Which Modal Machinery Should the Set Theoretic Potentialist Use?

REDACTED

Abstract

Accepting some form of potentialist set theory promises to help us solve puzzles about the intended height of the hierarchy of sets. However, philosophers have developed two different schools of potentialist set theory: minimalist and dependence-based approaches. In this paper, I will argue that minimalist formulations of potentialism have some important advantages over dependence-based formulations.

1 Introduction

Accepting some form of potentialist set theory promises to help us solve puzzles about the intended height of the hierarchy of sets. However, philosophers have developed two different schools of potentialist set theory, which I will call (following Neil Barton[1]) minimalist[20, 12, 13, 3] and dependence theoretic[18, 19, 17, 15, 23] formulations of potentialist set theory.

In this paper, I will review the motivations for potentialist set theory and the core differences between minimalist and dependence theoretic approaches to potentialism. I will then develop three arguments for favoring minimalist potentialism over dependence potentialism[1].

2 Background

So, let's begin by reviewing what potentialism about set theory is, and one important way this view can be motivated. Recall that, as a way of avoiding

[1]Earlier versions of some of these arguments appear in REDACTED.

Russell's paradox (among other things), contemporary set theorists embrace an iterative hierarchy conception of sets. On this view, all sets exist within a hierarchy of different layers (that satisfy the well-ordering axioms). There's the empty set at the bottom. And each layer of sets contains sets corresponding to 'all possible ways of choosing' from sets below that layer (i.e., all subsets of the set of sets occurring at prior layers).

This conception precisely characterizes the intended width of the hierarchy of sets. But what about the height of the hierarchy of sets? How many layers of sets are there?

Naively, it is tempting to say that the hierarchy of sets is supposed to extend 'all the way up' in a way that guarantees it satisfies the following principle.

> **Naive Height Principle:** If some objects are well-ordered by some relation $<_R$, there is an initial segment of the hierarchy of sets whose structure mirrors that of these objects under relation $<_R$ (in the sense that the objects related by $<_R$ could be 1-1 order preservingly paired onto the layers in this initial segment).

But this assumption leads to contradiction via the Burali-Forti paradox.[2] And, there's no widely accepted precise conception of how high the hierarchy of sets is supposed to go up, which can replace the above paradoxical idea. Furthermore, it seems arbitrary to say that the hierarchy of sets just happens to stop somewhere.

Note that the problem here is not simply that it might be impossible to define the intended height of the hierarchy of sets in other terms. After all, every theory will have to take some notions as primitive. Instead, we find

[2]For example, considering the relation '$x <_R y$ iff x and y are both layers in the hierarchy of sets and x is below y or y is the Eiffel tower and x is a layer' shows the above naive conception of the hierarchy of sets cannot be satisfied. We have a sequence of objects that is strictly longer than the hierarchy of sets (it's a theorem of ZFC that no well-ordering is isomorphic to a proper initial segment of itself), contradicting the Naive Height Principle above.

ourselves in the following situation. Our naive conception of absolute infinity (the height of the actualist hierarchy of sets) turns out to be incoherent, not just unanalyzable. And, once we reject this naive conception, there's no obvious fallback conception that *even appears* to specify a unique height for the hierarchy of sets in a logically coherent way.

In principle, one could reply by claiming to grasp a notion of 'absolute infinity' — the intended stopping point for the hierarchy of sets — as a primitive, and using that to replace our naive conception of the intended hierarchy of sets. But this doesn't seem to be a plausible, or popular, choice of primitive (unlike, e.g., the notion of 'all possible ways of choosing' which we use to pick out the width of the hierarchy of sets)[3].

One might try to avoid this problem (of apparent commitment to an arbitrary and mysterious stopping point for the hierarchy of sets) by saying the height of the hierarchy of sets is vague [4]. For example, you could say all interpretations of set theory that make the standard ZFC axioms (Zermelo Fraenkel with Choice) come out true (or all that satisfy the conception of the *width* of the hierarchy of sets above) are acceptable precisifications of our current vague concepts.

[3]Also note that we cannot address this arbitrariness challenge merely by citing familiar arguments against the existence of a universal set. For one can distinguish two questions about the height of the hierarchy of sets analogous to two senses of the question 'why am I not taller than I (actually) am?' There's one sense of this question which can be answered by logical-metaphysical reflections implying that no one has the property of being taller than themselves. This is analogous to using Russell's paradox and Comprehension to explain why the hierarchy of sets lacks the property of containing some set which has all sets as members. However, there's another sense to 'Why am I not taller than I am' which cannot be answered by logic/metaphysics alone and might be elaborated by asking things like 'why am I 5'3 rather than 5'8?'. This latter question is analogous to the explanatory question about the height of the hierarchy of sets at issue. Once we abandon our conception of the hierarchy of sets as going 'all the way up darn it', we face a question about why it stops at one place rather than another.

[4]One could then say the truth-value for claims in the language of set theory is determined by supervaluation. In this paper, I'll focus on this way of thinking about vagueness, for concreteness. Note that attributing such vagueness to the height of the hierarchy of sets still lets one explain the acceptability of classical reasoning (as e.g., Field [6] points out). But plausibly a similar argument could be made for other ways of thinking how about vagueness and vague meanings affect the truth value of sentences.

