How (not) to construct worlds with responsibility

FABIO LAMPERT[1] and PEDRO MERLUSSI[2]

[1]Department of Philosophy - University of Greifswald
[2]Center for Logic, Epistemology and the History of Science - University of Campinas

Forthcoming in *Synthese*

Abstract

In a recent article, P. Roger Turner and Justin Capes argue that no one is, or ever was, even partly morally responsible for certain world-indexed truths. Here we present our reasons for thinking that their argument is unsound: It depends on the premise that possible worlds are maximally consistent states of affairs, which is, under plausible assumptions concerning states of affairs, demonstrably false. Our argument to show this is based on Bertrand Russell's original 'paradox of propositions'. We should then opt for a different approach to explain world-indexed truths whose upshot is that we may be (at least partly) morally responsible for some of them. The result to the effect that there are no maximally consistent states of affairs is independently interesting though, since this notion motivates an account of the nature of possible worlds in the metaphysics of modality. We also register in this article, independently of our response to Turner and Capes, and in the spirit of Russell's aforementioned paradox and many other versions thereof, a proof of the claim that there is no set of all true propositions one can render false.

1   Introduction

The *direct argument* for the incompatibility of determinism and moral responsibility, as formulated by van Inwagen (1983: 182-188), makes use of two deduction rules involving the notion of not being morally responsible for the

fact that p, where p is any true proposition.[1] Let NRp abbreviate that p and no one is, or ever has been, even partly morally responsible for the fact that p. The first rule, (A), says that if p is a necessary truth, then p and no one is, or ever has been, even partly morally responsible for the fact that p, that is:

$$\text{(A) } \Box p \vdash NRp.$$

The second rule, (B), says that NR is closed under Modus Ponens for material implication in the following sense:

$$\text{(B) } NR(p \supset q), NRp \vdash NRq.$$

It follows from determinism that the conjunction of the remote past and the laws of nature strictly implies every truth. So where P is a true proposition describing the complete state of the universe at some point in the remote (pre-human) past, L is the conjunction of all the laws of nature, and p is a true proposition about something that happened after P (i.e. during human times), the direct argument can be formulated thus:

1.	$\Box((P \wedge L) \supset p)$	Premise, from Determinism
2.	$\Box(P \supset (L \supset p))$	1, by Modal Exportation
3.	$NR(P \supset (L \supset p))$	2, by (A)
4.	NRP	Premise
5.	$NR(L \supset p)$	3, 4, by (B)
6.	NRL	Premise
7.	NRp	5, 6, by (B)

The argument concludes that if determinism is true, then no one is, or ever has been, even partly morally responsible for the fact that p.[2]

---

[1] In this article we will follow van Inwagen in saying sometimes that an agent is (or is not) morally responsible for *the fact that* p, and sometimes that an agent is (or is not) morally responsible for *the truth that* p. This rather loose way of speaking might suggest a commitment to the metaphysical thesis that facts are just the same as true propositions. But, to be sure, the intended relation here is invariably between agents and propositions. In particular, an agent being (or not) morally responsible for the truth of a proposition. For instance, if a person, say, Stephen, murders someone, then Stephen is (at least partly) morally responsible for the truth of the proposition that he murdered someone. Van Inwagen (1983: 66) is clear about this relation with respect to the consequence argument, and we believe that this is maintained in his formulation of the direct argument. We should also take this opportunity to point out that, whenever the context allows, unquoted expressions such as '$\Box p$' should be taken as referring to the propositions themselves which are expressed by them.

[2] Provided, of course, that no one is, or ever has been, even partly morally responsible for the fact that P—and similarly with respect to L.

A simpler formulation of the direct argument may be offered in the same spirit as that which was proposed by Warfield (1996). Following his suggestion, rules (A) and (B) may be replaced with

$$(\text{Beta } \square) \ \square(p \supset q), NRp \vdash NRq.$$

The argument can then be reformulated as follows:

1. $\square((P \wedge L) \supset p)$ Premise, from Determinism
2. $NR(P \wedge L)$ Premise
3. NRp 1, 2 by (Beta $\square$)

Even though (Beta $\square$) is the only rule that is required for this argument, it is straightforward to verify that, under plausible assumptions, (Beta $\square$) entails (A): If q is a necessary truth, then it is strictly implied by any proposition, including, in particular, any true proposition p such that NRp is true, and so (Beta $\square$) yields the conclusion that q and no one is, or ever has been, even partly morally responsible for the fact that q.

Much has been said about the direct argument and, of course, its counterpart, the *consequence argument* for the incompatibility of determinism and free will. Although the direct argument and the consequence argument are importantly distinct, they are also strongly related, for one is morally responsible for the fact that p only if one has a choice about it. In Lampert and Merlussi (2021) we argue that, in the context of the consequence argument, there are reasons for rejecting rule (Beta 2) below. This is the counterpart of (Beta $\square$) with the sole difference that occurrences of 'NR' are replaced by 'N', where Np abbreviates that p and no one has, or ever had, a choice about whether p, that is:

$$(\text{Beta 2}) \ \square(p \supset q), Np \vdash Nq.$$

The consequence argument can then be stated by simply replacing occurrences of 'NR' in the reformulated direct argument above with 'N', and using (Beta 2) instead of (Beta $\square$) to justify the last step. Obviously, under plausible assumptions, (Beta 2) entails (Alpha), which is the counterpart of (A):

$$(\text{Alpha}) \ \square p \vdash Np.$$

In Lampert and Merlussi (2021) we also argue against (Alpha).

Once upon a time, rule (A) seemed "entirely uncontroversial",[3] "beyond dispute... (and) unquestionably valid",[4] and "(nearly) as trivial and inconsequential as a rule of inference could be".[5] Nowadays, however, things have

---

[3]Stump and Fischer (2000: 53).
[4]Van Inwagen (1983: 184).
[5]Warfield (1996: 218-219).

changed. Kearns (2011) argues that there are several plausible counterexamples to (A) and, therefore, to (Beta □) as well. In Lampert and Merlussi (2021) we argue that one of his counterexamples to (A) is also a counterexample to (Alpha) and, therefore, to (Beta 2), both of which have also been deemed uncontroversial rules.[6] However, Turner and Capes (2018) have recently responded to the counterexamples proposed by Kearns, including the counterexample endorsed in Lampert and Merlussi (2021) in the context of the consequence argument. Given the relation between the direct and the consequence argument, and given the similarities between (A) and (Alpha), Turner and Capes' defense of (A) immediately provides the incompatibilist with a defense of (Alpha) against the counterexample presented in Lampert and Merlussi (2021). The question which therefore faces us concerns the persuasiveness of Turner and Capes' defense of rule (A).

The main purpose of this article is to argue that there are good reasons to think the counterexample to both (A) and (Alpha) still stands in light of the defense provided by Turner and Capes. In §2 we explain the counterexample, and in §3 we discuss Turner and Capes' response to it. We argue, first, that there is an orthodox and well-motivated conception of what it takes for a proposition to be true at a possible world that helps to motivate the counterexample in question. We believe this conception is preferable to the alternative endorsed by Turner and Capes, and we think that this conception is generally correct. In §4 we argue, additionally, that Turner and Capes' defense of (A) is unpersuasive since it also assumes a particular account of the nature of possible worlds that is demonstrably false given reasonable premises concerning the nature of states of affairs. As will become clear in due course, our argument will be quite similar to well-known versions of the Russellian 'paradox of propositions', which we will here take as a *reductio ad absurdum* of the conception of possible worlds advocated by Turner and Capes. Also registered here, in the appendix, and independently of our response to Turner and Capes, is the fact that a Russellian 'paradox' can be achieved with respect to a certain notion that is central to the free will problem, to wit, the notion of *rendering a proposition false*.

---

[6]See, for instance, Finch and Warfield (1998: 517) with respect to rule (Alpha). The present authors originally thought that they had discovered such a counterexample, only to find out, while Lampert and Merlussi (2021) was under review, that Kearns had already given it in print many years ago amongst his other interesting arguments against (A).

2 Responsibility and choice about necessities

Kearns' counterexample to (A), which he calls *Murder!*, makes use of the standard rigidified sense of 'actually'. Suppose that Stephen murders someone and is, consequently, morally responsible for the truth that he murdered someone. Stephen is thereby morally responsible, too, for the truth that he actually murdered someone. But given the standard rigidified sense of 'actually', if it is true that Stephen actually murdered someone, then it is a necessary truth that Stephen actually murdered someone. Therefore, Stephen is (at least partly) morally responsible for a necessary truth, contrary to what (A) says.

