The Philosophical Quarterly Vol. 00, No. 0 *2021*
ISSN 0031-8094 https://doi.org/10.1093/pq/pqab048

THE PROPER FORMULATION OF THE MINIMALIST THEORY OF TRUTH

By Thomas Schindler[1] and Julian J. Schlöder[2]

Minimalism about truth is one of the main contenders for our best theory of truth, but minimalists face the charge of being unable to properly state their theory. Donald Davidson incisively pointed out that minimalists must generalize over occurrences of the same expression placed in two different contexts, which is futile. In order to meet the challenge, Paul Horwich argues that one can nevertheless characterize the axioms of the minimalist theory. Sten Lindström and Tim Button have independently argued that Horwich's attempt to formulate minimalism remains unsuccessful. We show how to properly state Horwich's axioms by appealing to propositional functions that are given by definite descriptions. Both Lindström and Button discuss proposals similar to ours and conclude that they are unsuccessful. Our new suggestion avoids these objections.

Keywords: truth, minimalism, deflationism, propositions, axioms, Horwich.

I. THE MINIMALIST THEORY OF TRUTH

Minimalists about truth claim that the following is all there is to say about what it means for the proposition that snow is white to be true.

(1) The proposition that snow is white is true iff snow is white.

The theory of truth, the minimalist continues, should be some suitable generalisation of (1). It should say about what it means for any given proposition to be true what (1) says that it means for the proposition that snow is white to be true. Donald Davidson (1996: 272–4) argues that it is unclear what such a generalisation would be since 'the same sentence appears twice in [(1)], once after the words "the proposition that", in a context that requires the result to be a singular term, the subject of a predicate, and once as an ordinary sentence.' But if no such generalisation can be found, minimalism about truth cannot be stated, which is 'reason enough to reject' it (also see Blackburn and Simmons 1999).

As a first attempt, one may try to state the theory *schematically* by saying that its axioms are all sentences obtained by replacing the letter '*p*' in (2) by a declarative sentence of English.[1]

(2) The proposition that *p* is true iff *p*.

Although it is often presented in this way, this is not (and cannot be) the minimalist's official theory of truth. There are more propositions than can be expressed by sentences of present-day English. But a theory of truth must explain what it means for *any* proposition to be true (Horwich 1998b: 20). The schema (2) is, at best, an incomplete theory of truth. However, as we will rehearse in Section II, certain key arguments of the minimalist crucially depend on the availability of a complete theory of truth.

Neither can the minimalist express her theory by (3).

(3) For all propositions *x*, *x* is true iff *x*.

This involves *quantification into sentence position*: on the right-hand side of the biconditional in (3), the bound variable '*x*' occurs in the syntactic position of a sentence. But, at least in a typical first-order language, bound variables can only occur as arguments of predicates (i.e., in *object position*). So (3) is ill-formed. Whether *natural* language permits quantification into sentence position is controversial (Künne 2003: ch. 2 and 6; Williamson 2013: ch. 5). However, even if natural language permits some kinds of sentential quantification, it seems safe to say that there is no way of expressing (3) in natural language.

To be sure, the minimalist could invent an artificial language that permits quantification into sentence position. But this would only undermine her own project. For she claims that the purpose of the truth predicate 'is true' is to allow the expression of generalisations like (4).

(4) Everything the pope says is true.

If one could quantify into sentence position in the way required to state (3), one could formalize (4) without using a truth predicate as (5).

(5) $\forall\alpha$, if the pope says α, then α.

The whole point of minimalism is that we have a need to express sentences like (4) and our languages have come to contain a truth predicate to meet this need. As seen in (5), sentential quantifiers would also serve this purpose, but

> in that case there would be required a battery of extra syntactic and semantic rules to govern the new type of quantifier. Therefore, we might consider the value of our concept

[1] Something also needs to be said about the Liar paradox, e.g. that not *all* schematic substitutions for '*p*' in (2) amount to an axiom (Horwich 1998b: 40ff) or that the Liar is not disastrous in the background logic (e.g. Field 2008). Although this is an important issue, it is tangential to our present concerns. We will therefore set it aside.

> of truth to be that it provides, not the only way, but a relatively 'cheap' way of obtaining the problematic generalizations—the way actually chosen in natural language. (Horwich 1998b: 124–5)

As Picollo and Schindler (2018) have argued in a recent paper, the function that minimalism (or other akin forms of deflationism) ascribe to the truth predicate is best understood as a means to simulate higher-order quantification in a first-order framework, i.e., in a framework that allows quantification into object position only. Thus, if the minimalist were to allow quantifiers that can bind variables in sentence position, she would obviate her own reason for being a minimalist. Minimalists cannot phrase their theory as in (3).

The minimalist's woes result from the fact that she needs to generalize over 'snow is white' in (1), but 'snow is white' occurs in (1) once in sentence position and once in object position. As pointed out by Donald Davidson (1996), there is no adequate way to generalize over both occurrences simultaneously. But if the minimalist cannot phrase her theory as (3), what else can she do? Minimalists like Paul Horwich (1998b) bite the bullet and conclude that the minimalist theory (MT) cannot be phrased. However, Horwich continues, it is possible to *characterise* the axioms of MT, which is sufficient. The minimalist can find a function E whose range (i.e., the collection of its outputs) comprises exactly the axioms of MT. Then the axioms of MT can be characterised as follows:

(6) x is an axiom of MT iff there is a y such that y is a proposition and $x = E(y)$.

For instance, E maps the proposition that snow is white to the proposition expressed by (1). Notice that Horwich does not only take propositions rather than sentences as truth bearers, he also takes theories to be collections of propositions rather than sentences. Hence, E is a function from propositions to propositions. The minimalist's task is to state the function E in a way that avoids the issues that prevented her to state her theory outright.