However, such views have trouble accounting for mathematicians' tendency to favor taller over shorter candidates for the hierarchy of set structure. That is, mathematicians tend to regard claims that impose a lower bound on the height of the hierarchy of sets (and coherently extend the conception of the width of the hierarchy of sets above) as true — rather than indeterminate (as the theory of vague reference would suggest)[5]. Also taking this route would require rejecting a somewhat common metaphysical doctrine that existence is never vague[6]. For different hypotheses about the height of the hierarchy of sets imply different constraints on the total cardinality of the universe, not just vagueness about whether 'set' applies to certain objects that uncontroversially exist[7].

3 Potentialism

Potentialist approaches to set theory promise to solve the above problem (about the intended height of the hierarchy of sets) by denying that there is a single intended hierarchy of sets whose structure determines the truth value of set-theoretic claims. They reinterpret ordinary set-theoretic statements, to replace quantification over sets within some fixed structure with claims about what structures of a certain kind there could be.

[5] C.f. the tendency to accept large cardinal axioms and Gödel's appeal to Ackerman's reflection principle [24]

[6] c.f. chapter 4 section 9 of Sider's *Four Dimensionalism*, for example. .

[7] In principle, one could avoid the verdict that existence is vague (while also avoiding commitment to an arbitrary stopping point for the height of the hierarchy of sets) by saying that there is simply vagueness about which determinately existing objects the predicate "set" applies to, rather than any vagueness in what objects exist or what "there is" and similar quantifier expressions mean. So, for example, you might say that there are layers of things that uncontroversially exist but get (for whatever reason) less determinately set like and more proper class like (like swatches of color getting less red and more orange)? However saying this seems to only temporarily delay the worry that we are committed to an arbitrary stopping point of some kind. While this view can explain away a certain amount of arbitrariness in the height of the hierarchy of sets it implies that the total height can't be larger than the collection of all actual existing objects (including sets, proper classes/other mathematical objects and non mathematical objects). So the problem of commitment to an arbitrary stopping point in some sense (an arbitrary stopping point to the universe, if not for the hierarchy of sets) still remains.

This basic idea has been developed in two major ways, which I will call (following [1]) minimalist and dependence theoretic potentialism. I will explain each view in more detail below. However, in a nutshell, the key difference is this. Dependence theoretic potentialists understand set theory as talking about what *sets* there could (in some sense) be. In contrast, minimalists interpret set theory as talking about how it would be possible for there to be *objects (of any kind)* satisfying certain set-theoretic axioms.

3.1 Minimalist Potentialism

To my knowledge, the Minimalist approach to potentialist set theory originates with Hilary Putnam. In [21] Putnam sketches a way of thinking about set theory in terms of modal logic: as talk about what 'models' of set theory are, in some sense, *possible* and how such models can be extended.

Putnam considers (standard) models of set theory built of concrete objects like points marked in pencil that are related by arrows. Putnam suggests that we can understand set-theoretic statements as claims about what such concrete models are possible, and how they can be expanded.

- For example, we can paraphrase a set theoretic statement of the form $(\exists x)\phi(x)$, where ϕ is quantifier free, as saying the following. It is possible that there's a concrete model G and a point p within G (e.g., a dot in a system of physical dots related by physical arrows), such that $\phi(x)$ is true of the object x within the physical structure G (that is, ϕ expresses a property that's true of p, when when '.. is an element of..' is replaced by 'there is a physical arrow from... to ...').

- Similarly we can paraphrase a more complex set-theoretic statement of the form $(\forall x)(\exists y)(\forall z)\phi(x, y, z)$, where ϕ is quantifier free, as saying that the following. If G is a concrete model, and p is a point within G, then it is possible that there is a model G' which extends G, and a point y within G' such that necessarily, for any concrete model G'' which extends G' and contains a point z, $\phi(x, y, z)$ holds within G''.

And (on Putnam's proposal) can treat arbitrary quantified statements in set theory in an analogous fashion.

Now when we talk about the possibility of (extending) concrete standard models for certain axioms of set theory, what kind of possibility do we mean? Putnam is somewhat noncommittal. But in [12] Hellman suggests taking the key modal notion $\Diamond$ in Putnam's potentialist set theory to be a primitive modal notion of logical possibility, and taking standard models of set theory to be models which satisfy ZFC_2 (the second order logical version of the Zermelo Frankel with Choice axioms of set theory).

Thus, on Hellman's minimalist potentialism (circa [12]), set-theoretic statements can be expressed entirely in terms of modal and logical primitives. For example, the potentialist paraphrase of $(\forall x)(\exists y)(x \in y)$ might look like this (where quantification over all V_i is shorthand for quantification over all second order objects X_i, f_i satisfying ZFC_2)

$\Box(\forall V_1)(\forall x)[x \in V_1 \rightarrow \Diamond(\exists V_2)(\exists y)(y \in V_2 \wedge V_2 \geq V_1, \wedge x \in y)]$

Adopting such a potentialist approach to set theory lets us understand set-theoretic talk without imposing arbitrary limits on the size of structures (in a way that seems faithful to our intuitions about the generality of set-theoretic reasoning).

Later work on minimalist potentialism explores variations on the above proposal (using a logical possibility operator $\Diamond$ and second-order quantification). For example, Hellman considers replacing second-order quantification with plural quantification and mereology (and logical possibility with something like metaphysical possibility that preserves the laws of mereology). And [4] Berry argues that a certain natural generalization of the logical possibility operator $\Diamond$ (to a, so called, conditional logical possibility operator $\Diamond_{...}$), can do the work of second-order quantification in formulating minimalist potentialism[8]. Berry also advocates (in effect) replacing appeals to ZFC_2 with weaker axioms IH_2 expressing the conception of being an intended width initial segment of an iterative hierarchy of sets, as above.