It is worth laying out the argument in some formal detail. Let Rp abbreviate the claim that either p is false or someone is, at some time, at least partly morally responsible for the fact that p. That is, $Rp \equiv \neg NRp$, where '$\equiv$' is the material biconditional. Given the nature of the case under consideration, we also need the actuality operator, written here as '@'. The argument starts with the assumption that Stephen murders someone and is thus morally responsible for this, that is:

(1) $\exists x Msx$,

from which it follows that

(2) $R\exists x Msx$.

But (1) is true at the actual world if and only if it is *actually* true, and so (3) follows from (1):

(3) $@\exists x Msx$.

The argument depends, furthermore, on the claim that Stephen *makes it the case* that he actually murders someone. That is, "The actual world is the way it is in part because of what Stephen does".[7] It is (at least partly) because snow is white, for example, that 'Snow is white' expresses a true proposition. Likewise, it is (at least partly) because Stephen actually murders someone that 'Stephen actually murders someone' is true. If that is so, Stephen should also be held morally responsible for the truth of (3), in which case (4) is true:

(4) $R@\exists x Msx$.

---

[7]Kearns (2011: 309-310).

Now, the formula $@p \supset \Box@p$ is valid in any normal modal logic extending the minimal modal logic **K** where the basic modal language is augmented with the actuality operator.[8] And this implication remains valid for every formula in the standard first-order extension of this language. So, in particular, it will be true that

(5) $@\exists x Msx \supset \Box@\exists x Msx$,

and so (6) can be deduced from (3) and (5) by Modus Ponens:

(6) $\Box@\exists x Msx$.

But, given (6), it follows by (A) that

(7) $NR@\exists x Msx$,

which is impossible by the definition of R and (4). Kearns concludes that (A) is what is at fault here, since he believes that (4) is true. We do so as well. As it is claimed in Lampert and Merlussi (2021), a similar argument can be developed against (Alpha): Replace occurrences of 'NR' in the above argument with 'N', and justify step (7) of the new argument with (Alpha) instead of (A).[9]

In Lampert and Merlussi (2021) we anticipate and reply to a number of objections, and it is unnecessary to repeat those arguments here. Yet, it is worth mentioning that we also formulate a counterexample to (Alpha) by means of rigidified descriptions. For example, the proposition that the murderer of Warren Street is Emmanuel Barthélemy is a contingent truth that was up to Emmanuel Barthélemy. But the proposition that the actual murderer of Warren Street is Emmanuel Barthélemy is a necessary truth that was also up to Emmanuel Barthélemy.

---

[8]See Crossley and Humberstone (1977) for an axiomatization of an **S5** modal logic with @. Note that $@p \equiv \Box@p$ (in fact, $\Box(@p \equiv \Box@p)$) will be valid in any normal modal logic extending the system **D**, i.e. **K** $+ \Box p \supset \Diamond p$, alongside the usual axioms for the actuality operator. This equivalence is, however, not valid in the system **K** with the addition of the actuality operator (and its corresponding axioms), for a counterexample can be provided by considering a relational model $M = (W, w^*, \mathcal{R}, V)$ consisting of a unique dead-end (actual) world, that is, $W = \{w^*\}$, where w^* is the designated *actual* world of the model, $(w^*, w^*) \notin \mathcal{R}$, that is, no worlds are $\mathcal{R}$-related, and $w^* \notin V(p)$. In this model, any formula prefixed with a box will be vacuously true at w^*. So, in particular, $\Box@p$ will be vacuously true at w^*, even though $@p$ will be false at w^*, and so $\Box@p \supset @p$ will also be false at w^*. Of course, the left-to-right direction of the biconditional, i.e. $@p \supset \Box@p$, still holds.

[9]The example in Lampert and Merlussi (2021) was formulated differently but, of course, Kearns' original case can be simply adapted to our purposes.

So it seems that such deduction rules relating necessity to lack of choice or moral responsibility collide with certain portions of modal orthodoxy, namely, the standard semantics for 'actually' and rigidified descriptions. This is not good news for incompatibilists, for they need such rules if determinism is to be understood as implying the necessary truth playing the role of the first premise in their arguments, for any true proposition p. And it seems to many that such is the correct understanding of the notion of determinism.

But not all is yet lost for the incompatibilist. Before turning to Turner and Capes' response we must recall that in Lampert and Merlussi (2021) we also offer a response to the counterexample on behalf of the incompatibilist that was *consistent* with one having choices about some necessary truths. This was done as follows.[10] Pruss (2013) suggests a formal version of the 'no-choice-about' operator N and proves that (Beta 2) is valid in the Lewisian semantics for counterfactuals and the modal logic **T**. His formal version of N is as follows. Where '$\Box\!\!\rightarrow$' is the counterfactual conditional, 'x' is a variable ranging over human agents, and 'a' is a variable ranging over action types:

$$Np := p \wedge \neg\exists x \exists \mathsf{a}[\mathrm{Can}(x,\mathsf{a}) \wedge (\mathrm{Does}(x,\mathsf{a})\Box\!\!\rightarrow\neg p)]$$

This formulation tries to capture the 'counterfactual sufficiency' interpretation of N, according to which there is no choice about p just in case p is true and there is nothing anyone could have ever done such that, if one were to do it, p would be false.[11] It follows immediately from this definition of N, given the semantics accepted by Pruss, that there is no choice about necessary truths, simply because there is no counterfactual world at which a necessary truth is false. In particular, the inference from @p to N@p will be valid: Since @$p \supset \Box$@p is true at every possible world, if @p is true at a world w, $\Box$@p will also be true at w, and so the right conjunct of N@p will also hold at w, as there is no possible world according to which @p would be false. This result clashes with the intuition that some necessities—in particular, some necessities involving 'actually'—may be up to us.

Our suggestion on behalf of the incompatibilist was to revise the no-choice-about operator to the effect of leaving open the possibility that some necessary truths are up to us. The incompatibilist could bite the bullet on the 'actually' cases as long as some of the central notions in the consequence argument are understood a bit differently—in a way, we believe, that may still be helpful

---

[10]The rest of this section can be skipped without loss if the reader is only interested in our response to Turner and Capes.

[11]See also Huemer (2000). Pruss also proves the validity of a corresponding rule of inference for the *might* version of N. The distinction between these conditionals has no bearing on the 'actually' case, however.

for incompatibilists. Instead of the counterfactual conditional with the usual Lewisian semantics involving a similarity relation between possible worlds, we have suggested using a novel conditional called *counteractual*, written as '$\blacksquare\!\!\rightarrow$', in order to define a novel no-choice-about operator. The semantic entry for this conditional uses double-indexing. It is a two-dimensional restatement of the Lewisian entry for the counterfactual conditional in the following sense:[12]

> $p\,\blacksquare\!\!\rightarrow q$ is true at a possible world w relative to a world v taken as actual if and only if at the most similar possible worlds z to w at which p is true relative to z taken as actual, q is true at z relative to z taken as actual.

So not only one varies the 'counterfactual' world (or world of evaluation) with the $\blacksquare\!\!\rightarrow$ conditional: One also varies the actual world. We then suggested the following definition of a no-choice-about operator notated as '$\mathbf{N}$':

$$\mathbf{N}p := p \wedge \neg\exists x\exists\mathsf{a}[\mathrm{Can}(x,\mathsf{a}) \wedge (\mathrm{Does}(x,\mathsf{a})\blacksquare\!\!\rightarrow\neg p)]$$

The advantage of $\mathbf{N}$, apart from the fact that it is very similar to N, is that it is consistent with some actual truths being up to us: The semantics does not validate the inference from $@p$ to $\mathbf{N}@p$. To see this, suppose that $@p$ is true at a world w relative to a world v taken as actual. That is, let $@p$ be true at the pair (w,v). Then $\square @p$ is also true at (w,v), since $@p \supset \square @p$ will be valid too in this semantics.[13] Yet, this time $\mathbf{N}@p$ will not follow. To see this intuitively, suppose that $@p$ is the proposition that Stephen actually murders someone. At the most similar worlds to this one considered as actual at which Stephen does not murder anyone, it is false that Stephen *actually* murders someone. This is so because the proposition that Stephen actually murders someone will be evaluated relative to *different worlds taken as actual*, that is, worlds at which he does not murder anyone. That is, even if $@p$ is true at a pair (w,v), which is sufficient for $\square @p$ to hold there as well, this does not make it so that $@p$ is true at a pair, say, (z,z), where $z \neq v$. For if z is a world close to w where, when considered as actual, Stephen does not murder anyone, then $\neg p$ is true at the pair (z,z). But $@\neg p$, and hence $\neg@p$, will then be true at (z,z) as well. To sum up: Stephen can have a choice about some world-indexed truths if we understand the relevant notion of *having a choice*

---

[12]In Lampert and Merlussi (2021) definitions like this one were given model-theoretically, with the aid of an accessibility relation between worlds. We refer the reader to Lampert and Merlussi (2021) for more details.