This task is not trivial. Horwich's own attempt to define E makes use of a certain piece of notation, the angle bracket. Lindström (2001) and Button (2014) have independently argued that Horwich's attempt to define E fails due to confusions surrounding this notation. We explain these arguments in Section II. In Section III, we discuss two ways to characterize MT without the angle bracket, one based on possible worlds semantics and one suggested by Lindström. Unfortunately, these formulations are of little use for the minimalist (as Lindström readily admits) because they presuppose a prior concept of truth. In Section IV, we provide a formulation of MT that avoids the defects of previous attempts. Button anticipated a strategy similar to ours and has reason to believe it is futile. However, we will show in Section V that his argument is unsuccessful. We conclude in Section VI that the minimalist is able to phrase her theory in a way that avoids the known problems.

II. HORWICH'S FORMULATION

Horwich's attempt to characterize MT makes use of a piece of notation that, when surrounding a sentence, produces a singular term denoting the proposition expressed by the sentence. Typically, this is written as angle brackets: the singular term '⟨snow is white⟩' *denotes* the proposition *expressed* by the sentence 'snow is white' (Horwich 1998b: 18, fn. 3). That is, ⟨snow is white⟩ is the proposition that snow is white. One may hence think of angle brackets as the formal analogue to the word 'that', as the expression 'that snow is white' denotes the proposition expressed by 'snow is white'.[2]

With this notation, we can easily state some individual axioms of MT. For example, the expression in (7) denotes the proposition expressed by the sentence in (1).

(7) ⟨⟨snow is white⟩ is true $\leftrightarrow$ snow is white⟩

Having conceded that it is impossible to state all the axioms of MT explicitly, Horwich's goal is to find the function that maps a proposition to the corresponding axiom of MT, e.g. it should map ⟨snow is white⟩ to (7). (Again, note that Horwich considers theories to be collections of propositions.)

To find this function, Horwich (1998b: 17–20) avails himself of the angle bracket. He defines the axioms of MT to be all the propositions whose structure is

(E^*) ⟨⟨p⟩ is true iff p⟩.

Now, continues Horwich, 'when applied to any proposition y, this structure (or function)' yields an axiom of the minimal theory. That is, Horwich characterises the axioms of MT as follows.

(8) x is an axiom of MT iff there is a y such that y is a proposition and $x = E^*(y)$.

But this is confused, as has been pointed out independently by Lindström (2001) and Button (2014). The variable 'y' is bound by a quantifier, so it must be a singular term. But E^* cannot be applied to singular terms for propositions; if 'y' is a singular term, '⟨⟨y⟩ is true iff y⟩' is ill-formed. To see this, replace the variable 'y' by another term denoting a proposition, say '⟨snow is white⟩'. The result is

(9) ⟨⟨⟨snow is white⟩⟩ is true iff ⟨snow is white⟩⟩.

On the left-hand side of the biconditional, angle brackets surround a name for a proposition (i.e., a singular term), but they were explained as surrounding

[2] Whether the word 'that' actually functions as a term forming operator in natural language is controversial, but we will bracket that issue and focus our attention on the formal device ⟨.⟩.

sentences; and on its right-hand side, the singular term '⟨snow is white⟩' occurs in sentence position. Either results in ill-formedness.

The only coherent interpretation of (E^*) is as it being a *schema*: if we *substitute* an English sentence for the letter 'p' in (E^*), the result is an expression that denotes a proposition. But the set of instances of (E^*) is not an adequate theory of truth, for this theory would be incomplete.

A theory of truth should apply even to propositions we cannot (yet) express by a sentence. For example, given any real number x, there is the proposition that x is a real number. So there are uncountably many propositions, whereas there can only ever be countably many sentences in a (natural) language. But for a proposition that cannot be expressed by a sentence, there evidently is no sentence one could substitute for 'p' in (E^*) to obtain an axiom that states what it means for that proposition to be true. So MT of truth cannot be stated by defining its axioms as a schema on sentences, since this would amount to an incomplete theory.

It would be a non-starter to say that one substitutes *open formulae* (sentences with free variables, e.g. 'x is a real number') in (E^*) and considers all propositions schematically formed by instantiating free variables. We might believe that there is a true proposition d describing the composition of dark matter. This is an intelligible use of the concept of truth and so the minimalist should admit into her theory an axiom that explains what it is to predicate truth on d. But d might involve yet undeveloped concepts, so it is not an instance of a presently expressible open formula.[3] Considering this proposition d in more detail further elucidates the dilemma the minimalist is in. The minimalist may want to say that (10) is an axiom of MT.

(10) $\langle T(\langle d\rangle) \leftrightarrow d\rangle$.

But 'd' is a singular term, so '$\langle d\rangle$' is ill-formed. This was the problem with Horwich's function E^*. It is equally hopeless to write (11).

(11) $\langle T(d) \leftrightarrow d\rangle$.

Here, '$T(d)$' is well-formed, but the whole biconditional is not, as the singular term 'd' appears in sentence position on its right-hand side.

Finally, for the sake of argument, suppose that $\{.\}$ is a sort of inverse of $\langle.\rangle$. That is, $\{.\}$ takes a singular term for a proposition and produces a sentence

[3] Horwich (1998b: 18, fn. 3) suggests that, in case his theory (8) does not pan out, one should use all instances of (E^*) for sentences in *possible extensions* of English. Conceivably, there is a possible extension of English that comes with the concepts needed to describe dark matter; and for every real number r, one can imagine an extension of English that has a name for r. So this definition has the potential to result in a complete theory of truth, but it seems that it commits us to there being uncountably many possible extensions of English. Button thinks this is hopeless. We are more optimistic about Horwich's prospects here, but as we go on below to phrase a proper version of MT where E takes propositions as arguments, we need not press that point here.

expressing that proposition. We might be tempted to claim that (12) is an axiom of MT.