3.2 Dependence Potentialism

In contrast to minimalist potentialism, dependence theoretic potentialists like Linnebo[14, 15, 17] and Studd[23] acknowledge the existence of special objects called 'sets' (like traditional actualists), but interpret set theory as talking about what sets there could (in some sense) be[9].

---

[8]That is, she argues one can formulate minimalist potentialist set theory using only first-order language and this conditional logical possibility operator.

[9]In [9] and [8] Kit Fine also discusses a notion of interpretational possibility, in connection with his postulationism about mathematical objects. I won't engage with this proposal much, because I agree with Studd in section 4.1 of [23], in being unsure how to understand Fine's talk of a third parameter associated with a sentence that's neither the content nor the context of utterance, but can change the truth-value (and possible worlds truth conditions) for that sentence.

In [9] pg 32 Fine motivates appeal to his notion of interpretational possibility by three specific arguments that one can't use the notion of logical possibility to cash out claims about absolute generality and the (supposed) possibility of talking in terms of more sets. So far as I can tell, most of his worries concern how to understand first and second-order *quantifying into* the diamond of logical possibility. So they prima facie don't apply to my preferred form of minimalist potentialism[4], which replaces all quantifying in, with claims about structure-preserving logical possibility $\Diamond_{...}$. That is, it replaces claims about what's logically possible for specific objects with claims what's logically possible given the structural facts about how some subscripted list of relations $R_1 \ldots R_n$ apply.

But perhaps some version of Fine's worry that if we use logical possibility when developing potentialism, we won't be able to capture the contrast between domains of quantification which are "inextendable" and those which are not, still applies. However, I'm not sure why we should want to say that any domains of quantification are inextendable. If we just think about what kinds of language change are possible in terms of sentences, it's clear that we could start speaking a language in which any word in our current language was associated with a larger extension than it currently has. In this sense you could start talking in terms of more "sets" but you could also start talking in terms of more "cats".

Admittedly, all potentialists will want to acknowledge some important difference between our current practices of "set" and "cat" *talk*, which makes potentialist explication of one but not the other attractive. But do we need to attribute this difference to our *concept* set (or all 'domains of quantification' including sets) having a special property of indefinitely extensiblility, which our concept cat (and some domains of quantification including only objects like cats) lacks? A minimalist potentialist might simply say "set" and "cat" talk differ in that the the Burali-Forti considerations above suggest the ideal Carnapian explication of our set talk

The claim that there could (in some sense) be more sets than there actually are can initially seem mysterious. And Linnebo and Studd allow that facts about what pure sets exist are *metaphysically* necessary. However, they cash out potentialist set theory by appealing to a notion of interpretational possibility. They appeal to facts about how it would be possible for us to successively reinterpret our language (or reconceptualize the world) so as to talk/think in terms of more and more sets.

When explaining interpretational possibility, Linnebo particularly focuses on ways of changing acceptable interpretations of our language by introducing *abstraction principles*[10]. Adopting such abstraction principles allows someone who is currently talking in terms of one actualist hierarchy of sets to start talking in terms of a taller one that adds a new layer of sets. Specifically, one can do this by adopting abstraction principles which (in effect) require that for every plurality of sets xx in the old sense of the term, there's to be a set (in the sense of the new language) which has all and only this plurality of old sets as elements[11]. Linnebo explains his key notion of interpretational possibility as follows.

> We may think of '$\Box\phi$' as meaning 'no matter what abstraction steps we carry out, it will remain the case that ϕ', and '$\Diamond\phi$' as 'we can abstract so as to make it the case that ϕ'. Obviously, this interpre-

(i.e., the one that best fits and explains intuitive truth conditions, assertability conditions, practical usefulness etc.) must diverge from surface grammar in a potentialist way, while no analogous puzzles and paradoxes arise for 'cat' talk.

[10] In [17] Linnebo suggests that some objects, including the sets, are 'thin', in the sense that we can come to know things about them by introducing abstraction principles that specify identity conditions relating them to antecedently understood objects. For example, in Frege's classic case, if you are already talking about lines, you can start talking in terms of the abstract objects we call 'directions', by stipulating that two lines have the same direction iff they are parallel.

[11] Here I use scare quotes because I take it that (on Linnebo's picture) when speaking your new language (post acceptance of abstraction principles) it's straightforwardly true to say there's a set whose elements are exactly the plurality of objects xx. However, I use scare quotes when speaking prospectively, about what the impact of adopting certain abstraction principles would be, since we are not talking about objects in the extension of the term 'set' in our current language. I don't take much to hang on this difference.

> tation of the modal operators is different from the more familiar one in terms of metaphysical modality.

Specifically, Linnebo takes the "modal operators $\Box$ and $\Diamond$ [interpretational possibility and necessity] to describe how the interpretation of the language can be shifted — and the domain expanded — as a result of abstraction."[17]. Even more specifically, an interpretational possibility claim $\Diamond\phi$ is supposed to be true iff you could make ϕ true via some well-ordered sequence of acts of reconceptualization (whether or not it would be metaphysically possible for anyone to make such a sequence of abstractions)[12].

In this way, Linnebo and Studd can make dependence potentalist claims that there 'could' be more sets in one sense (i.e., we could talk/think in terms of more sets), while also agreeing with common intuitions that all facts about pure sets existence are metaphysically necessary[13].

