(T < A > & ~A) Premise
(9) Let A be a Liar sentence *a* = <~T*a*>
(10) ~(~T*a* & ~T < ~T*a* >) (7) [A/~T*a*] the result of uniformly replacing all occurrences of A in (1) by ~T*a*
(11) ~(~T*a* & ~T*a*) (10), (9) = E
(12) ~~T*a* (11), '~(P & P)' entails '~P'
(13) ~(T < ~T*a* > & ~~T*a*) (8) [A/~T*a*]
(14) ~(T*a* & ~~T*a*) (11), (9) = E
(15) ~(T*a* & T*a*) (14) Double Negation
(16) ~T*a* (15) '~(P & P)' entails '~P'
(17) ~T*a* & ~~T*a* (16), (12) &I; Contradiction#

The logic on which this contradiction relies is fairly minimal: substitution of identicals (=E), that '~(P & P)' entails '~P', and Double Negation Elimination. If these rules are valid, then the premises are brought into question. If one of the premises is not true, then sometimes truth and what is the case come apart.

As per the sophism I adapted from one of Buridan's, in some possible worlds, some sentences are the case but untrue. I take this to support my view that (7) does not hold of Liar sentences, among others. That is, in general, Liar sentences are the case but are not true. I believe this would be broadly in line with the kind of approach advocated by Buridan (Hughes 1982).

I hope this contributes something towards the semantics of Liar sentences, but the Liar has, I think, a complex cause. As a referee of this journal astutely points out, the question emerges: Does this mean that the contradiction (6) 'Pinocchio's nose grows and Pinocchio's nose does not grow' is the case but not true? More is required to address another aspect of the Liar paradox, Liar arguments. (10) is untrue; however, without a modified logic, "revenge paradoxes" will regain a contradiction. Just so, I think that given a possible world in which the Pinocchio scenario holds, Pinocchio's nose is growing but 'Pinocchio's nose is growing' is not true. Nevertheless, it will require further argument to avoid inferring a contradiction, at least in what is the case. I would make the following observations though. Unless (2) is a necessary truth, then the Pinocchio sentence has the same semantic value as other Liar sentences in a possible world in which the Pinocchio scenario holds. This, as I show in subsequent sections, leaves hierarchical approaches exposed to the Pinocchio paradox, whereas those approaches claim to be contradiction free. Other approaches aim to restrict true contradictions to semantic dialetheia, but I argue that the Pinocchio paradox still forces a gross physical dialetheia. In partial contrast, on my Buridanal approach, if the Pinocchio sentence is the case but is untrue, the biconditional (4) is false. So, (4) is either a dialetheia or simply false. Given that (2) is not a necessary truth, there is a possible world in

[10] This derivation of the Liar may usefully be compared with Heck 2012; although Heck's Liar is premised on . ~(A & T < ~A>) and ~(~A & ~T < ~A>). These premises might be justified by the thoughts that nothing is the case and is false, and nothing is not the case but is not false. The advantage of Heck's Liar is that it does not require double negation.

which the Pinocchio scenario holds and therein, if (4) is simply false, then (2) is untrue, in which case it is unsound to infer (4), (5), or (6) from (2) and (3). I think (4) is simply false in such a possible world because the Pinocchio sentence is untrue – that makes the left-hand-side of (4) simply false while its right-hand-side is true. So, (2) is simply untrue in such a possible world, and the contradiction (6) does not follow soundly from (2) and (3). If, however, some revenge argument shows that (4), (5) or (6) is nevertheless a dialetheia, a paraconsistent logic is still required. I grant that there is not much to choose between the Pinocchio sentence being the case but being untrue, and the Pinocchio sentence being a dialetheia; my semantics for Liar sentences is only a hair-breadth from dialetheia.[11]