(12) $\langle T(d) \leftrightarrow \{d\} \rangle$.

But this is ill-formed because $\{d\}$ does not (presently) exist. The curly brackets are supposed to map a proposition to a sentence expressing it. But there is (presently) no sentence that expresses *d*.[4] This was the problem with reading Horwich's E^* as a schema.

The minimalist appears to be out of options. In the axiom stating what it means to predicate truth on *d*, '*d*' occurs once as being subject to a predicate and once as being equivalent to a proposition. These distinct uses of '*d*' cannot be equivocated, so the minimalist must *somehow* make this distinction. Davidson (1996: 274) does 'not see how this can be done' and we have just seen that the angle brackets are of no help (see Button (2014) and Lindström (2001) for further discussion on the angle bracket).

Before we move on, let us be clear why the minimalist needs a *complete* theory of truth in the first place. For, one might wonder, wouldn't it be sufficient to consider (E^*) as an open-ended schema? The problem, as Button (2014: 285–6) observes, is that certain key arguments of the minimalist crucially rely on the availability of a complete theory of truth. For example, consider how the minimalist typically explains general facts about truth like 'true beliefs facilitate success'. Very roughly, start by considering the proposition that snow is white. The idea is that we can observe that if snow is white, then believing that snow is white facilitates success. Using MT, it follows that if it is true that snow is white, then believing that snow is white facilitates success. We can do this for any other proposition as well, so to derive the *general* claim that true beliefs facilitate success, we merely 'need to collect all these conclusions together' (Horwich 1998b: 137).[5] If MT were not complete, then this strategy would only show the value of *some* true beliefs (namely, about propositions covered by MT). But the minimalist claims to have explained the value of true belief *tout court*. To maintain this claim, she needs to present a complete theory.

III. FORMULATIONS PRESUPPOSING TRUTH

Horwich's attempt to characterize the axioms of MT is unsuccessful. The angle brackets are of little help because they can only be used to form terms

[4] Again, one might consider that $\{.\}$ maps a proposition to a sentence in a possible extension of English that expresses *d* (according to the semantics of this possible extension of English). See the previous footnote.

[5] There are some well-known criticisms of this strategy (e.g. Gupta 1993; Raatikainen 2005). We set this aside here, as the matter at hand is whether the minimalist can get her own arguments off the ground—not whether these arguments are successful.

for propositions that are expressible in the actual language. Of course, this does not mean that MT cannot be phrased *at all*.

If, for example, the minimalist makes use of a sufficiently powerful conception of propositions, she is able to describe the axioms of her theory. Many conceive of propositions as sets of possible worlds. Given this conception, we can characterize the axioms of MT as follows (where Tr is the function that sends a proposition x to the proposition that x is true).

(13) x is an axiom of MT iff there is a proposition y such that for all worlds w: $w \in x$ iff ($w \in y$ iff $w \in Tr(y)$).

That is, x is an axiom of MT when there is a proposition y such that x contains exactly those worlds in which y is equivalent to the proposition that y is true. Accepting all axioms of MT would then leave one with all and only those worlds in which all propositions are equivalent to their own truth, which is as desired.

But this gives no succor to the minimalist. If she conceives of propositions as sets of possible worlds, she has taken on certain metaphysical commitments that include an understanding of a proposition's *truth conditions* and what it means for a proposition to be true at a world. The minimalist cannot, of course, buttress her theory of truth on a theory of propositions with substantial commitments about what it means for propositions to be true. It is for that reason that Horwich (1998b) searches for an alternative conception of propositions. (He endorses a use-based theory of propositions, whose details are tangential to the matter at hand.)

A more sophisticated attempt along similar lines is developed by Lindström (2001). He assumes that there is a class P of all propositions whatsoever, be they expressible or not. He further assumes that there is a subclass T of P that contains all and only the true propositions. On the class T, we can define the Boolean operations by their usual recursive definitions. For example, we may stipulate that disjunction (+) and negation (−) satisfy the following laws.

(14) $x + y \in T$ iff $x \in T$ or $y \in T$.
(15) $-x \in T$ iff $x \notin T$.

Following the same idea, we can also define the truth predicate and the material biconditional.

(16) $Tr(x) \in T$ iff $x \in T$.
(17) $x \Leftrightarrow y \in T$ iff ($x \in T$ iff $y \in T$).

From (16) and (17), we immediately obtain, for all propositions $x \in P$.

(18) $(Tr(x) \Leftrightarrow x) \in T$.

Now, we can define MT to be all propositions of the form (18). To wit:

(19) A proposition y is an axiom of MT iff there is an $x \in P$ such that $y = (Tr(x) \Leftrightarrow x)$.