[12]Linnebo describes this as follows

> "the dynamic approach allows abstraction to be iterated, which turns out to be an extremely powerful tool. The idea is simple. One application of plural Law V [i.e., the abstraction principle Linnebo proposes to add one layer of sets] takes us from some initial domain to a larger one. Since this larger domain gives rise to more pluralities than the initial one, a second application of the law gives rise to even more objects. We can continue in this way indefinitely. At limit stages, we take the union of all the objects generated thus far. Since each round yields something new—as Russell's paradox would otherwise re-emerge—the process never terminates." pg 60 [17].

.

[13]Studd explicitly references Linnebo's notion of interpretational possibility as similar when explaining his favored approach to interpretational possibility. However, Studd puts less emphasis on language change *via (specifically) acts of Fregean abstraction* than Linnebo. For Studd describes "admissible interpretations" (i.e., the possible reinterpretations of our language relevant to the interpretational possibility operator) more broadly, writing that such interpretations "result from shifts of interpretation of the kind that a relativist may bring about in her attempt to expand the universe."[23]

When initially explaining and motivating the notion of interpretational possibility Studd is very generous about the range of language change events considered (he gives the example of historical narrowing in the extension of the word "meat"). However, Studd specifies that his proposed modal operators are only intended to generalize over reinterpretaions of our language which

1. assign "intended" interpretations to non-logical vocabulary like 'set' and 'element' (Studd takes there to be "An open-ended sequence of [such] intended interpretations"[23] for the language of set theory)
2. are 'expansive', going from a smaller universe to a larger one in such a way that the original smaller universe can be represented as a quantifier restriction of the larger one

Dependence potentialist paraphrases of set-theoretic sentences have a similar large-scale structure to the minimalist potentialist paraphrases discussed above. However, (as noted above) they take the relevant notion of possibility to concern what sets could, in some sense, be formed (e.g., what is interpretationally possible). Additionally, dependence potentialists don't specify the structure of the iterative hierarchy in their paraphrases. Instead, they take the fact that the sets form an iterative hierarchy to fall out of the interpretational essences of sets (certain principles, like extensionality, that must continue to hold true when we change our language to start talking in terms of more sets) plus the idea of a well-ordered sequence of reinterpretation events[14].

For example, we said that a minimalist potentialist paraphrase of $(\forall x)(\exists y)(x \in y)$ might look like this (where quantification over all V_i is understood as shorthand for quantification over all second-order objects X_i, f_i satisfying some axioms like ZFC_2)

$\Box(\forall V_1)(\forall x)[x \in V_1 \rightarrow \Diamond(\exists V_2)(\exists y)(y \in V_2 \wedge V_2 \geq V_1, \wedge x \in y)]$

In contrast, the dependence theorist will paraphrase the same sentence more simply, as follows (using 'set' as a primitive):

$\Box(\forall x)[set(x) \rightarrow \Diamond(\exists y)(set(y) \wedge x \in y)]$

(see Studd's axioms of monotonicity and stability on pg 149 of [23] for more detail).

Another difference between Linnebo's interpretational possibility operator and Studd's is that (where Linnebo has a single $\Box$ and $\Diamond$) Studd has two basic modal operators $\Box_>$ and $\Box_<$ whose meaning he explains in terms of what could be got by a well-ordered sequence of purely expansive language re-interpretation events, as follows.

- $\Box_>$ 'however the lexicon is interpreted by succeeding interpretations'
- $\Box_<$ 'however the lexicon is interpreted by preceding interpretations'

So Studd (in effect) considers what you could get to by going either forward ($\Box_>$) or backwards ($\Box_<$) in a sequence of quantifier meaning 'expansion' events, whereas Linnebo only considers what's possible looking forwards. Despite these differences, I take it to be clear how the same concerns about commitment to something like a notion of logically possible well orderings which I've raised for Linnebo also apply to Studd.

[14]The dependence potentialist imagines a hierarchy of sets which could grow (with new sets somehow being formed) as follows. The empty plurality always exists. So an empty set could be formed. Form it/Reconceptualize to recognize it. Now there's a plurality xx whose sole member is the empty set, so a set $\{\{\ \}\}$ could be formed. Form that. Now that both these sets exist, there are four pluralities xx of sets. And two of them correspond to sets we don't already have. So we could form $\{\{\{\ \}\}\}$ and $\{\{\{\ \}\}, \{\}\}$ etc.

Because the dependence theorist takes facts about the interpretational essence of sethood to ensure that the hierarchy of sets has certain a certain structure in all interpretationally possible scenarios, their paraphrases don't need to include a description of this structure.

Accordingly, we should note that when the dependence potentialist talks about interpretational possibility, they aren't generalizing about what claims can be made true via arbitrary changes to the meaning currently associated with symbols "set" and "element of". Rather, they take there to be something about the current *meaning* of the word "set" which allows for certain changes to its interpretation, but rules out others. For example, reinterpretations of our language and "set" which merely (so to speak) add an extra layer of sets are allowed/intended. But those that would make the axiom of extensionality false are unintended (e.g., it is not interpretationally possible for there to be two distinct sets with no elements)[15].

I admit that dependence potentialism has the advantage of nicer looking paraphrases. But typesetting isn't all[16].

Also note that, as part of this story, dependence potentialists take ordinary set-theoretic talk to have two different legitimate readings.