[13]The necessity operator $\square$ is only allowed to modify the counterfactual world, that is, the first coordinate of a pair (w,v), just as in the usual modal semantics where formulas are evaluated relative to a single parameter w.

about a proposition counteractually, instead of counterfactually. Needless to say, this can be adapted to the case originally developed by Kearns involving actual truths for which we are morally responsible.

This is, on the face of it, a plausible resolution for those compelled by the claim that we have choices and are morally responsible for some world-indexed truths, while being at the same time inclined to accept incompatibilist views. For in Lampert and Merlussi (2021) it is shown that a novel Beta-type rule can be formulated with $\mathbf{N}$, besides a novel consequence argument that is demonstrably valid in the semantics for counteractual conditionals. This new Beta-type rule was formulated as follows:

$$\text{(Beta 3)} \quad \blacksquare(p \supset q), \mathbf{N}p \vdash \mathbf{N}q,$$

where the operator written as '$\blacksquare$' quantifies over all 'diagonal' points (or pairs of possible worlds) of a model: $\blacksquare p$ is true at (w, v) if and only if for every world z, p is true at (z, z). This notion of necessity is not strange to philosophical semantics. The operator $\blacksquare$ receives the same semantic entry as the deep necessity operator *fixedly actually* defined by Davies and Humberstone (1980), although, in their formal language, this was not a single operator, but the result of prefixing @ with their *fixedly* operator. It was shown, additionally, that $\blacksquare$ and $\square$ agree about the truth values of all formulas in their respective scopes except possibly for formulas involving @, which motivated the authors to run a reformulated consequence argument as follows:

1. $\blacksquare((P \wedge L) \supset p)$ Premise, from Determinism*
2. $\mathbf{N}(P \wedge L)$ Premise
3. $\mathbf{N}p$ 1, 2 from (Beta 3)

As $\blacksquare$ and $\square$ are 'almost' equivalent relative to the class of models for this enriched two-dimensional modal language, we thought the same considerations leading to the first premise under the usual notion of determinism would underpin its two-dimensional counterpart as seen above. However, we have also expressed second thoughts about this solution. We argued that there are reasons to believe that the move to a two-dimensional semantics might come with its own challenges. It turns out that new (though possibly more contrived) counterexamples can be generated in the two-dimensional case that are similar in spirit to the original counterexamples to (Alpha) and (Beta 2). This can be done by defining a new actuality operator in the two-dimensional framework that, as it were, points invariably to the *real* actual world of the model(s)—a world which was demoted from this new semantics in place of a second parameter of evaluation for formulas. We omit the relevant details here, which

the reader can find in Lampert and Merlussi (2021). But it is worth pointing out that if those new cases are plausible counterexamples, they affect once more the intuitive principle that necessary truths are not up to us, which is encapsulated by (Alpha) and, mutatis mutandis, by (A) as well.

3 The Turner and Capes defense of rule (A)

In their response to Kearns, Turner and Capes use 'α' to designate the actual world, and consider the world-indexed truth that Stephen murders someone in α instead of that Stephen actually murders someone. They say the following:

> We deny that Stephen makes it the case that he murders some-one in α, for the following reason. Possible worlds are maximally consistent states of affairs. α is thus a maximally consistent state of affairs a defining feature of which is that it has the state of affairs 'Stephen murders someone' as a component. We contend that what makes world-indexed truths of the form 'X happens in world w' true is the fact that the state of affairs 'X happens' is compossible with the conjunction of all the other states of affairs that make up w. What makes 'Stephen murders someone in α' true on this account is the fact that the state of affairs 'Stephen murders someone' is compossible with the conjunction of the other states of affairs of which the maximally consistent state of affairs α is composed. Notice, though, that Stephen doesn't make it the case that 'Stephen murders someone' is compossible with the other states of affairs of which α is composed. Which states of affairs are consistent with which is not determined, even in part, by human behavior. Hence, Stephen doesn't make it the case that he murders someone in α. (2018: 585)

Turner and Capes say that the sentence 'Stephen murders someone' is a state of affairs, which it is not. We believe what they mean is that the sentence 'Stephen murders someone' expresses a state of affairs or, better, a proposition, which is what is typically taken to be a truth-bearer. States of affairs, by contrast, are usually said to obtain or fail to obtain.[14] We should note too that Turner and Capes do not use the gerund form for states of affairs, as it is

---

[14]It is also common to distinguish between a state of affairs *existing* and *obtaining*. The state of affairs *Stephen's failing to murder someone* does not obtain relative to or in α, but this does not need to mean or entail that it does not exist in it. This distinction between obtaining and existing is important in order to distinguish states of affairs from facts the

commonly done, and so it is not clear whether, for them, states of affairs are really distinct from propositions. Either way, the arguments presented here, in particular, the arguments in the next section, still seem to be applicable.[15]

Now the proposition (or state of affairs) that Stephen murders someone in α is apparently taken to be the same as, or at least equivalent in some sense to, the proposition that Stephen actually murders someone. In fact, the whole argument developed by Turner and Capes is intended as a defense of the claim that Stephen is not even partly morally responsible for the truth of the *former* proposition. (It is worth pointing out that Kearns (2011: 310) himself thinks the propositions in question are equivalent.) But whether these propositions are the same or, in some sense, equivalent, is not entirely obvious, since the latter proposition would be formalized as (3) above, in a first-order modal language with operators, whereas the former proposition would seem to require a richer language supplemented with a binary 'truth-at-a-world' predicate relating formulas to worlds according to which, roughly, $T(p, w)$ would state that p is true at w. But even if we could derive the biconditional $T(p, \alpha) \equiv @p$ or its necessitation in this richer language, this would not mean that $T(p, \alpha)$ and $@p$ are the same proposition, although they would be materially or maybe necessarily equivalent relative to the structure(s) in question. Indeed, since both propositions would presumably be necessarily true if true at all, and so true at every possible world, the interesting question here is, we think, whether they are identical, and not just true at the same worlds.[16] And depending on how fine-grained one takes propositions to be, these propositions could be distinguished, for example, simply on the basis of the different syntactic constructions of the sentences expressing them. For example, King (2007) provides an account of structured propositions according to which, roughly, the syntactic structure of the sentence expressing the proposition constitutes the structure of the proposition, and so difference in syntactic construction is sufficient for difference in proposition expressed. According to this account, '$T(p, \alpha)$' and '$@p$' would express distinct propositions. But, of course, Turner and Capes could reject this very fine-grained view of propositions. They could hold instead that propositions are coarse-grained in the sense of being identi-

---

existence of which makes propositions true. A similar distinction is usually made with respect to propositions being true at a world and existing at a world, but this distinction has no bearing on the arguments presented here.

[15]For more on states of affairs, see §4.

[16]We are just assuming that this is the notion of equivalence Turner and Capes have in mind. It might turn out that it is not. It might be that they are equivalent in the sense of sharing the same class of models (for the language in question). But no criterion of equivalence was provided by the authors.

fied with the sets of possible worlds at which they are true, or with functions from possible worlds to truth-values. Proponents of this view, for example, include Stalnaker (1976: 79-80) and Lewis (1986: 53). According to this view, there is only one necessary truth, i.e. the set of all possible worlds. So the proposition expressed by, say, '$2 + 2 = 4$', is the same as that which is expressed by '$@p$', provided that the latter is true. But this view of the nature of propositions, combined with Turner and Capes' claim about what makes a world-indexed proposition true, would result in the claim that what makes the proposition expressed by '$2 + 2 = 4$' true is the fact that the state of affairs *Stephen's murdering someone* is compossible with the conjunction of the other states of affairs composing α, which seems to be false. Indeed, it does not seem to us that Turner and Capes presuppose this coarse-grained view of the nature of propositions. Perhaps they assume a broadly *Fregean* view of propositions according to which they are composed of Fregean senses, or something in the vicinity. This is a much more fine-grained account of propositions in comparison to the *sets of possible worlds* view. For example, 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' would express different propositions according to this view, even though both propositions are necessarily true. Yet, Fregean propositions are not as fine-grained as propositions construed according to King's account, for instance. While 'Hesperus is Phosphorus' and 'Phosphorus is Hesperus' presumably express the same Fregean proposition, these sentences do not express the same proposition according to King's view, since the constituents at the terminal nodes in these propositions will differ. However, '$T(p, \alpha)$' and '$@p$' may plausibly fail to express the same Fregean proposition. If one overhears a conversation in which a reliable source says that Stephen has actually murdered someone, one will have good reasons to believe that Stephen is actually a murderer. But if all one hears is that Stephen has murdered someone in α, one might need further information to believe that Stephen has actually committed a murder. One might need to know that 'α' is being used as a name for the actual world. More generally, a rational agent could believe that $@p$ without also believing that $T(p, \alpha)$, provided the agent fails to believe that α is the actual world, and so the sentences '$T(p, \alpha)$' and '$@p$' can differ in their cognitive significance, in which case the thoughts or propositions expressed by them can also be different.[17]