However, definition (19) cannot satisfy the minimalist, as Lindström readily admits (Lindström 2001: 171). The problem is that the minimalist has helped herself to a substantive prior understanding of truth when defining the Boolean operations by recursion on the set of true propositions. It is important to be precise here. The problem of Lindström's approach is *not* that it assumes the (mere) existence of the class T or the existence of the predicate Tr. The minimalist assumes that these exist and aims to state their proper theory. Rather, the problem is that the characterisation of the axioms of MT, (19), presupposes that the class of true propositions, T, satisfies some very specific laws, viz (14)–(17). For instance, the last clause, (17), stipulates that a biconditional proposition $x \Leftrightarrow y$ belongs to the class of *true* propositions iff x belongs to it exactly when y belongs to it. This is naturally interpreted in one of two ways. Either as expressing a truth-theoretic law (i.e., that the set of true propositions is closed under biconditionalising two members and under biconditionalising two non-members) or as a clause fixing the meaning of the biconditional (*à la* Davidson). Both readings are problematic. If (14) is read as a truth-theoretic law, then one should expect it to *follow* from the minimalist's theory of truth, rather than being *required* for phrasing it. If it is a clause intended to fix the meaning of the biconditional, then the meaning of the biconditional depends on a prior conception of truth. This cannot be the minimal theory (19). Since (19) depends on (17), one cannot have (17) depend on (19) on pain of circularity.[6]

IV. THE PROPER FORMULATION

Lindström's attempt to state MT gestures at the right direction: we need to characterize the axioms of MT in terms of our prior understanding of the Boolean connectives. But we must not characterize these connectives by their usual recursive definitions, or algebraically in terms of the class of all true proposition. We will now show how to do so in a way that is germane to the minimalist's needs.

As a warm-up, consider the function NOT that maps a proposition to the proposition that is its negation. For propositions that *can* be expressed by sentences, we can state NOT explicitly by using the angle brackets.

(20) $\mathsf{NOT}(\langle p \rangle) = \langle \neg p \rangle$.

Again, this is a *schema* where the letter 'p' is *not* a singular term, but a placeholder for a declarative sentence of English.

[6] We thank an anonymous reviewer for discussion on this point.

Now, we contend that it is possible to apply the function NOT to propositions that cannot (yet) be expressed. That is, given some proposition x that is currently not expressible by some declarative sentence of English, we can refer to its negation by using the expression 'NOT(x)'. Indeed, when we have a proposition, we can always talk about *its negation* without making anything further explicit. Making this thought more formal, we may define the function NOT for arbitrary propositions by a definite description.

(21) NOT(x) $=_{Df}$ the proposition that is the negation of x.

Using this definite description to refer to negations of potentially inexpressible propositions is in line with typical minimalist commitments about language and propositions. Suppose we are working with a use-theory of meaning and propositions, favoured by Horwich (1998a). From being acquainted with the use of the word 'not' in a range of sentences, we are acquainted with many instances of the principle in (20). From this acquaintance, a use-theorist would maintain, we come to open-endedly accept a *rule* by generalisation.

Open-endedness is an assumption about how we accept and apply certain rules. For example, logicians typically consider our acceptance of inference rules to be open-ended. Here's Vann McGee:

> Our acceptance of the classical rules of inference surely is open-ended. . . . When we introduce a new predicate into the language—say, when we discover a new species or introduce a new product line—we do not have to inquire whether [a rule] remains valid for inferences involving the new predicate. (McGee 2000: 66)

That is, accepting the inferential relations between propositions is not accepting something about any particular class of propositions expressible in a particular language, but about propositions *tout court*. Forming negations is open-ended in this sense. Our acceptance of the rule to form negations does not rest on empirical knowledge about some subset of the propositions we are acquainted with or which we can phrase. It is accepting a rule about propositions *tout court*. This rule in particular allows us to apply 'NOT' to singular terms, including variables 'y' bound by a quantifier '$\exists y$' that ranges over possibly inexpressible propositions. This is analogous to our ability to learn the addition function and apply it to refer to the sum of numbers that we have not encountered before. If 'x' and 'y' are singular terms for real numbers, we can talk about the sum of x and y without actually computing it. Likewise, if 'x' is a singular term for a proposition, we can talk about the negation of x without actually taking a sentence expressing x and inserting 'not'.

All this is just to say that we accept rules about things whose total or potential extents we cannot survey and indeed do not need to survey. This, arguably, includes logical rules, mathematical operations—and negating propositions. The open-ended nature of our acceptance of these rules is distinct from the process of *acquiring* a rule. Above we told a story, germane to minimalism, that this is done by generalisation. The question of *how* such generalisation works is of course a gnarly one (Kripke 1982) and one we cannot answer here. There are other possible stories—e.g. that some rules are *a priori* (Wedgwood 1999)—that face similarly gnarly questions. In any case, we take it to be largely uncontroversial that one can talk about propositions *tout court* by using singular terms and then also talk about their negations.

We cannot, however, state $\mathsf{NOT}(x)$ in functional terms for propositions x that are not expressible by a sentence. That is, we cannot complete the equation '$\mathsf{NOT}(x) = \ldots$' like we could state (20).[7] The angle brackets allow us to state schematic equations like (20), but such equations are not needed to talk about the proposition that is the negation of some proposition (that may or may not be expressed by a sentence).

Now, we can similarly consider for each predicate F the function that maps a first-order object x to the proposition that x is F. For example, since 'is green' is a predicate, we may define the function GREEN that maps an object x to the proposition that x is green.

(22) $\mathsf{GREEN}(x) =_{Df}$ the proposition that x is green.

Note that GREEN is not the same as the predicate 'is green'. '$\mathsf{GREEN}(x)$' is a singular term for a proposition, but 'x is green' is a formula. Neither is GREEN identical to the property of being green, as, again, the former is a function that outputs propositions whereas the latter is something else, depending on one's theory of properties.[8]

Like we defined NOT, we can also define a two-place function BC that, given two propositions x and y, returns their biconditional. And like we defined GREEN, we can consider the function TRUE that maps a proposition x to the proposition that x is true.[9]

(23) $\mathsf{BC}(x, y) =_{Df}$ the proposition that is the biconditional of x and y.