- Set talk expresses potentialist modal claims about what sets could be

[15]Studd references an open-ended sequence of different "intended" interpretations of 'set' and 'element'[23]. And Linnebo suggests that our concept set is unusual in having a precise intension which fails to determine a precise extension. He writes

> "Suppose we have formulated a perfectly precise notion of a star. For any object whatsoever, this notion enables a definitive verdict as to whether or not the object is a star. When this precise intension is applied to the world, reality answers with a determinate extension, namely the plurality of objects that satisfy the intension. And there is nothing unusual about stars in this regard. *In most ordinary empirical cases, a precise intension determines an extension when applied to the world. But in mathematical cases, and other cases involving abstraction, this is no longer so. Here a precise intension often fails to determine an extension.*"[17].

[16]If concerned about typesetting, we could easily introduce conventions by which statements that look like a dependence theoretic paraphrase are used to abbreviate a corresponding minimalist potentialist paraphrase.

'formed' (i.e., what sets we could start talking/thinking in terms of). This is what determines the assertability of set-theoretic claims in normal mathematical contexts.

- Set talk also has an actualist reading. There is an actualist hierarchy of sets that we are currently (in some sense) talking in terms of.

For example, consider the claim that every set is an element of some other set.

- On the actualist reading, this says that every existing set is a member of some other set. So, e.g., this claim will be false if the actual hierarchy of sets cuts off at V_{15} (since certain sets first occur at the successor stage 15) and true if the actual hierarchy has the structure $V_{\omega+\omega}$
- On the potentialist reading, this means that for every existing set x, it would be interpretationally possible to extend the hierarchy of sets to contain another set y which x is an element of (e.g., by reinterpreting 'set' to apply to an extra layer of objects).

4 Commitment to Minimalist Potentialism about the Ordinals?

With this contrast between minimalist and dependence potentialism in mind, let's turn to the question of which version to favor. In the remainder of this paper, I'll propose and develop some arguments for favoring the minimalist approach.

The first worry I want to mention arises from the fact that dependence potentialists like Linnebo and Studd seem to accept the meaningfulness and correct behavior of all the logical machinery needed to give minimalist potentialist paraphrases.

For example, Linnebo's explanations of the notion of interpretational possibility (used to articulate dependence theoretic potentialism about *set theory*) suggest acceptance of (something like) a minimalist potentialist conception of *the ordinals*, in a way that can raise questions.

Recall that Linnebo takes an interpretational possibility claim $\Diamond\phi$ to be true if and only if you could make ϕ true via some *well-ordered sequence* of acts of reconceptualization (whether or not it would be metaphysically possible for anyone to perform such a sequence of acts).[17]

However, in doing this, Linnebo seems to be assuming something very much like a notion of logical possibility. To the extent we understand this appeal to a notion of in principle possible well-ordered sequences of *language reinterpretation* events, it seems we should exactly understand analogous facts about the in principle possible well-ordering of many other kinds of objects. Thus Linnebo seems to accept the meaningfulness of something very close to the extra logical machinery neeeded/used to formulate potentilist set thoery.

Indeed (more directly) Linnebo seems to be invoking something very similar to the minimalist potentialist understanding of the ordinals when explaining his notion of interpretational possibility. For when we compare Linnebo's talk of possible well-ordered sequences of reconceptualization events with Putnam's talk of in principle possible pencil points and arrows forming standard models of set theory, the difference does not seem great.

[17] He writes as follows.

> the dynamic approach allows abstraction to be iterated, which turns out to be an extremely powerful tool. The idea is simple. One application of plural Law V [i.e., the abstraction principle Linnebo proposes can add one layer of sets] takes us from some initial domain to a larger one. Since this larger domain gives rise to more pluralities than the initial one, a second application of the law gives rise to even more objects. We can continue in this way indefinitely. At limit stages, we take the union of all the objects generated thus far. Since each round yields something new—as Russell's paradox would otherwise re-emerge—the process never terminates." pg 60 [17].

Accordingly minimalist potentialists may raise the following challenge. If we're willing to (essentially) presume understanding of minimalist potentialism about *the ordinals* (as Linnebo seemingly is), wouldn't it be simpler and more elegant to just extend this story to minimalist potentialism about **the sets**? Why bother with dependence theory?

Similarly, Studd seems to accept without objection, both the basic notion of logical possibility, and Hellman's use of this notion to provide a potentialist explication of set theory that at least gets truth values right. So he faces an analogous question about why we should bother with dependence potentialist paraphrases (e.g., why bother with introducing the somewhat less familiar notion of interpretational possibility as a primitive)?

To sharpen this challenge, let me make some quick clarifications. First, I'm not claiming Linnebo and Studd literally accept minimalist paraphrases *as correct explications for mathematicians' talk of the ordinals* — and so must accept corresponding minimalist paraphrases of set theory in order to treat the sets and ordinals analogously. For Linnebo and Studd would presumably reply that, although (as argued above) they seem to accept both minimalist and dependence paraphrases as meaningful and getting truth-values right, further desiderata for choosing conceptual analyses/Carnapian explication favor dependence paraphrases.

Second I'm not claiming that dependence potentialists absolutely **must** accept minimalist potentialist explications of set and ordinal talk as using acceptable primitives and getting truth-values right.[18]. I'm only noting that actual

[18]In principle, dependence theorists could take interpretational possibility as a metaphysical and conceptual primitive, and explain apparent appeals to something like logical possibility when explaining interpretational possibility as an illusion or a temporary expedient. For example, dependence theorists could say they use appeals to logical possibility etc. to explain their terms and then recommend that we kick away the ladder. After all, both forms of potentialists use appeal to set models to introduce and explain their modal notions.