---

[17]We are, of course, bypassing complications which may arise from a truly Fregean opinion concerning the propositions in question given Frege's views about modality *and* the truth predicate which, when applied to a sentence 'p', as in "'p' is true', is eliminable, as that sentence expresses nothing over and above what is expressed by 'p'—although this view concerning the truth predicate, as well as Frege's argument for it, would likely not transfer over to the truth-at-a-world predicate unless, at the very least, the world in question is the

It should be noted, however, that there are views according to which '@p' would *express* the proposition that p is true at the actual world α. Soames (2007) defends a view according to which '@p' expresses the *Russellian* proposition that p is true at α, where α is the actual world-state. This would raise questions about the commitment of Turner and Capes (and Kearns, for that matter) to Russellian propositions, although it could provide enough motivation for their claim. Notwithstanding, this account may not be necessary for the argument presented by Turner and Capes. The identification of $T(p, \alpha)$ and @p might turn out to be inessential for their purposes, as Turner and Capes could have said instead that, more generally, what makes the proposition expressed by '@p' true is the fact that the state of affairs corresponding to the proposition expressed by 'p' is compossible with the conjunction of the other states of affairs of which the actual world is composed, where a state of affairs corresponds to a proposition if it is *described* by that proposition. We are not sure if Turner and Capes would accept this revised proposal, although it seems more natural to us in the context of a response to the 'actually' counterexample to (A). Still, there would need to be a plausible account of the relation posited here between propositions and states of affairs. This relation might not be one-to-one, for example. If Stephen is also named 'Bill', then the propositions that Stephen is a murderer and that Bill is a murderer seem to describe a unique state of affairs, namely, *Stephen's being a murderer*. On the other hand, if propositions are Russellian, and names directly refer, these propositions will presumably be identical. Regardless, let us just assume for the sake of argument that $T(p, \alpha)$ and @p are the same proposition, if only to keep changes relative to Turner and Capes' argument to a minimum.

Following our previous remarks, the counterexample involving 'actually' depends on the claim that Stephen *makes it the case* that the proposition that he actually murders someone, or murders someone in α, is true. If so, then it will follow that he is indeed (at least partly) morally responsible for the truth that he murders someone in α, since this proposition will be true (at least partly) because of what he did. Turner and Capes do not seem to deny this.

world of the context according to which the sentence is being said to express the relevant proposition. Additionally, the sentences in question present very different syntactic structures which, given some remarks from Frege (e.g. in his 'Logic in Mathematics' in Frege (1979: 225)), should coincide with difference in senses (see also Frege (1984: 390)). A well-known passage offering a *logical* criterion for the identity of thoughts can be found in a letter from Frege to Husserl written in 1906 (see Frege (1980: 70)). A passage containing Frege's *equipollence* criterion can be found in Frege (1979: 197). The sentences in question would also express distinct propositions according to Carnap's criterion for sameness of intensional structure, since the two sentences are constructed from their sub-sentential expressions in different ways. See Carnap (1956), especially pages 56-59.

What they deny is that Stephen makes it the case that he murders someone in α. Why? One crucial premise is their account of what makes world-indexed truths true, that is:

(P) What makes [the proposition or state of affairs expressed by] 'Stephen murders someone in α' true is the fact that the state of affairs *Stephen's murdering someone* is compossible with the conjunction of the other states of affairs of which the maximally consistent state of affairs α is composed.

And underlying (P) is a certain notion of possible worlds, to wit:

(PW) Possible worlds are maximally consistent states of affairs.

This premise concerning the nature of possible worlds motivates a more general account of truth-at-a-world: The proposition expressed by 'Snow is white' is true at a world w if and only if the state of affairs *snow's being white* is compossible with the conjunction of the other states of affairs of which the maximally consistent state of affairs w is composed. Now the claim is that if one accepts (PW) and (P), thereby agreeing with Turner and Capes on what makes world-indexed truths true, one should also agree that Stephen will not make it the case that he murders someone in α.

Let us attempt to spell out the above argument in a bit more detail, while hoping that our reconstruction remains faithful to what Turner and Capes originally had in mind. Say that states of affairs s and t are compossible if and only if there exists a possible world relative to which both s and t obtain, whereas propositions p and q are compossible if and only if there exists a possible world where both p and q are true. (Similarly for any (possibly infinite) number of states of affairs and propositions.) Premise (P) seems to require, additionally, that there be a conjunction of *all* the states of affairs composing a world *minus* a single state of affairs—in this case, *Stephen's murdering someone.* This in turn seems to call for a certain 'subtraction' operation on worlds according to which, for every possible world w and for every state of affairs s, there is a state of affairs $\bigwedge w^{-s}$ which is the conjunction of all the states of affairs composing w with the exception of s. Turner and Capes provide no argument in favor of this subtraction principle. And, indeed, this principle brings to light a further possible difficulty for this defense of rule (A). For even though they say that a possible world is just a maximally consistent state of affairs, the principle of subtraction just stated seems to suggest that possible

worlds are *composed* of states of affairs *as* conjuncts.[18] And even if possible worlds are maximally consistent states of affairs, it is not obvious why one should take them to be composed in this way. Notwithstanding, the principle of subtraction can be modified to avoid this worry, by saying only that those conjunctions are *included* in the corresponding possible worlds, that is:[19]

(Sub.) For every possible world w and for every state of affairs s, there is a conjunctive state of affairs $\bigwedge w^{-s}$ which is the conjunction of all the states of affairs included in w with the exception of s.

Let m stand for the state of affairs *Stephen's murdering someone*. Then, by (Sub.), the state of affairs $\bigwedge \alpha^{-m}$ exists. (Note that if m were not included in α, $\bigwedge \alpha^{-m}$ would just be α.) Now Turner and Capes' principle (P) can be reformulated as follows:

(9) The compossibility of m and $\bigwedge \alpha^{-m}$ is what makes it the case that $T(\exists x M s x, \alpha)$,

from which they seem to suggest we can draw the following consequence:

(10) Stephen does not make it the case that m and $\bigwedge \alpha^{-m}$ are compossible,

and so:

(11) Stephen is not even partly morally responsible for the fact that $T(\exists x M s x, \alpha)$.

Given the identification of propositions $T(\exists x M s x, \alpha)$ and $@\exists x M s x$, Turner and Capes conclude that Stephen is not even partly morally responsible for the fact (or truth) that $@\exists x M s x$. In fact, they claim that *no one* is responsible for this fact, and so (4), from the argument given by Kearns, is false.

That seems to be the main argument from Turner and Capes in defense of rule (A). Now, is it *sound*? We do not think so. At the very least, we believe that this defense of rule (A) is not persuasive, and in the rest of this article we will lay out our reasons for why we believe so.

First, this defense of rule (A) relies on a certain conception of possible worlds, namely, (PW). But whether the direct argument is valid is a question which is in principle *independent* of idiosyncratic conceptions concerning the

---

[18]Thanks to an anonymous reviewer for pointing this out, as well as for mentioning the need for a subtraction principle of the sort mentioned above.

[19]We say more about what it takes for a possible world to include or preclude a state of affairs in §4.

nature of possible worlds. In §4 we discuss a more detailed view according to which possible worlds are maximally consistent states of affairs. But it must be pointed out that there are multiple views concerning the nature of possible worlds or, more generally, of what type of entity is involved in what makes a proposition true. For instance, one could hold, as Lewis (1986) did, that possible worlds are concrete entities instead of abstract states of affairs, and this view does not seem to require principles such as (P). By contrast, one could hold that propositions are true (false) relative to possibilities (Humberstone (1981)), situations (Barwise and Perry (1983)), or states (Fine (2017a, 2017b)), for example. That is, entities which might sometimes be taken to 'make up' a world but which are not themselves full possible worlds. There is nothing in those views that seem to require an explanation of world-indexed truths along the lines of (P). Of course, whether the 'actually' counterexample to rules (A) and (Alpha) holds in alternative frameworks is also a further issue that will depend on the various semantic entries for the actuality operator. Still, there is also nothing in the traditional semantics for modal logics built upon a set of 'possible worlds' requiring these to be identified intuitively with maximally consistent states of affairs. It would be surprising if the direct argument (or the consequence argument) demanded this, and it would be desirable that it did not.