[7] It *would* be possible to do so if we were to commit to certain theories of propositions. If, say, propositions are sets of worlds, then $\mathsf{NOT}(y)$ is the complement of y. Using a recursive, truth-conditional definition of negation would likewise allow a definition of NOT by an explicit equation. But neither minimalists nor ourselves want to commit to such accounts, for reasons given in the previous section.

[8] Note that we have some leeway as to the precise nature of propositions, properties, and individuals, and how they relate to each other. For instance, we have a choice as to whether 'x' ranges over individuals or Fregean senses. If 'x' ranges over individuals, then the proposition $\mathsf{GREEN}(x)$ is a singular or Russellian proposition, having the individual x itself as a constituent; if we assume 'x' ranges over Fregean senses, $\mathsf{GREEN}(x)$ is something like a Fregean thought. Note that, at least according to Horwich, minimalism is compatible with either type of propositions, Russellian and Fregean ones, so we need not decide this matter here (Horwich 1998b: 91). However, as said, Lewisian propositions as sets of possible worlds are incompatible with minimalism.

[9] We contend that predicating truth on a proposition is just like predicating green on a leaf, i.e., that propositions are first-order objects. Some might treat quantification over propositions as, say, second-order quantification. But this permits quantification into sentence position, which the minimalist will simply reject.

(24) $\mathsf{TRUE}(x) =_{Df}$ the proposition that x is true.

For those propositions that *can* be expressed by sentences, we can use angle brackets to make these functions explicit (again, 'p' and 'q' are schematic placeholders for English sentences).

(25) $\mathsf{BC}(\langle p\rangle, \langle q\rangle) = \langle p \leftrightarrow q\rangle$.
(26) $\mathsf{TRUE}(\langle p\rangle) = \langle\langle p\rangle \text{ is true}\rangle$.

But these equations do not (and cannot) define what it means to apply BC and TRUE to propositions that cannot be expressed by sentences. The *general* definition of these functions is by definite description.

Now we can define MT. To wit:[10]

(27) x is an axiom of MT iff there is a y such that y is a proposition and $x = E(y)$, where $E(y) =_{Df} \mathsf{BC}(\mathsf{TRUE}(y), y)$.

This is well-defined, since the functions TRUE and BC take propositions as their arguments and return propositions as their outputs. All occurrences of the singular term 'y' in (27) are in object position. The resulting theory is complete, as for even those propositions y that cannot be expressed by a sentence, we can refer to the corresponding axiom of MT, $E(y)$.

Using definite descriptions to define E allows the minimalist to make the distinction between propositions occurring as subjects to predicates and as being equivalent to (other) propositions. To wit, given some proposition x we can refer to a proposition like $\mathsf{TRUE}(x)$ in which x occurs as an object and we can also refer to a proposition like $\mathsf{BC}(x, y)$ in which x occurs as being equivalent to y. This answers Davidson's (1996) challenge about how to generalize over both types of occurrences. The core observation is that using definite descriptions to define E in (27) eliminates the need to put a singular term in sentence position. This need previously arose when the minimalist attempted to form the biconditional in E. The definition of BC, however, *describes* a biconditional proposition by a description in which only singular terms occur.

Since our E does not make use of the angle brackets and can properly be applied even to propositions not expressible by sentences, Lindström's (2001) and Button's (2014) concerns are resolved as well. It might be objected at this point that, like Lindström, we have illicitly appealed to a prior understanding of truth. Our definition (26) of the function TRUE is given in terms of a definite description containing the word 'true'. But this is not a problem. All that is required

[10] With it being clear now that we can grasp arbitrary propositions formed by logical operations and predication, we could also speak directly of the proposition obtained by 'T-biconditionalising' a proposition directly. We assume, however, that in the order of explanation grasping T-predication and biconditionalisation come first.

to get the function TRUE going is an understanding that 'true' is a predicate that can be applied to propositions. Nothing further about the nature of truth is presupposed by the definite description in (26). Everything regarding what the truth predicate actually is or contributes to a proposition comes from MT. And our definitions of the functions like BC are neither truth-theoretic laws (as they do not involve truth at all) nor definitions of the connectives. All that is required for these functions is an understanding of the logical connectives. This understanding must of course not be grounded in their truth-functions, but the use-theorist has other options, e.g. that the meaning of 'and' is an abstraction of the tendency to accept '*p* and *q*' if and only if one accepts both *p* and *q* (Horwich 1998a: 45).

That is, our formulation of MT rests on the following assumptions. We (i) have an understanding of the logical connectives that does not involve truth and (ii) this understanding includes the open-ended ability to negate propositions, biconditionalize them etc. Both are provided by a use-theory like the one suggested by Horwich.

While the angle brackets were not useful to state E, we can find them precisely where they make sense. Suppose a proposition y is expressible by some sentence, e.g. y is ⟨snow is white⟩. Then we can compute $E(y)$ as follows:

$$\begin{aligned} E(y) &= E(\langle\text{snow is white}\rangle) = \mathsf{BC}(\mathsf{TRUE}(\langle\text{snow is white}\rangle), \langle\text{snow is white}\rangle) \\ &= \mathsf{BC}(\langle\langle\text{snow is white}\rangle \text{ is true}\rangle, \langle\text{snow is white}\rangle) \\ &= \langle\langle\text{snow is white}\rangle \text{ is true} \leftrightarrow \text{ snow is white}\rangle. \end{aligned}$$

This is exactly the desired axiom (7). So, for those propositions that are expressible by a sentence (and only those), the angle brackets can be used to explicitly compute the functions TRUE and BC. But it would be a mistake to use angle brackets to define these functions in general. The same goes for E. For expressible propositions (and only those), E can be explicitly written in angle brackets, but its *general* definition is as in (27). This, we submit, is the proper formulation of the minimalist theory of truth.