3 Pinocchio's Logical Adventures

3.1 Pinocchio's Previous Logical Adventures

Pinocchio has had a number of logical adventures across the logical pluriverse (Eldridge-Smith 2011; Eldridge-Smith 2012b). The following is a synopsis of his logical adventures so far. Using his ability to face the truth, Pinocchio became a politician at a parliamentary congress of possible world counterparts. He was beguiled however by a Dialetheic Lord into asserting the Pinocchio sentence, causing logical explosion. His parliament dissolved into triviality and he became a denizen of an impossible world. In another possible world, Pinocchio appeared in a dream to the 29th incarnation of Chuang Tzu as an avatar. In that dream, Pinocchio related how his own possible world had logically exploded into triviality when he asserted the Pinocchio statement due to the Pinocchio paradox. Perplexed about avatars, Chuang Tzu asked the fourth incarnation of Dr. Pangloss about the structure of the logical pluriverse. Dr. Pangloss explained that just as each possible world in the logical universe is intended to represent a way in which the world might be, each logical universe in the logical pluriverse is intended to represent a way in which logic might conceivably be – for our purposes, particularly with respect to truth principles. Also, just as reincarnation is an identity relation thought to conserve lives in the world, counterparts conserve identities across possible worlds, and avatars conserve conceivable identities across logical universes. In any possible world in which Pinocchio has an avatar or counterpart, the Pinocchio principle holds. Always the optimist, Dr. Pangloss assured Chuang Tzu that Pinocchio would never escape from his Trivial world to propagate his paradox to another. However, Pinocchio mastered logical existence, and managed to access his avatars in other logical universes. (I haven't quite got my mind around how he does this, but it is something like discussing logical pluralism.) In each of his avatars, Pinocchio is usually heroic in the classical model. For example, in one logical universe his avatar is Horatio defending by way of the bridge, one of the Latin bridges anyway. Pinocchio is not particularly concerned whether the biconditional in his principle or the T-schema is material or not; he is more concerned with whether or

[11] However, I think paraconsistent logics are weak logics. One wants a strong non-ampliative alethic logic that supports maximal expressiveness. That is why one wants a strong deductive logic that can use its own truth predicate. But whether my semantics works with a stronger logic than dialetheism is a question that depends on further argument addressing other inferential aspects of the Liar paradox, and a question that is beyond the scope of this article.

not they are necessary principles. He does, however, in a sequel have an identity crisis, but today is concerned with his semantic adventures. Here are four of them.

3.2 Pinocchio on Twi-Nerf

Twi-nerf is like our world except that the Pinocchio principle holds. In all cases where an instance of the T-schema is inconsistent with the Pinocchio principle, the Pinocchio principle takes precedence.[12] So that when Pinocchio says my nose is growing, if it is growing, what he said is untrue (we have a statement that is untrue but is the case); and if it is not growing, what he said is true (and we have a statement that is true but is not the case).

3.3 Pinocchio in Tarski's Hierarchy

Pinocchio pops into existence in a Tarskian universe. He and Tarski are the only multi-logical-dimensional gods there; they have a god's eye view of this logical universe that consists of a hierarchy of worlds. Tarski explains that in the base world things are the case but there are no truths. Worlds are related by the hierarchical T-schema (a version of the Indexed T-schema). So that what is actually the case in the base world is true in the meta-world. As an avatar of Pinocchio is in this universe, the Pinocchio principle also holds. Tarski guides Pinocchio to the base world. Eager to express himself, Pinocchio asserts that his nose is growing. The base world is fine, but the meta- world and all the hierarchy of meta-worlds collapse into triviality.

3.4 Pinocchio in a Kripke Hierarchy

Pinocchio pops into existence in a Kripkean universe. He has a god's eye view, and there to guide him is the ghost of Tarski. This logical universe also consists of a hierarchy of worlds. They are also related by the hierarchical T-schema. Nevertheless, the base world contains many truths and falsehoods. In this logical universe, each world is like a planet with a logical stratosphere. The way in which a world contains truths is for them to be written in the sky as green clouds, and the way falsehoods are contained in a world is for the them to be written in the sky as red clouds. There are sentences in these worlds that are not written in the sky at all, such as Liar sentences; but the Liar sentences themselves have counterparts that are written in the sky in red in the next world in the hierarchy. Nevertheless, as an avatar of Pinocchio is in this universe, the Pinocchio principle holds. The ghost of Tarski guides Pinocchio to the base world. Pinocchio expresses his sentence, and the base world and all the hierarchy of worlds explode into triviality.