Second, and apart from this issue, it is not clear how exactly is (10) supposed to follow from (9), if at all. The only thing said in favor of this inference is that "Which states of affairs are consistent with which is not determined, even in part, by human behavior". (2018: 585) And this seems to be considered an a priori truth for Turner and Capes. We confess, however, that whether agents can be responsible for the consistency or compossibility of certain states of affairs is not the kind of thing we have any robust intuitions about. Perhaps the claim is motivated by the assumption that it is a necessary truth whether a certain state of affairs is consistent with another, and therefore no one could be even partly morally responsible for such things, since no one could be even partly morally responsible for necessary truths. But if this is the reason, then it seems to be question-begging, for the claim that no one could be even partly morally responsible for necessary truths is precisely what is at stake here. It might be that there is some such principle in the vicinity that is warranted a priori, and it might be that this principle would justify the inference in question. But we are not sure what principle would serve this purpose, much less if it is warranted or even plausible.[20]

---

[20]As an anonymous reviewer suggested, perhaps the following principle would justify the inference from (9) to (10):

Third, even if the inference from (9) to (10) is warranted a priori and the argument is valid, why should one accept (P) and so (9)? We are persuaded by *Murder!* simply because we think the proposition expressed by 'Stephen murders someone in α' is made true at α (at least partly) because Stephen murders someone in α. The truth of that proposition is said to *depend on the world*. It is because Stephen murders someone in α that the proposition expressed by 'Stephen murders someone in α' comes out true. This seems to be in agreement with the traditional Aristotelian idea that propositions are true or false (at a world w) because of the way the world (w) is:

> This is clear, in the first place, if we define what the true and the false are. To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true; so that he who says of anything that it is, or that it is not, will either say what is true or what is false (*Metaphysics* 1011b, 24-29)

> If there is a man, the statement whereby we say that there is a man is true, and reciprocally—since if the statement whereby we say that there is a man is true, there is a man. And whereas the true statement is in no way the cause of the actual thing's existence, the actual thing does seem in some way the cause of the statement's being true: It is because the actual thing exists or does not that the statement is called true or false. (*Categories* 14b, 15-22)

The addition we make to this Aristotelian idea is the relativization to worlds. We say:

> It is because the thing exists or does not in the actual world that the statement (or proposition) is called true or false at the actual world.

(S) Necessarily, if the obtaining of a state of affairs u makes another state of affairs v obtain, then no individual also makes v obtain.

But (S) faces counterexamples. Suppose that Ed turns on the light by flipping the switch. Presumably, what makes the state of affairs *Ed's turning on the light* obtain is the obtaining of the state of affairs *Ed's flipping the switch*. If (S) holds, then it should follow that no individual makes *Ed's turning on the light* obtain, but it seems it is Ed himself who makes this obtain—at the very least, it seems clear that Ed is at least partly, even if not *morally*, responsible for the obtaining of this state of affairs, or the truth of the corresponding proposition.

And since whether Stephen murders someone in α is, presumably, a matter of his own choice, we think the proposition made true because of this is too a matter of his own choice: By (freely) doing something one (freely) makes a proposition true.[21]

So, why should we reject this explanation in favor of the Turner and Capes account? Unfortunately, the authors provide no argument for (P). What we have here seems to be a conflict between two answers to the question of what makes the world-indexed proposition that Stephen murders someone in α true. And the answer to this question is crucial to determine whether or not Stephen is (at least partly) morally responsible for the truth of that world-indexed proposition. The first, the Aristotelian answer, says that what makes the proposition expressed by 'Stephen murders someone in α' true is Stephen murdering someone in α, which is something that is at least partly up to Stephen. According to this view, truth depends on the world. The truth of a proposition at a world depends on how that world turns out. The second, Turner and Capes' answer, appeals to the compossibility of states of affairs. The truth of a proposition at a world depends on whether a corresponding state of affairs is compossible with the conjunction of every other state of affairs which make up or compose (or are included in) that world, which is in turn motivated by (PW). Which one should we accept?

We think one should favor the Aristotelian view. It is uncommitted with respect to a particular conception of the nature of possible worlds. It is weaker than some versions of the correspondence theory, to wit, those requiring that for every truth there is something corresponding in the world in question. And it is also weaker than some versions of truthmaker theory, to wit, those requiring that for every truth there is a truthmaker in the world in question. But it motivates such views (although they are not necessary for our point), and it does so precisely because it amounts to a trivial fact about truth, namely, that it is contingent upon how the world turns out—at the very least, in a great number of cases. We claim that this is true of certain world-indexed propositions as well: They are true because of the way the world in question turns out; and they are up to someone if their truth depends on what someone freely does.

_____

[21]To be sure, we are not the first to propose this explanation. See Schnieder (2004: 421) for a similar proposal. Additionally, the Aristotelian idea that truth depends on the world—with an application to free will—is discussed in Merricks (2009), who also mentions Aristotle. Claims such as that Stephen murders someone is true because Stephen murders someone (and not vice-versa) are also made in unpacking the premise that every truth has a truthmaker, for example, in Armstrong's (2004: 4) main argument for the existence of facts.

In Lampert and Merlussi (2021) we suggest that the claim that truth depends on the world can be understood via a truthmaker relation, or the positing of some sort of grounding relation. But it is clearly unnecessary to endorse either of these views in order to hold on to the more basic and ecumenical Aristotelian view described here. We aim to be neutral with respect to such views, for we think the idea that the truth of propositions expressed by sentences such as '@p' depends on how the actual world turns out is compatible with many developments of truthmaker theory, correspondence theory, and different understandings of ground.

It might be objected that while the direct argument should not rely on one specific account of the nature of possible worlds, it should not rule one out either. However, while the Aristotelian answer seems unproblematic, we believe things are different with respect to (P). Apart from the issues we have raised here concerning some of the principles and inferences in Turner and Capes' defense of rule (A), and the fact that there is an orthodox Aristotelian account of what makes a proposition true that combines very well with our claim, namely, that one can be (at least partly) responsible for the truth of certain world-indexed propositions, we shall argue in the next section that the Turner and Capes answer is, given plausible assumptions, subject to more serious difficulties. It deserves mention that our critique is not intended in the spirit of hostility. Rather, our point is that the Turner and Capes response (which is the most up-to-date incompatibilist reply to the objection) appeals to a framework that lacks detail in terms of the specific theory of the nature of possible worlds that is assumed. And when we attempt to fill in the gaps, by introducing some principles about states of affairs which seem to us at least as plausible as principles such as (Sub.), we can show that there are no possible worlds in the sense envisaged by (PW), i.e. the theory of possible worlds motivating (P).

4 Are worlds maximally consistent states of affairs?

Turner and Capes do not say much about the account of possible worlds which they appeal to, but what they say about it resounds well-known theories of the nature of possible worlds offered some decades ago by philosophers such as Adams (1974) and Plantinga (1974). According to such theories, possible worlds are abstract entities of some sort.

Adams defines a possible world w as a *maximally consistent set of propositions*. A set S of propositions is *maximal* just in case for every proposition p,

either $p \in S$ or $\neg p \in S$. And a set S of propositions is *consistent* just in case it is possible for all of its elements to be jointly true. So there is a sense in which a possible world 'makes a decision' regarding every proposition's truth value. This is the sense in which we say possible worlds are 'complete', for they leave no information out. Now if one identifies propositions and states of affairs, and some have done just that,[22] the present account would say that possible worlds are maximally consistent sets of states of affairs. But this does not seem to be quite exactly the account that Turner and Capes have in mind, since they do not claim that possible worlds are *sets* of any kind. Rather, they claim that possible worlds are states of affairs which are maximally consistent.

Plantinga, in contrast to Adams, does define a possible world w as a *maximally consistent state of affairs*.[23] According to him, a state of affairs S is *maximal* just in case for every state of affairs s, S either includes s or it precludes s, where a state of affairs S *includes* a state of affairs s if it is not possible for S to obtain while s fails to obtain, and S *precludes* s if it is not possible for both to obtain. And a state of affairs is *consistent* (or *possible*) just in case it is possible in the 'broadly logical sense'. Plantinga (1976: 145) was inclined to reject the identification of propositions and states of affairs, since the former have truth values whereas the latter are said to obtain or fail to obtain, although he does say they are 'intimately related'.[24] States of affairs in Plantinga's sense are things like *Clara's being a philosopher* and *3's being an odd number*; they have complements, that is, *Clara's not being a philosopher* and *3's not being an odd number*; are said to obtain or fail to obtain—in which case their complements might obtain—that is, *3's being an odd number* obtains, but its complement does not; sometimes they include other states of affairs, for example, *David's being an old engineer* includes *David's being old* and *David's being an engineer*; they participate in logical relations with other states of affairs, for example, *David and Clara's being witty* is a conjunction of *David's being witty* and *Clara's being witty*; they bear a close relation to propositions, for example, the state of affairs *Clara's being witty* and the proposition expressed by 'Clara is witty'. Finally, states of affairs exist as primitive entities, that is, they are not defined in terms of other things.