Alternately, dependence theorists could say the modal notion invoked by their talk of possible sequences of language re-interpretation is itself interpretational possibility. They could claim their explanations have a kind of (putatively non-vicious) circularity: explaining in-dependence theorists like Linnebo and Studd *do actually* seem to allow that minimalist explications of potentialist set theory have these features (and face some pressure to do so)[19] – and noting that this fact raises a potential challenge. If you already accept something like the minimalist potentialist understanding of the ordinals as using meaningful vocabulary and getting truth conditions right, why bother with dependence potentialism about the sets? Why not treat set talk and ordinal talk alike, by going minimalist about both?'

Third, I admit that the above challenge is not a decisive criticism on its own. For the dependence theorist will probably agree that we should treat the sets and ordinals analogously but maintain that both minimalist and dependence paraphrases satisfy the minimal requirements specified above (using acceptable primitives and getting truth values for pure mathematical sentences right) and that we should prefer dependence paraphrases for some other reason. Accordingly, I'll devote the remainder of this paper to supplementing my challenge for dependence theorists with some positive arguments that we should favor minimalist potentialism – once it's granted that both minimalist and dependence potentialism pass the low bar specified above. [20]

terpretational possibility by appeal to interpretational possibility. Indeed, Studd's remarks about Kripke models for interpretational possibility may have something of this character[23].

So, in principle, one could accept interpretational possibility while denying that there's a meaningful notion of logical possibility. However this would be an uncomfortable position. For once we accept a notion of possible well ordered sequences of reinterpretation events that's unconstrained by metaphysically necessary cardinality limits on space and time, it seems hard to resist the meaningfulness of the notion of logical possibility which (in effect) tries to generalize this to other kinds of objects and relations.

[19] Admittedly in [16], Linnebo raises a worry for Hellman's minimalist potentialism, via suggesting the possibility of "'metaphysically shy' objects, which can live comfortably in universes of small infinite cardinalities, but which would rather go out of existence than cohabit with a larger infinite number of objects"[16]. The existence of such shy objects would pose a problem for Hellman, because it could block us from saying that every plurality of objects forming a hierarchy of a certain kind could be extended in a certain way.

However, Linnebo himself seems to regard this problem of metaphysically shy objects as not too serious. For he notes that the minimalist can plausibly hope to solve it by understanding potentialist set theory as making claims about structure preserving extendability rather than de re possibility. And the version of dependence potentialism in [3, 4] which eliminates quantifying in can be seen as implementing this strategy.

[20] One might criticize minimalist potentialism (cashed out using logical possibility) on the grounds that, "logical possibility and necessity are not genuine modalities acting on propositions, but at best quasi-quotational devices sensitive to modes of presentation." For example,

5 Better Fit with Logicist and Structuralist Intuitions

My second concern centers on a claim that minimalist paraphrases better fit with some enduringly popular - broadly logicist and structuralist - ideas about the nature and purpose of set theory (or mathematics as a whole) than dependence theoretic ones do. I have three specific points in mind.

First, note that minimalist paraphrases attractively fit common broadly logicist intuitions – that mathematics is a part of logic, or somehow closely related to it. Minimalist explications for set theory seem to support this claim, at least if developed using a logical possibility (or conditional logical possibility) opera-

---

there's a prima facie awkward question about whether it's logically necessary that all furze is gorse. Since 'furze' and 'gorse' are different names for the same plant, it's appealing to say that 'all fuze is furze' and 'all furze is gorse' express the same proposition. Yet it seems that 'it's logically necessary that all furze is furze' is true, while 'it's logically necessary that all furze is gorse' is false. So there's pressure to say that logical possibility applies to something like sentences, not propositions – and hence logical possibility is 'a mere quasiquotational device'. If we accept this argument (and think 'quasi-quotational devices' can't be used to formulate potentailist set theory), the kind of minimalist potentialism using logical possibility I've defended in this paper will be ruled out.

However, I don't think the above argument against treating logical possibility as a genuine modality works. For one thing, the minimalist potentialist could deny that "furze" and "gorse" really express the same concept by appeal to Chalmersian two dimensionalism (Williamson seems to allow this as a currently live option[25]).

Alternatively, the minimalist potentalist could accept that 'all furze is furze' and 'all furze is gorse' express the same proposition, but say that 'it is logically necessary that all furze is gorse' is actually true. Admittedly, taking this approach would require saying that there are some presentations of logically necessary truths - like 'all furze is gorse'- which competent English speakers won't be in a position to learn a priori. However minimalist potentialists (who employ a notion of logical possibility) already have strong independent reason to allow the possibility of such a priori argument transcendent logically necessary truths. For, they think facts about logical possibility and necessity can be used to explicate set theory and thereby number theory. And well known arguments about mechanizable minds and Gödel incompleteness suggest that, for every thinker with minds like ours, there will be some claim in the language of arithmetic which is true but not discoverable via a priori reasoning methods this thinker accepts. So they have independent reason to accept the existence of some logical truths which (can be presented in such a way that) are not knowable a priori or by linguistic competence alone. Taking 'all furze is gorse' to express a logical truth would just require accepting that more logical truths have this feature.

Finally, the above line of thinking might also be used to challenge whether interpretational possibility is a real modal operator (vs. mere quasi-quotational device). For it's not obvious whether it's interpretationally necessary that all furze is gorse (similar strategies to the above could be used). So it's not clear this kind of criticism of minimalist potentialism (invoking logical possibility) can be leveled by dependence potentialists (invoking interpretational possibility).

tor as discussed above. For such paraphrases explicate set theoretic sentences as making claims about what is logically possible and necessary. In contrast, Linnebo and Studd's dependence potentialism connects the meaning of set theory to facts about possibilities for ontologically inflationary *language change* (interpretational possibility), rather than facts about logical possibility. In contrast with logicism, there is (to my knowledge) no comparable tradition of expecting a close relationship between mathematics and possibility via ontologically inflationary language change.