3.5 Pinocchio in a Dialetheic World

Guided by the ghost of Tarski, Pinocchio crawls out of the Kripke universe through the navel of some logical beast. Pinocchio emerges in a dialetheic universe, dusting off some navel lint; Tarski does not make it through. Again, Pinocchio is in a world with

[12] This can be thought of along the lines of precedence of operators, so that it is conceivably formally specifiable. Essentially, the semantic value of any statement made by Pinocchio in Twi-nerf is determined by the Pinocchio principle and not the T-schema.

green and red clouds, but it also has yellow clouds, which represent true contradictions (dialetheia). Again, as an avatar of Pinocchio is there in this universe, the Pinocchio principle holds. Pinocchio does not know whether he is in a semantic dialetheic world or a metaphysical one. He asserts his sentence. The dialetheic world either explodes into triviality or at least now contains one gross physical dialetheia.

4 Some Analysis of the Cases Posed in Pinocchio's Logical Adventures

I make some comments as to why each of the hierarchical worlds Pinocchio visited collapsed into triviality through logical Explosion. I also comment on why I think dialetheic worlds are not designed to contain a gross physical contradiction like the Pinocchio contradiction as a dialetheia.

4.1 Pinocchio's Logical Explosion of the Meta-Language Worlds in a Tarskian Universe

Here is a demonstration of the demise of the Tarskian meta-language hierarchy.

For convenience, here is the hierarchical T-schema again in a general form as the Indexed T-schema[13]:

(Indexed T-schema) $T_i(t_\Phi)$ iff Φ

where each instance of the biconditional is an $(i+1)$th level sentence, T_i is a truth predicate, t_Φ is a canonical name for a sentence of the i th level language whose translation into the $(i+1)$th level language is Φ.

The Pinocchio principle is similarly indexed.

(Indexed Pinocchio principle): Pinocchio's nose grows iff what he is saying in the i th level is not true_i

where the whole biconditional is an $(i+1)$th level sentence and 'is true_i' is a truth predicate.

(1) What Pinocchio is saying in the base level = 'Pinocchio's nose is growing' Premise

(2) 'Pinocchio's nose is growing' is true_0 iff Pinocchio's nose is growing (indexed T-schema)

(3) Pinocchio's nose is growing iff 'Pinocchio's nose is growing' is not true_0 (1), (Indexed Pinocchio principle) =E

[13] Let us say that the hierarchical T-schema makes standard assumptions about the range and ordering of the index, i.

(4) ‘Pinocchio’s nose is growing’ is true$_0$ iff ‘Pinocchio’s nose is growing’ is not true$_0$ (2), (3) SL

Alternatively, (5) Pinocchio’s nose is growing iff Pinocchio’s nose is not growing (2), (3) SL (but remember this is a statement in the metalanguage).

This is an argument in the first meta-language. Truth for the object language is defined in the metalanguage and therefore all instances of the Indexed T-schema that use ‘is true$_0$’ are asserted in the metalanguage, in particular (2). Thus, the meta-language contains contradictions (4) and (5); however, the object language does not seem forced to have a contradiction. Nevertheless, the contradiction will spread by logical contagion (i.e. ECQ) in the meta-language, and thence to all the higher-level metalanguages.

4.2 Pinocchio’s Logical Explosion of the Whole Hierarchy of Worlds in a Kripkean Theory of Truth Universe

In a Kripkean logical universe, the meta-language hierarchy suffers the same fate as in the Tarskian universe. Also, the base world suffers from logical Explosion because of the original version of the Pinocchio paradox as set out in section 1, given that the Pinocchio principle holds in this possible world scenario.