V. IS THE PROPER FORMULATION SELF-UNDERMINING?

Button (2014), in his discussion of the minimalist's problem with the angle brackets, anticipated something akin to our response. He claims that the minimalist who makes use of functions like BC to phrase E as in (27) has again undermined herself by implicitly endorsing quantification into sentence position. Someone who endorses functions like BC, Button argues, should also endorse a corresponding function for universal quantification, ALL. He defines it as follows (p. 279, with notation slightly adjusted by us):

(28) when F is a predicate and h_F is a one-place function mapping any x to the proposition that x is F, ALL(h_F) is the proposition that predicates F of everything.

Now, in addition to the function BC that forms biconditionals, we can consider the function C that forms material conditionals.

(29) C(x, y) $=_{Df}$, the proposition that is the conditional with antecedent x and consequent y.

With this, Button continues, we have everything we need to refer to the proposition expressed by 'Everything the pope says is true', but *without* appealing to the concept of truth. Here is Button's argument, as it would go for our formulation of MT. Consider the following one-place function POPE.

(30) POPE(x) $=_{Df}$ the proposition that the pope said x.

This means there is the one-place function C(POPE($\underline{x}$), $\underline{x}$). We use underlined letters '$\underline{x}$' and '$\underline{y}$' to indicate the arguments of a function (where it matters), distinguished from singular terms 'x' and 'y' for propositions. That is, C($\underline{x}$, $\underline{x}$) is a one-place function, C($\underline{x}, \underline{y}$) is a two-place function, and C(x, x) is the proposition returned by C($\underline{x}$, $\underline{x}$) on the input x.

Now, apply ALL to obtain the proposition ALL(C(POPE($\underline{x}$), $\underline{x}$)). Button claims that this proposition is equivalent to the one expressed by 'Everything the pope says is true'. But ALL(C(POPE($\underline{x}$), $\underline{x}$)) does not involve the concept of truth. Hence, he concludes, the minimalist has undermined herself: the minimalist claimed that the truth predicate is needed to form universal generalisations over propositions, but to phrase her own theory has endorsed machinery to obviate that need.

At first glance, it may seem that the minimalist has to concede the point. She herself has claimed that one can form BC(TRUE($\underline{x}$), $\underline{x}$), so it must be fair to form C(POPE($\underline{x}$), $\underline{x}$). She can hardly deny that there is a function such as ALL, as forming universal generalisations is something that we can do just as well as forming biconditionals and there is no reason to think that the description in the definition of ALL does not refer. Finally, it would be a non-starter for the minimalist to insist that the predicate 'is true' is not eliminable from our language (where there is no quantification into sentence position), while accepting Button's demonstration that she has committed herself to accepting that the concept of truth (or the property *being true*) is redundant on the level of propositions. This would amount to a redundancy theory of truth, not minimalism. So it appears that when using propositional functions to characterize the axioms of MT, minimalism is either directly self-undermining or, if the minimalist insists on the non-eliminability of 'is true' on the linguistic level, collapses into the redundancy theory of truth.

The appearance is misleading. Button claims that ALL applied to the function C(POPE($\underline{x}$), $\underline{x}$) returns a proposition equivalent to *everything the pope says is true*. To see that it does not, first consider a simpler case. What do we get when we apply ALL to the function C($\underline{x}$, $\underline{x}$)? If Button is right, we should get a proposition equivalent to the claim that all conditionals with identical antecedent and consequent are true. But this is not the case, as ALL cannot be applied to C($\underline{x}$, $\underline{x}$) at all! The function C($\underline{x}$, $\underline{x}$) is not a one-place function that maps each x to a proposition that predicates something of x. Let x be any proposition. C(x, x) does not *predicate* anything of x because x does not occur in object position (i.e., under a predicate). So one cannot apply ALL to it.

To see this, consider a concrete proposition, like ⟨snow is white⟩ or ⟨siw⟩ for short. When talking about such concrete propositions that are expressed by sentences, we can make the result of applying C explicit. To wit (again, the letters 'p' and 'q' are schematic placeholders for sentences):

(31) $\mathsf{C}(\langle p\rangle, \langle q\rangle) = \langle p \rightarrow q\rangle$.

So we can compute that C(⟨siw⟩, ⟨siw⟩) is ⟨siw → siw⟩. This is not a proposition where ⟨siw⟩ is in object position. Nothing is predicated on ⟨siw⟩.

Now turn to Button's application of ALL. Like above for C, we can make POPE explicit for concrete propositions like ⟨siw⟩.

(32) $\mathsf{POPE}(\langle p\rangle) = \langle \text{pope-says}(\langle p\rangle)\rangle$.

Now compute:

(33) $\mathsf{C}(\mathsf{POPE}(\langle \text{siw}\rangle), \langle \text{siw}\rangle) = \langle \text{pope-says}(\langle \text{siw}\rangle) \rightarrow \text{siw}\rangle$.

This does predicate something on ⟨siw⟩. We can find the predicate F by substituting a free variable 'y' for '⟨siw⟩' in the right-hand side of (33). The result is this:

(F) pope-says(y) → siw.