Second, minimalist paraphrases of set theory appealingly fit *structuralist intuitions* that mathematics is somehow 'the science of structure' and so that facts about the individual natures and essences of objects are irrelevant to mathematics. For, these minimalist paraphrases appeal to logical possibility constraints (which are supposed to apply equally to all objects and relations), rather than to interpretational possibility facts which are constrained by the supposed interpretational essences of our set and element concepts.

And minimalists who use the conditional logical possibility operator approach of[3, 4] may be able to claim an especially good fit with structuralist intuitions. For these minimalists replace traditional potentialist claims about de re modality (what's possible for a specific object) with claims about structure-preserving logical possibility $\Diamond_{...}$ (what's possible given structural facts about how the relevant relations ... apply). So they may claim to fit structuralist intuitions (that facts about the essences and natures of particular objects are irrelevant to mathematics) especially well in this way. In contrast, dependence paraphrases a la Linnebo and Studd don't attempt to appeal to a modal notion that treats all relations of the same arity alike, and do appeal to facts about the interpretational essence of sets.

Thirdly, one might argue that contemporary formulations of dependence theory by Linnebo and Studd have the disadvantage of packing intuitively irrelevant and controversial claims about *modal logic* into their explications of basic set theoretic claims (while minimalist paraphrases avoid doing this)[21] . For both Linnebo and Studd endorse the controversial converse Barcan-Marcus principle, which implies everything exists necessarily[22]. They thus risk giving paraphrases of ordinary set theoretic claims, on which the truth of these claims depends on controversial principles about de re modality hold. In contrast, minimalist potentialists can, and sometimes have, avoided taking a stance on this topic, by replacing de re possibility claims with claims about structure-preserving conditional logical possibility (as per [3, 4])[23]

For, there's a popular and enduring structuralist impulse to say that, even if mathematical objects exist, questions about their essence and persistence conditions are irrelevant to mathematics (c.f. the story of Hilbert saying that geometry was just as much about salt shakers and forks as about points and lines as standardly understood). For those who share this intuition, Linnebo's paraphrases of claims in basic set theory can seem to introduce intuitively irrelevant content *in the form of de re claims about what's interpretationally possible for specific sets.*

Admittedly, dependence potentialist arguably have one advantage as regards

[21]Specifically, you might think, potentialist explications for set theory should try to preserve the current intuitive content and interest of set-theoretic practice (by avoiding adding intuitively irrelevant commitments to its paraphrases of set-theoretic claims), so far as possible. C.f. "One might say: the axis of reference of our examination must be rotated, but about the fixed point of our real need"[26].

And my intended criticism of dependence potentialism in the passage above is somewhat analogous to the objection many would have to explicating heat talk with sentences of the form $M \wedge A$, where M is a claim about molecular motion and A is some metaphysically necessary aesthetic truth. Even if such paraphrases get the possible worlds truth conditions for each heat sentence right, they would still seem unappealing, because they seem to introduce novel and intuitively irrelevant content (in a way that could be easily avoided).

[22]Maybe they can reply that 'everything exists *interpretationally* necessarily' is analyticish, and therefore not controversial. But saying this makes the notion of interpretational possibility $\Diamond$ look less natural, hence a less attractive primitive.

[23]Admittedly dependence-potentialist paraphrases of set theory sentences don't mention any irrelevant *kind of object or property* like beauty in an analysis of heat above. But arguably dependence theoretic paraphrases for set theory do bring in intuitively mathematically irrelevant *features* of mathematical objects like sets like *de re* modality.

conservatism, as follows. Dependence theorists can agree with the commonsensical-sounding claims that 'set theory is about things called sets, not just some structures with such and such structural relations on them'[24] (while minimalists cannot). However, I think the existence of widespread and longstanding structuralist impulses noted above (suggesting that facts about the natures and essences of objects are irrelevant to mathematics, and only structure matters) makes it -at best- unclear that preserving the truth of the above commonsensical-sounding claim counts as a feature rather than a bug.

So, overall, I think minimalist potentialists can plausibly claim to introduce less intuitively irrelevant content to their regimentation of set theory. They can claim to provide paraphrases which do a better job of 'rotating around the axis of real need' (or real curiosity) than dependence paraphrases – even if both kinds of paraphrases are granted to use meaningful primitives and get possible worlds truth conditions right.

6 Conceptual Parsimony

Next, I will explore an argument that minimalist potentialism has an advantage in conceptual parsimony. I think this argument is, ultimately, less powerful (and more dependent on controversial premises) than the other criticisms of dependence potentialism presented in this paper. However I include it has some nuances and connections worth exploring.

Minimalist potentialists can claim to develop their theory using only independently motivated conceptual tools, like a logical possibility operator that's familiar (approximately interdefinable with validity c.f. [7][25]) and independently

[24]Thanks to an anonymous referee for this phrasing.