4.3 Pinocchio’s Logical Explosion of the Dialetheic World

Basically, the original argument for the Pinocchio paradox still holds. While (4) could reasonably be construed as a semantic dialetheia, (5) cannot – it is about empirical matters concerning Pinocchio’s nose – a gross physical contradiction over whether it is growing or not. Actually, dialetheic logics resist Explosion. Being paraconsistent, they do not support ECQ. However, semantic dialetheism only constrains semantic dialetheia. A semantic dialetheic world is either trivialised by containing a gross physical contradiction, or, at very least, a gross physical contradiction is a counterexample to all contradictions being merely semantic in such a world. Moreover, I think that a gross physical contradiction is not acceptable in most dialetheic theories, as per the discussion in the next subsection.

4.4 Is the T-Schema a Metaphysical Principle?

D’Agostini and Ficara (2016) argue the T-schema is a necessary metaphysical principle. Their argument appeals to truth-makers. Nevertheless, there may be truths that do not have truth makers, whether or not the associated instance of the T-schema is a metaphysical truth.[14] In particular, on their analysis, Liar statements do not have truth makers, and Liar arguments do not prove metaphysical contradictions, but merely semantic contradictions. In this way, a Liar paradox gives rise to semantic dialetheia but not any metaphysical dialetheia. They support Eldridge-Smith’s view that the Pinocchio paradox is a Liar paradox, but on their analysis, that means it only gives rise to semantic but no metaphysical dialetheia.

[14] D’Agostini and Ficara (2016) are not committed to truth maker maximalism, the view that every truth has a truth maker.

A moot point in D'Agostini and Ficara (2016)'s analysis is whether the instance of the T-schema in every Liar argument is a metaphysical truth or merely a semantic truth. Does their claim that the T-schema is a necessary metaphysical principle commit them to every instance of the T-schema being a metaphysical truth for every sentence? Or does it merely commit them to the weaker claim that every instance of the T-schema for a sentence with metaphysical significance is a metaphysical truth? In any case, the instance of the T-schema for the Pinocchio sentence is a metaphysical truth on D'Agostini and Ficara (2016)'s analysis.

The Pinocchio paradox is a variation of the Liar paradox. Nevertheless, the Pinocchio statement is an empirical statement, and it has empirical truth makers, that, as per D'Agostini & Ficara's analysis, will make the relevant instance of the T-schema a metaphysical truth. In the Pinocchio scenario, the Pinocchio principle also holds. And D'Agostini & Ficara agree this will result in a logical contradiction (at least they agree it will result in a semantic one). Nevertheless, I would argue that given that the physics of a logical world will conform to its logic about empirical sentences, there is therefore a gross physical contradiction, namely (5), in such a world (Cf. Gams et al. 2016). That is, given the truth of (5) in that possible world and that (5) entails (6),

(6) Pinocchio's nose is growing and Pinocchio's nose is not growing,

it follows that that is the case in that logical world.

D'Agostini & Ficara's argument, as I understand it, relies on there being facts in the world, whether as another metaphysical layer or simply by supervenience on what is the case. And Graham Priest (1987 /2006), as a non-semantic dialetheists, does appeal to facts. There may well be facts or states of affairs, as well as what is the case. But these, if they play their intrinsic function, will not differ from what is actually the case in that world. So, as I see it, if there are facts in that world, and the facts are such that they make (6) true, then the facts are as (6) says because (6) is the case. In other words, that world contains a gross physical dialetheia. That is, *unless*, we make a distinction between what is the case and what is true, so that we have one sentence that is untrue but is the case – but that would be more like the Twi-nerf world than a dialetheic world.

Appendix: Pinocchio's search for a Truth-teller: the missing Pinocchio hypodox

Eldridge-Smith has conjectured that every paradox of a large class, broader than just the paradoxes of self-reference, has a related hypodox (Eldridge-Smith 2007; Eldridge-Smith, 2012a, 2015: Appendix). A hypodox being a generalization of the Truth-teller phenomenon. As a hypodox, the Truth-teller is either true or not, but there is lack of a principle determining which. The class of paradoxes with dual hypodoxes includes the Liar, Russell's, Grelling's, and their truth-functional variations; it also includes time travel paradoxes.[15] As a variation of the Liar paradox, one might expect the Pinocchio paradox to have a related Pinocchio hypodox. Then a point against considering the Pinocchio paradox as a Liar paradox is that it does not have a corresponding hypodox.