There remains an occurrence of 'siw' because this occurrence is in sentence position. When we apply a substitution to find the predicate F, we only substitute 'y' for the singular term '⟨siw⟩', but not for the sentence 'siw'. Although we have a predicate, applying ALL to C(POPE($\underline{x}$), $\underline{x}$) is impossible. The consequent of F depends on the particular input ⟨siw⟩. So what is predicated varies with the input. C(POPE(⟨siw⟩), ⟨siw⟩) predicates F on ⟨siw⟩. But, writing 'wiw' for 'water is wet', C(POPE(⟨wiw⟩), ⟨wiw⟩) predicates F^* on ⟨wiw⟩.

(F^*)pope-says(y) → wiw

Thus, there is no determinate predicate F such that C(POPE($\underline{x}$), $\underline{x}$) is a function that maps a proposition x to the proposition that x is F. Different x give different F. Hence one cannot apply ALL to it.

The problem is subtle, vexing—and familiar. In the proposition $\mathsf{C}(\mathsf{POPE}(\langle \text{siw} \rangle), \langle \text{siw} \rangle)$ (i.e., ⟨If the pope said ⟨siw⟩, then siw⟩), the expression 'siw' occurs once in sentence position and once as part of an expression that is in object position: '⟨siw⟩'. Button's plan is to apply ALL to generalize over both occurrences of 'siw'. But this is illicit, as the two occurrences are in different syntactic positions and so cannot be generalised in the same variable. This was Davidson's observation: one cannot generalize a variable in sentence and object positions. To universally generalise, we must ensure that what we generalize over is always in object position. For example, we obtain the proposition that everything the pope says is true by applying ALL to $\mathsf{C}(\mathsf{POPE}(\underline{x}), \mathsf{TRUE}(\underline{x}))$.

This is what the minimalist claimed all along. And she already conceded that she is unable to state a universal generalisation over all outputs of functions like $\mathsf{C}(\mathsf{POPE}(\underline{x}), \underline{x})$. The minimalist's problem with generalising over her function $E(\underline{x})$ was that in the output of E for any x, x appears both in object and in sentence position. Button's argument shows that the minimalist *must* accept that she is unable to form such universal generalisations over propositional functions, if her theory is not to collapse into redundancy. This is why she was forced to state her theory as an infinite collection of axioms instead of a single universal statement.

But minimalism is not yet out of the woods. Yes, one may intervene on Button's behalf, the function ALL applied to $\mathsf{C}(\mathsf{POPE}(\underline{x}), \underline{x})$ does not allow the generalisation that Button had in mind and so does not lead to truth being redundant. But this may merely be a problem with the particular function ALL. There could be another function ALL' that when applied to $\mathsf{C}(\mathsf{POPE}(\underline{x}), \underline{x})$ yields a proposition equivalent to the one expressed by 'everything the pope says is true'. Indeed there is. It can be defined as (34).

(34) $\mathsf{ALL}'(h) =_{Df}$, the proposition that for all propositions y, $h(y)$ is true.

But ALL' uses the truth predicate, so truth is not eliminated when we form $\mathsf{ALL}'(\mathsf{C}(\mathsf{POPE}(\underline{x}), \underline{x}))$.

No variant of ALL will make make truth redundant. Taking a step back from the specifics of how one forms universal generalisations, Button's core idea is that the following two functions are not all that different. They both map a proposition to another one.

(35) $h_1(y) =_{Df} \mathsf{C}(\mathsf{POPE}(y), y)$.
(36) $h_2(y) =_{Df} \mathsf{C}(\mathsf{POPE}(y), \mathsf{TRUE}(y))$.

Either function applied to, say, the proposition that snow is white outputs a proposition equivalent to the one expressed by 'if the pope said that snow is white, then snow is white'. The sole difference between h_1 and h_2 is that h_1 does not mention a truth predicate.

Everyone agrees that one can universally generalize over the outputs of h_2 to obtain the proposition that everything the pope says is true. But, however, it is that one forms this generalisation, it seems surprising that one cannot do the same for h_1 to form a proposition stating that everything the pope says is true, but without involving a truth predicate. Why should one *not* be able to take any function mapping a proposition to another and form an universal generalisation over, roughly put, all its outputs?

The minimalist insists that this is indeed just roughly put. Once one properly unpacks the idea of generalising over 'the outputs' of a propositional function, the truth predicate will be revealed to not be redundant; in fact, its central role in quantifying over propositions will be highlighted. Certainly it *is* possible to universally generalize over all outputs of a propositional function, but the only precise way to do so is to say that all the function's outputs *are true*. This is our function ALL'. It is also possible to universally generalize over a predication function. This is Button's function ALL. Both functions are useful and represent important parts of our conceptual vocabulary, but neither makes truth redundant. No good alternatives are forthcoming.

Perhaps one can insist that there just *is* a generalisation function that when applied to h_1 returns, roughly put, the universal generalisation over all outputs of h_1. It might not be possible to *define* this function without involving the truth predicate or even to non-roughly explain what this function does without using the word 'true'. But such a function may nevertheless *exist* and thereby cause trouble for the minimalist who states her theory using proposition-forming functions. Such a move would need to be supplemented with at least *some* argument as to why the minimalist should concede the existence of such a generalisation function. And any such argument would likely be question begging.

It was always clear that if the truth property is eliminable on the propositional level (even if the predicate 'is true' is not eliminable from language), minimalism collapses into the redundancy theory of truth. If the required uses of 'is true' in language are not replicated on the propositional level, then our function TRUE is just identity, mapping any proposition to itself. But minimalists assume that propositions are the primary truth bearers and that the truth property is not redundant (albeit deflationary) on the propositional level as well. That is, say, the sentences 'snow is white' and '⟨snow is white⟩ is true' have different *meanings*, i.e., express the distinct propositions ⟨snow is white⟩ and TRUE(⟨snow is white⟩), respectively. If TRUE is identity, these propositions are identical and minimalism collapses into redundancy—the MT would just be the set of trivial biconditionals $\mathsf{BC}(y, y)$. Appealing to a generalisation function to establish this is just a detour.