We do not know how much of this is accepted by Turner and Capes. But we think they should hold most of it, as none of those features seems particularly controversial in light of the claim that possible worlds are maximally consistent states of affairs. Yet, such theory of possible worlds faces deep problems, with

---

[22]For instance, Chisholm (1970, 1971).

[23]Plantinga (1974: 44) calls them 'possible' rather than 'consistent'.

[24]Plantinga (1974: 44).

similar issues appearing for Adams' account too, as we note below.

One initial problem with a Plantinga-style theory of possible of worlds is that it is far from obvious why one should believe there are maximally consistent states of affairs in the first place. We might all be believers in possible worlds, whatever they are, but whether there are maximally consistent states of affairs is, on the face of it, a substantially different question. This latter notion mimics more or less the construction of maximally consistent sets of formulas which are constructed out of a formal vocabulary as in a language L for, say, classical propositional logic. A set Γ of L-formulas is said to be maximally consistent just in case it is consistent in the sense of there being at least some L-formula which is not derivable from it, and maximal in the sense of having no consistent proper supersets. Maximally consistent sets of formulas in this sense are useful in proving completeness theorems. For it can be shown by standard arguments that any consistent set of L-formulas can be extended to a maximally consistent set, and that these are always satisfiable, from which it follows that every consistent set of L-formulas is also satisfiable.[25] This is familiar territory for logicians and many philosophers alike. A reason for mentioning this is that the existence of such sets must be demonstrated, since it is not simply evident by their definitions that consistent sets of formulas can be extended to large maximally consistent sets. Analogously, even if one grants the existence of states of affairs, why should one believe that there are maximally consistent such states?

Plantinga (1985: 327-329) acknowledges that the question of the existence of maximally consistent states of affairs is not a trivial one, and he attempts to prove that such entities exist by arguing for the claim that for every state of affairs s, s is possible if and only if there is a possible world in which s obtains. But the argument has well-known issues.[26] In particular, besides the principles mentioned thus far governing states of affairs (i.e. that they exist, have complements, obtain or their complements obtain, etc.) Plantinga appeals to the following:

(Exist.) For every possible state of affairs s, there exists a set S of states of affairs whose members are possible and include s.

(Conj.) For every set S of states of affairs, there is a state of affairs $\&(S)$, i.e. *all of S's members having obtained*, that obtains if and only if every member of S obtains.

---

[25]See Enderton (2001: 131-145) for a standard completeness theorem for *first-order* logic using maximally consistent sets of formulas.

[26]Plantinga's argument is contained in a reply to an article by Pollock (1985), who attempts to show that necessarily, some possible worlds obtain.

Additionally, Plantinga assumes the following principle called 'Quasi-compactness':

(QC) For every set of possible states of affairs S, if S has a maximal linearly ordered subset, then S has a maximal linearly ordered subset S^* which is such that if every finite subset S^{**} of S^* is possible—i.e. such that $\&(S^{**})$ is possible—then so is S^* itself.

Menzel (1988) has shown (QC) to be false. Chihara (1998: 120-126), on the other hand, has shown how one can construct a proof of the existence of at least one maximally consistent state of affairs, that is, a possible world, without using (QC). The problem is that Plantinga's assumptions, including (Exist.) and (Conj.), are then shown to be inconsistent in the sense that a contradiction can be derived from them by using standard set-theoretic principles.[27] We direct the reader to Menzel and Chihara's texts for more details. The main point, however, is this: We have seen no positive and convincing argument for the claim that maximally consistent states of affairs, which are possible worlds according to (PW), exist. And we shall see in what follows that there actually is what seems to be a plausible argument against the existence of such entities.

Grim (1984) and Bringsjord (1985) have independently argued against the existence of a maximally consistent *set* of states of affairs or propositions, although Grim's main target was to argue that there is no set of all truths. The argument is more or less the same in all three cases, be it against a set of all states of affairs, propositions, or truths. In fact, Menzel (2012) has shown how arguments like these can be reformulated under very basic set-theoretic assumptions in a way that is almost identical to Russell's (1903: Appendix B, §500) 'paradox of propositions'. Another way to put it: While Russell thought the argument was a paradox, without explicitly rejecting any of its premises, Menzel has shown it can be taken as an argument to the effect that there is no set of all propositions—or truths, or states of affairs. Such arguments can be formulated as follows. Consider the following premises, directed against the existence of a maximally consistent set S of states of affairs:

(12) There is a maximally consistent set S of states of affairs.

(13) For every $P \subseteq S$ there is a unique state of affairs s_P (say, *P's being a subset of S*).

(14) Every state of affairs s has a complement s'.

---

[27]Jubien (1988: 310-312) attempts different ways of arguing for the existence of a maximally consistent state of affairs, but, as he notes, they all fail as well.

(15) For every $P, Q \subseteq S$, if $s_P = s_Q$, then $P = Q$ (similarly for s'_P and s'_Q).

A contradiction follows from (12)-(15). By Cantor's theorem, where the cardinality of S is k, the cardinality of the power set of S, $\wp(S)$, is 2^k. But it follows from (12)-(14) that for every $P \subseteq S$, either $s_P \in S$ or $s'_P \in S$, and so, by (15), there is a one-to-one function $h : \wp(S) \to S$, which is impossible. This is in essence the argument given by Bringsjord. Grim's argument against a set of all truths is similar, although it does not explicitly need a premise such as (14). Menzel's formulation against a set of all propositions does not use the Power Set Axiom or even Cantor's theorem. As he shows, only the Axiom of Separation is needed. Chihara (1998: 126-127) argues similarly to Bringsjord, showing that at least in the possible worlds theory given by Plantinga (1974, with further assumptions from (1985: 327-329)), the contradiction above seems inescapable. Also, one can simply replace 'states of affairs' in the argument above with 'propositions', and 'complement' with 'negation', and so one has an argument directly against Adams' notion of a possible world.[28] If possible worlds are maximally consistent sets of propositions, then there are no possible worlds. If possible worlds are maximally consistent sets of states of affairs, then there are no possible worlds.

If one accepts the principle that for every state of affairs s there is a set containing all the states of affairs included in s, then there are no possible worlds if these are just maximally consistent states of affairs. This entails a contradiction since Turner and Capes explicitly assume that *there is* the actual world, which is a maximally consistent state of affairs. And if their defense of rule (A) depends on an inconsistent notion of possible worlds, this is a good reason for not accepting it.

The above argument depends on the principle that for every state of affairs s there is a set containing all the states of affairs included in s. And this principle could as well be rejected. Yet, we find this rejection difficult to motivate. Maximally consistent states of affairs are, after all, intended to be totalities of a certain kind, and so why would it be impossible in principle to collect all of the states of affairs—which are ur-elements—included in a maximally consistent state of affairs into a single set?

Still one could take the above argument to show that for at least one kind of state of affairs, namely, a maximally consistent one, the principle that there is a set collecting all (and only) those states of affairs included in it is

---

[28]Note: Plantinga's *books* defined on worlds are explicitly defined as maximally consistent sets of propositions, where each world has its own book. So this version of the argument directly shows that Plantinga's metaphysical assumptions in question lead to a contradiction.

inapplicable. Whether Turner and Capes are committed to this is not obvious to us since the framework they appeal to lacks detail. Notwithstanding the controversy surrounding the existence of a maximally consistent set of states of affairs or, more precisely, whether one should believe that there is such a set *if* one already believes in maximally consistent states of affairs, we shall argue directly against the existence of the latter, without explicitly involving any sets. That is, we shall argue that, given plausible premises concerning states of affairs, the assumption that there is a possible world in the sense of (PW) entails a contradiction.

In order to write the argument in a more precise way, we use individual variables such as 's' and 't' as well as plural variables such as 'ss' and 'tt', since we shall quantify over pluralities of states of affairs, where the expression '$s \prec tt$' means that s is one of the tts.

Consider the following premises:

(16) There is a maximally consistent state of affairs w.

(17) For every plurality of states of affairs ss, if, for every state of affairs t, $t \prec ss$ only if w includes t, then there is a conjunctive state of affairs $\bigwedge ss$ such that (i) for every state of affairs t, $t \prec ss$ if and only if t is a conjunct of $\bigwedge ss$, and (ii) $\bigwedge ss$ is included in w.

(18) For every plurality of states of affairs ss included in w, there is a unique state of affairs $s_{\bigwedge ss}$ which is also included in w (say, $\bigwedge ss$'s *being a state of affairs*).

(19) For every pair of pluralities of states of affairs ss and tt, if $s_{\bigwedge ss} = s_{\bigwedge tt}$, then $\bigwedge ss = \bigwedge tt$.