[15] Non-truth functional variations of Curry's paradox do not seem to have related hypodoxes.

Had Pinocchio said “my nose is not growing,” it might either of been the case that his nose is growing or it isn’t. In either case there is no conundrum (given the T-schema). A similar criticism, if it is a criticism, applies to close relatives of the Pinocchio paradox, such as the Blushing liar. If Victoria had said “I am not blushing,” there would be no conundrum. Either she would be blushing or not.[16]

References

Beall, J. C. (2009). *Spandrels of truth*. Oxford: Oxford University Press.

Beall, J. C. (2011). Dialetheists against Pinocchio. *Analysis, 71*(4), 689–691.

Cook, R. T. (2012). The T-schema is not a logical truth. *Analysis, 72*(2), 231–239.

D’Agostini, F., & Ficara, E. (2016). The Blushing Liar. *Proceedings of the International Conference on Paradoxes, Logic and Philosophy*. University of Peking: Beijing pp. 252–257.

Eldridge-Smith, Peter 2004. The Cretan Liar Paradox. *An ABC of Lying: Taking Stock in Interesting Times*. eds: Jan Lloyd-Jones, Patricia and Livio Dobrez. Melbourne: Australian Scholarly: 72–92.

Eldridge-Smith, P. (2007). Paradoxes and hypodoxes of time travel. In J. L. Jones, P. Campbell, P. Wylie, et al. (Eds.), *Art and time* (pp. 172–189). Melbourne: Australian Scholarly http://philpapers.org/rec/ELDPAH

Eldridge-Smith, P. (2011). Pinocchio against the Dialetheists. *Analysis, 71*(2), 306–308.

Eldridge-Smith, P. (2012a). A Hypodox! A Hypodox! A disingenuous Hypodox! *The Reasoner, 6*(7), 118–119.

Eldridge-Smith, P. (2012b). Pinocchio beards the barber. *Analysis, 72*(3), 749–752.

Eldridge-Smith, P. (2015). Two paradoxes of satisfaction. *Mind, 124*(493), 85–119.

Eldridge-Smith, P., & Eldridge-Smith, V. (2010). The Pinocchio paradox. *Analysis, 70*(2), 212–215.

Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.

Gams, M., Cerncic, E., & Montanari, A. (2016). A temporal perspective on the Paradox of Pinocchio's Nose. *Informatica; Ljubljana, 40.3*(Sep 2016), 365–367.

Halbach, V. (2014). *Axiomatic Theories of Truth. Revised Edn*. Cambridge: Cambridge University Press.

Heck Jr, R. G. (2012). A Liar Paradox. *Thought, 1*, 36–40.

Hughes, G. E. (1982). *John Buridan on Self-Reference: Chapter Eight of Buridan's Sophismata. An edition and translation with an introduction, and philosophical commentary*. Cambridge: Cambridge University Press.

Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy, 72*, 690–716.

Luna, L. (2016). Physicalism, truth, and the Pinocchio paradox. *Mind & Matter., 14*(1), 77–86.

Priest, Graham (1987). *In Contradiction*. 2nd edition Oxford: Oxford University Press 2006.

Prior, A. N. (1958). Epimenides the Cretan. *The Journal of Symbolic Logic., 23*(3), 261–266.

Quine, Willard V. O. (1995). Truth, Paradox and Gödel’s Theorem. Selected Logic Papers, second edition. Cambridge: Harvard University Press (1995): 236–41.

Tarski, Alfred (1933). The Concept of Truth in Formalized Languages. *Logic, Semantics and Metamathematics*. Tr. J.H. Woodger. Indianapolis: Hackett (1983): 152–278.

[16] Nevertheless, to the best of my knowledge there is no known counterexample or exception to Eldridge-Smith’s remaining conjecture that each hypodoxical conundrum has a related paradox.