To be sure, one *can* eliminate truth by appealing to certain other predicates. Consider the following.

(37) $\mathsf{ME}(x, y) =_{Df}$, the proposition that x materially entails y.

Applying ALL to the function $\mathsf{ME}(\mathsf{POPE}(\underline{x}), \underline{x})$ outputs a proposition equivalent to the proposition expressed by 'everything the Pope says is true'. But it is not surprising or troubling to the minimalist that one can eliminate truth by helping oneself to the property of material entailment.

The material entailment predicate is interdefinable with the truth predicate. Instead of saying that some proposition is true, we can say that it is materially entailed by some logical truth; and instead of saying that some proposition materially entails another, we can say that either the first is not true or the second is true. So both the truth predicate and the material entailment predicate can be used to fulfill the same needs.[11] Button's challenge was that if truth is eliminable, minimalism is undermined. The undermining challenge seemed troubling because minimalists claim that the truth predicate is *needed* for certain purposes and by eliminating truth one shows that it is not. But to show that truth is eliminable by introducing another device to meet that need does not undermine anything. Eliminating truth by appealing to an equivalent device for disquotation seems to only highlight the minimalist point that one such device is needed.

VI. CONCLUSION

Davidson (1996) argued that if minimalism about truth cannot be stated properly, this is reason enough to reject it. Horwich (1998b) conceded that it might not be possible to state the (infinitely many) axioms of MT, but it is sufficient to characterize them by stating a function E that takes a proposition and returns the corresponding axiom of MT. Lindström (2001) and Button (2014) independently pointed out that Horwich's attempt to state such an E is unsuccessful due to his use of the angle brackets.

We argued that the function E characterising MT of truth can and indeed should be stated without making use of the angle brackets. Instead, we suggest that the minimalist appeal to propositional functions given by definite descriptions. MT can then be characterised as follows.

(38) x is an axiom of MT iff there is a y such that y is a proposition and $x = E(y)$, where $E(y) = \mathsf{BC}(\mathsf{TRUE}(y), y)$.

Button argued that when articulated as (38), MT of truth collapses into the redundancy theory, as the minimalist has allegedly committed herself to a generalisation function that allows one to eliminate the truth predicate altogether. However, there is no definition of such a function that does not itself appeal to the concept of truth, so the argument succeeds only if TRUE is the identity function on propositions. But to assume this begs the question against the minimalist.

[11] We thank an anonymous reviewer for discussion on this point.

For all propositions y expressible by some sentence (i.e., $y = \langle p \rangle$ where 'p' stands for a sentence), $E(y)$ is indeed just $\langle \langle p \rangle$ is true $\leftrightarrow p \rangle$, exactly as the minimalist claimed all along. But she was mistaken to then generalize over this schema as if 'p' were a singular term. Instead, we have argued, she should define E by definite description without assuming anything substantive about the denotations of these descriptions. This, we submit, puts Davidson's worries about generalisation, Button's worries about notational confusions, and Lindström's worries about presupposing truth to rest.

ACKNOWLEDGEMENTS

We are grateful to Tim Button for in-depth discussions of his 2014 paper and our response. We thank two anonymous referees for helpful comments. This work has received funding from the European Commission through a Marie Sklodowska-Curie Individual Fellowship for the project LOFUPRO, The Logical Function of Property Talk (grant agreement number 792202), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program within the projects TRUST, *Truth and Semantics* (grant agreement number 803684), and EXPRESS, *From the Expression of Disagreement to New Foundations for Expressivist Semantics* (grant agreement number 758540).

REFERENCES

Blackburn, S. and Simmons, K. (1999) 'Introduction', in S. Blackburn and K. Simmons (eds) *Truth*, 1–28. Oxford: OUP.
Button, T. (2014) 'The Weight of Truth: Lessons for Minimalists from Russell's Gray's Elegy Argument', in *Proceedings of the Aristotelian Society*, vol. 114, 261–89. Oxford: OUP.
Davidson, D. (1996) 'The Folly of Trying to Define Truth', *The Journal of Philosophy*, 93: 263–78.
Field, H. (2008) *Saving Truth from Paradox*. Oxford, OUP.
Gupta, A. (1993) 'A Critique of Deflationism', *Philosophical Topics*, 21: 57–81.
Horwich, P. (1998a) *Meaning*. Oxford: Clarendon.
———. (1998b) *Truth*. Oxford: Clarendon.
Kripke, S. A. (1982) *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Cambridge, MA: Harvard University Press.
Künne, W. (2003) *Conceptions of Truth*. Oxford: OUP.
Lindström, S. (2001) 'Horwich's Minimalist Conception of Truth: Some Logical Difficulties', *Logic and Logical Philosophy*, 9: 326–51.
McGee, V. (2000) 'Everything', in G. Sher and R. Tieszen (eds) *Between Logic and Intuition*, 54–78. Cambridge: CUP.
Picollo, L. and Schindler, T. (2018) 'Deflationism and the Function of Truth', *Philosophical Perspectives*, 32: 326–51.
Raatikainen, P. (2005) 'On Horwich's Way Out', *Analysis*, 65: 175–7.
Wedgwood, R. (1999) 'The A Priori Rules of Rationality', *Philosophy and Phenomenological Research*, 59: 113–31.
Williamson, T. (2013) *Modal Logic as Metaphysics*. Oxford: OUP.

[1] *University of Bristol, UK*
[2] *University of Amsterdam, The Netherlands*