Given (18), for every plurality of states of affairs ss included in w, either $s_{\bigwedge ss}$ is a conjunct of $\bigwedge ss$ or it is not. Consider the plurality rr of states of affairs such that, for all s, $s \prec rr$ if and only if, for some plurality of states of affairs tt included in w, $s = s_{\bigwedge tt}$ and s is not a conjunct of $\bigwedge tt$. The plurality rr of states of affairs must be included in w, since each s such that $s \prec rr$ is so included. Now we argue that $s_{\bigwedge rr}$ is a conjunct of $\bigwedge rr$ if and only if it is not.[29] If $s_{\bigwedge rr}$ is a conjunct of $\bigwedge rr$, then $s_{\bigwedge rr} \prec rr$, by (17), and so for some plurality tt of states of affairs in w, $s_{\bigwedge rr} = s_{\bigwedge tt}$ and $s_{\bigwedge rr}$ is not a conjunct of $\bigwedge tt$. By (19), it then follows that $\bigwedge rr = \bigwedge tt$, and so $s_{\bigwedge rr}$ is not a conjunct of $\bigwedge rr$. Contradiction. If, on the other hand, $s_{\bigwedge rr}$ is not a conjunct of $\bigwedge rr$, then it is not the case that $s_{\bigwedge rr} \prec rr$, by (17), and so for every plurality tt

---

[29]The argument here is analogous to the one found in Menzel (2012).

of states of affairs included in w, if $s_{\bigwedge rr} = s_{\bigwedge tt}$, then $s_{\bigwedge rr}$ is a conjunct of $\bigwedge tt$. Since the plurality rr of states of affairs is included in w, it follows that $s_{\bigwedge rr}$ is a conjunct of $\bigwedge rr$. Contradiction. Therefore, there is no maximally consistent state of affairs w. That is, (16) is false.

We think the argument above is a strong one for abandoning the maximally consistent states of affairs conception of possible worlds. At the very least, it shows this conception of possible worlds to be problematic apart from its notorious issues with set-theoretic assumptions. So it remains to discuss the premises of the argument. Interestingly, (17) follows from a slightly modified version of (Sub.). Consider the following plural variation of (Sub.), namely:

(P. Sub.) For every possible world w and for every plurality of states of affairs ss, there is a conjunctive state of affairs $\bigwedge w^{-ss}$ which is the conjunction of all the states of affairs included in w with the exception of the sss.

By restricting the plural quantifier to pluralities of states of affairs which are already included in w, (P. Sub.*) then follows from (P. Sub.) under the assumption that any conjunction of states of affairs each of which is included in w is also included in it:

(P. Sub.*) For every possible world w and for every plurality of states of affairs ss, if, for every state of affairs t, $t \prec ss$ only if w includes t, then there is a conjunctive state of affairs $\bigwedge w^{-ss}$ such that (i) $\bigwedge w^{-ss}$ is the conjunction of all the states of affairs included in w with the exception of the sss, and (ii) $\bigwedge w^{-ss}$ is included in w.

Let us first state an explicitly more general version of (17), that is:

(17*) For every possible world w and for every plurality of states of affairs ss, if, for every state of affairs t, $t \prec ss$ only if w includes t, then there is a conjunctive state of affairs $\bigwedge ss$ such that (i) for every state of affairs t, $t \prec ss$ if and only if t is a conjunct of $\bigwedge ss$, and (ii) $\bigwedge ss$ is included in w.

We define the notion of a *plural complement of a plurality relative to a world* as follows:

(Def.) For every possible world w and for every plurality ss of states of affairs included in w, let ss' be that plurality of states of affairs such that for every state of affairs t, $t \prec ss'$ only if w includes t, and $t \prec ss'$ if and only if $t \nprec ss$ (where $t \nprec ss$ if and only if it is not the case that $t \prec ss$).

So for every possible world w and for every plurality ss of states of affairs included in w, the plurality ss' of states of affairs will also be included in w since each one of the ss's is so included. Now let u be a possible world and aa a plurality of states of affairs such that, for every state of affairs t, $t \prec aa$ only if u includes t. Consider the plural complement of the aas relative to u, i.e. aa'. For every state of affairs t, $t \prec aa'$ only if u includes t. Then, by (P. Sub*), there is a conjunctive state of affairs $\bigwedge u^{-(aa')}$ such that (i) $\bigwedge u^{-(aa')}$ is the conjunction of all the states of affairs included in u with the exception of the aa's, and (ii) $\bigwedge u^{-(aa')}$ is included in u. And so for every state of affairs t,

$$t \text{ is a conjunct of } \bigwedge u^{-(aa')} \text{ if and only if } t \not\prec aa' \text{ and } u \text{ includes } t.$$

Now, for every state of affairs t, if $t \prec aa$, then $t \not\prec aa'$ and u includes t, and so t is a conjunct of $\bigwedge u^{-(aa')}$. Conversely, if t is a conjunct of $\bigwedge u^{-(aa')}$, then $t \not\prec aa'$ and u includes t, and so $t \prec aa$. Hence, for every state of affairs t, $t \prec aa$ if and only if t is a conjunct of $\bigwedge u^{-(aa')}$, where $\bigwedge u^{-(aa')}$ is included in u. Since u and aa were chosen arbitrarily, it follows that for every possible world w and for every plurality of states of affairs ss, if, for every state of affairs t, $t \prec ss$ only if w includes t, then there is a conjunctive state of affairs, that is, $\bigwedge w^{-(ss')}$, such that (i) for every state of affairs t, $t \prec ss$ if and only if t is a conjunct of $\bigwedge w^{-(ss')}$, and (ii) $\bigwedge w^{-(ss')}$ is included in w. Therefore, (17*) is true, and so is (17).

What about premises (18) and (19)? We have no arguments that could establish the truth of these premises. Notwithstanding, we believe they are natural assumptions in the metaphysics of states of affairs. They appear to be as reasonable as their set-theoretic counterparts (13) and (15), and so it would be surprising to us if either of them were found to be false. Concerning the present discussion involving the validity of rule (A) (and (Alpha)), one wonders whether there is a satisfying, non-question-begging argument against either of (18) or (19). It seems to us that there is not. Nonetheless, and in all fairness, this also means that what we have here presented might not consist in a knockdown argument against Turner and Capes' defense of rule (A), since we have not shown that the premises leading to the paradox *must* be accepted by them on pain of inconsistency. All the same, we think the present argument suffices to the end of motivating a rejection of (16) and, consequently, of the claim that there are possible worlds in the sense of (PW). At the very least, it yields a coherent and motivated rejection thereof. It is an open question, moreover, whether a theory rejecting either of (18) or (19) would be robust enough as to identify certain kinds of states of affairs with entities such as possible worlds and, of course, whether one would have principled and independent grounds for doing so.

5 Final remarks

We now go back to a question we have asked before. Why should one accept Turner and Capes' account of what makes world-indexed truths true? Perhaps because it has the consequence that Stephen does *not* make it the case that he murders someone in the actual world (thus avoiding Kearns' counterexample to rule (A)), contrary to the Aristotelian view that we have mentioned. However, such an account of what makes world-indexed truths true faces its own difficulties. First, this account is difficult to motivate. It may appear natural under the conception of possible worlds according to which they are maximally consistent states of affairs. But if one accepts that possible worlds are maximally consistent states of affairs together with the principle that for every state of affairs s there is a set containing all the states of affairs included in s, then there are no possible worlds. Second, even if one denies the latter principle, there is still a plausible argument showing that there is no maximally consistent state of affairs that makes use of plural quantifiers, instead of quantifiers over sets. The Aristotelian view, on the other hand, does not suffer from any of these problems. Granted, we have not said what possible worlds are, as the Aristotelian view of what it takes for propositions to be true at possible worlds (or even at different entities altogether) is neutral in that respect—just as the direct argument should be. For the purposes of the semantics and most of the arguments offered in Lampert and Merlussi (2021), possible worlds are just points in a model. For the purposes of metaphysics, they are whichever entity the existence of which ensures that propositions are true or false in virtue of how they are.[30]

Appendix

Before concluding we would like to mention another set-theoretic result to be added to those already demonstrated by Grim, Bringsjord, and Menzel,[31] which is inspired by a central notion in the problem of free will. The claim to be shown is this: *There is no set of all true propositions one can render false.* First we shall explain this notion more carefully, and then we shall go through the argument following Menzel's formulation against a set of all propositions.

Van Inwagen's consequence argument involves essentially the relation of *rendering a proposition false*, which was introduced (see his (1983), especially pages 65-68) with the purpose of investigating "the conceptual relations", as he puts it, between free will and determinism.

> Determinism is a thesis about propositions, but the free will thesis is a thesis about agents. If we are going to investigate the conceptual relations between these two thesis, we shall do well to state the free will thesis as a thesis about agents and propositions. I propose to do this by devising a way to describe our powers to act—and, by acting, to modify the world—as powers over the truth-values of propositions. (65-66)

This relation of rendering a proposition false holds, for example, between agents, such as the authors of this article, and the propositions (expressed by sentences) (a) and (b):

(a) No one has ever read all of van Inwagen's *An Essay on Free Will* aloud, and

(b) This article was written during quarantine.

For what (a) and (b) have in common (in the salient sense) is that whether they are true or false is something that is, or once was, up to us—assuming, of course, that we do have free will. That is, the fact that (a) and (b) are true is something that is, or once was, within our power to change: We have, or once had, a *choice* about whether (a) and (b) are true. By contrast, we never once had a choice about whether

(c) $2 + 2 = 4$, and

(d) Galileo thought he was taller than he actually was

---

[31]With a similar argument, which is attributed to David Kaplan, Davies (1981: Appendix 9) has shown that there is no set of all worlds.

are true. For there is nothing that we could have ever done to the effect of modifying the truth values of (c) and (d). We can make choices, and thereby 'modify the world', to the effect that it is up to us whether some propositions are true, even though many other propositions are true regardless of what we do. There is free will, according to van Inwagen, only if someone has, or once had, a choice about a true proposition; that is, only if someone can render, or could have rendered, a true proposition false.[32]

Now it would be more than natural to expect that for each human agent a there corresponds a set S of all true propositions a can render, our could have rendered, false. After all, in life, there are only so many things we can do—so many choices we can make. But, as it happens, there is no such thing as S. For suppose there was a set S of that kind for an agent a. Suppose further that for each $P \subseteq S$ there corresponds a true proposition p_P, say, *all elements of P are true*, such that a can or could have rendered it false. Note that it suffices in order for a to be able to render p_P false to have the ability to render any single element of P false—an ability a does possess after all, by the nature of the set S under consideration. Finally, suppose that for every $P, Q \subseteq S$, if $P \neq Q$, then $p_P \neq p_Q$, which holds under the assumption that propositions about distinct sets are themselves distinct.[33] A contradiction ensues. By the Axiom of Separation, there is a set R collecting all and only the propositions that are not also members of their corresponding set, that is:

$$R = \{x \in S \mid \exists P \subseteq S(x = p_P \text{ and } x \notin P)\}.$$

Because $R \subseteq S$ there is also a corresponding proposition p_R. Assume $p_R \notin R$. Then p_R is not an element of its corresponding set R. Since this is exactly what it means for something to be an element of R, it follows that $p_R \in R$. So, discharging the assumption, we have shown $p_R \notin R$ *only if* $p_R \in R$, which implies $p_R \in R$. So we know $p_R \in R$. Now it follows by definition of R that for some $P \subseteq S$, $p_R = p_P$ and $p_R \notin P$. But then $R \neq P$, and therefore $p_R \neq p_P$.

Not only it is interesting that van Inwagen's notion of choice comes with its own set-theoretic 'paradox'; this result yields the nonexistence of other mythical sets. For if there is no set of all truths one can render false, then

---

[32]There are different ways of understanding the notion of rendering a proposition false, some of which are discussed and explored by Lewis (1981), Huemer (2000), and many others. This point, however, is completely irrelevant to the present result, as any interpretation of what this notion means will do for our purposes.

[33]This premise does not hold under the assumption that propositions are sufficiently coarse-grained, that is, if they are sets of possible worlds. So one could alternatively take the present result as showing certain limitations on the granularity of propositions. For more on this, see Uzquiano (2015).

there is no set of all contingent truths, and so there is no set of all truths. This finally entails that there is no set of all propositions.

References

Aristotle. (1984). *The Complete Works of Aristotle: The Revised Oxford Translation*, vols. 1-2, edited by J. Barnes. Princeton, NJ: Princeton University Press.

Armstrong, D. M. (2004). *Truth and Truthmakers*. Cambridge: Cambridge University Press.

Barwise, J. and Perry, J. (1983). *Situations and Attitudes*. Cambridge, Mass., Bradford Books, The MIT Press.

Bringsjord, S. (1985). Are There Set Theoretic Possible Worlds? *Analysis* 45(1): 64.

Carnap, R. (1956). *Meaning and Necessity*. Chicago: University of Chicago Press.

Chihara, C. (1998). *The Worlds of Possibility*. Oxford: Clarendon Press.

Chisholm, R. (1970). Events and Propositions. *Nous* 4(1): 15-24.

Chisholm, R. (1971). States of Affairs Again. *Nous* 5(2): 179-189.

Crossley, J. N. and Humberstone, L. (1977). The Logic of "Actually". *Reports on Mathematical Logic* 8: 11-29.

Davies, M. (1981). *Meaning, Quantification, Necessity: Themes in Philosophical Logic*. London: Routledge & Kegan Paul.

Davies, M., and Humberstone, L. (1980). Two Notions of Necessity. *Philosophical Studies* 38(1): 1-30.

Enderton, H. B. (2001). *A Mathematical Introduction to Logic*. Academic Press; 2nd edition.

Finch, A. and Warfield, T. A. (1998). The MIND Argument and Libertarianism. *Mind* 127: 515-528.

Fine, K. (2017a). A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation. *Journal of Philosophical Logic* 46(6): 625-674.

Fine, K. (2017b). A Theory of Truthmaker Content II: Subject-matter, Common Content, Remainder and Ground. *Journal of Philosophical Logic* 46(6): 675-702.

Frege, G. (1979). *Posthumous Writings*. Translated by P. Long and R. White, Chicago: University of Chicago Press.

Frege, G. (1980). *Philosophical and Mathematical Correspondence*. Translated by H. Kaal, and edited by G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart, Chicago: University of Chicago Press.

Frege, G. (1984). *Collected Papers on Mathematics, Logic and Philosophy*. Oxford: Basil Blackwell.

Grim, P. (1984). There is No Set of All Truths. *Analysis* 44(4): 206-208.

Huemer, M. (2000). Van Inwagen's Consequence Argument. *Philosophical Review* 109(4): 525-544.

Humberstone, L. (1981). From Worlds to Possibilities. *Journal of Philosophical Logic* 10: 313-339.

Jubien, M. (1988). Problems with Possible Worlds. In D. F. Austin (ed.), *Philosophical Analysis*, Dordrecht: Kluwer, 299-322.

Kearns, S. (2011). Responsibility for Necessities. *Philosophical Studies* 155: 307-24.

King, J. (2007). *The Nature and Structure of Content*. Oxford: Oxford University Press.

Lampert, F and Merlussi, P. (2021). Counterfactuals, Counteractuals,

and Free Choice. *Philosophical Studies* 178: 445-469.

Lewis, D. (1981). Are We Free to Break the Laws? *Theoria* 47: 113-21.

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Menzel, C. (1988). On an Unsound Proof of the Existence of Possible Worlds. *Notre Dame Journal of Formal Logic* 30(4): 598-603.

Menzel, C. (2012). Sets and Worlds Again. *Analysis* 72(2): 304-309.

Merricks, T. (2009). Truth and Freedom. *Philosophical Review* 118(1): 29-57.

Plantinga, A. (1974). *The Nature of Necessity*. Oxford: Oxford University Press.

Plantinga, A. (1976). Actualism and Possible Worlds. *Theoria* 42: 139-160.

Plantinga, A. (1985). Replies to My Colleagues. In Tomberlin and van Inwagen, 313-396.

Pollock, J.L. (1985). Plantinga on Possible Worlds. In Tomberlin and van Inwagen, 121-144.

Pruss, A. R. (2013). Incompatibilism Proved. *Canadian Journal of Philosophy* 43: 430-437.

Russell, B. (1903). *The Principles of Mathematics*. New York, W. W. Norton and Co., 1st edition.

Schnieder, B. S. (2004). Compatibilism and the Notion of Rendering Something False. *Philosophical Studies* 117(3): 409-428.

Soames, S. (2007). Actually. *Proceedings of the Aristotelian Society*, Supplementary Volume 81: 251-77.

Stalnaker, R. (1976). Propositions. In A. F. MacKay and D. D. Merrill (eds.), *Issues in the Philosophy of Language*. New Haven, CN: Yale University Press, 79-91.

Stump, E., and Fischer, J. (2000). Transfer Principles and Moral Responsibility. *Philosophical Perspectives* 14: 47-56.

Tomberlin J.E. and van Inwagen, P. (eds) (1985). *Alvin Plantinga*. Dordrecht: D. Reidel.

Turner, P. R., and Capes, J. (2018). Rule A. *Pacific Philosophical Quarterly* 99: 580-595.

Uzquiano, G. (2015). A Neglected Resolution of Russell's Paradox of Propositions. *The Review of Symbolic Logic* 8(2): 328-344.

Van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Oxford University Press.

Warfield, T. (1996). Determinism and Moral Responsibility Are Incompatible. *Philosophical Topics* 24: 215-226.