[25]An anonymous referee raised a worry that appeals to a notion of validity (which can be applied to arbitrary sentences) might allow formulation of something like the liar paradox. However, even if logical validity could generate paradoxes in unrestricted contexts, the minimalist potentialist need only apply it in restricted settings with simple vocabulary (c.f. the

motivated as a modal primitive. For example, works like [10, 11, 5] point out reasons why we plausibly can't analyze away such claims about logical possibility and entailment in terms of claims about the existence of set models[26]. Also note that treating the $\Diamond$ of logical possibility as a primitive *logical* operator lets us accommodate Boolos' intuition that, "it is somewhat peculiar [to say, as standard model theoretic approaches to validity do, that] if G is a logical truth, then the statement that G is a logical truth does not count as a logical truth, but only as a set-theoretical truth." And dependence theorists Linnebo and Studd seem to accept the legitimacy of this notion of logical possibility (or something very close to it)[27].

In contrast, (the minimalist might say) dependence potentailism requires

formal development of logical possibility-based potentialism in [3, 4])). They need not accept the meaningfulness of logical possibility/validity claims involving the paradox-prone semantic vocabulary above.

[26] In a nutshell, the problem with identifying logical possibility and entailment claims with assertions about the existence of mathematical objects is this. The claim that what's actual is logically possible is central to the notion of logical possibility, if anything is. For an argument to be valid surely at least requires that it doesn't *actually* lead from truth to falsehood. However, if we think about logical possibility in terms of set-theoretic models, then the actual world is strictly larger than the domain of any set-theoretic model (e.g., because it contains all the sets). So it's prima facie unclear why we should infer from the fact that ϕ isn't satisfied in any set-theoretic model, that ϕ isn't actually true. Thus, we seem to antecedently grip a notion of logical possibility (interdefinable with validity) that is different from the notion of having a set-theoretic model.

Admittedly it's possible for mathematicians talking about *first order logical sentences* to replace talk of logical possibility with talk of set-theoretic models for many normal mathematical purposes via the completeness theorem for first-order logic. For the completeness theorem shows that syntactic consistency and having a set-theoretic model are co-extensive which each other (and therefore with the intuitive notion of logical possibility) as regards first-order logical sentences. However, this does not apply beyond first order logic. Also as Boolos puts it, "it is rather strange that appeal must apparently be made to one or another non-trivial result in order to establish what ought to be obvious: viz., that a sentence is true if it is valid"[5].

[27] Even if you don't accept my argument that Linnebo is (in effect) invoking a minimalist potentionalist conception of the ordinals, he does seem clearly committed to there being a fact of the matter about what well-ordered sequences are possible — in some sense that's obviously meant to be free of any purely physical or even metaphysical limitations. And (as argued above) it's hard to see how one could understand such possibility claims, without a background notion of something like logical possibility.

And Studd explicitly introduces his notion of interpretational possibility by contrasting it with logical possibility, saying that interpretational possibility is " importantly similar and importantly different to logical necessity. Like logical necessity, it concerns possible shifts in interpretation rather than circumstance. But unlike logical necessity, the shifts in interpretation that are admissible are more closely constrained: not every logically-possible interpretation need be counted admissible". Studd also explicitly connects his notion of interpretational possibility to Linnebo's notion.

accepting new conceptual machinery (the interpretational possibility operator and a notion of interpretational essences) which we would not otherwise need. So if we take for granted that minimalist paraphrases are meaningful and get possible worlds truth conditions right (as we saw that dependence theorists seem to in §4) then considerations of conceptual parsimony favor going minimalist.

However, this argument is a bit too quick (or at least debatable). For dependence theorists could argue that the interpretational possibility operator is also independently motivated. In the remainder of this section, I'll argue that dependence theorists can (at least somewhat plausibly) argue notions are independently needed to

- best state quantifier variance/neo-carnapian philosophy of language claims.
- best explain certain kinds of social agreement in response to stipulations introducing new kinds of objects.

6.1 Needed to Best State Quantifier Variance Theses?

Let's start with assessing the idea that we need interpretational possibility to **best state neo-carnapian theses** about the possibility of talking in terms of more objects.

The basic neo-carnapian idea that we can start talking in terms of new kinds of objects objects has great appeal to many people. It just seems plausible that we could choose to start talking in terms of in-cars or an extra layer of classes, by making appropriate stipulations. Furthermore, allowing the possibility of such language change promises to help us address access worries about our knowledge of certain causally weird or wimpy objects (by saying that if we'd accepted different principles we would have speaking the truth using a variant quantifier sense). It also promises to help us resist a certain broadly Quinian on-ramp to traditional ontology (via combining the questions 'what is there?'

with presumption of a uniquely favored maximal quanitifier sense, that all other quantifier meanings are quantifier restrictions of of)[28]

However, there's a prima facie problem about how to state neo-Carnapian theses about the possibility of talking in terms of more objects sufficiently clearly to do the jobs above, while avoiding paradoxical declarations like 'there are some things I'm not now quantifying over'. And (I admit that) the interpretational possibility operator might provide one good way of non-paradoxically stating relevant neo-Carnapian claims, as argued in [23, 17][29].

However, I claim it's debatable whether that we *need* interpretational possibility to do this job. For example, work like [3, 2] argues that we can use the conditional logical possibility operator $\Diamond_{...}$ to do the same jobs. Specifically, it proposes that we can systematically give truth conditions for (portions of) the variant language we'd speak if we accepted certain kinds[30] of ontologically inflationary postulated axioms P, by saying something like the following

- For all sentences S with certain vocabulary[31] and all possible worlds w,
  - S expresses a claim (in the new language) which is true at possible world w iff $\Box_{R_1,\ldots,R_n}$[P $\rightarrow$ S] is true at w
    - where $R_1 \ldots R_n$ are antecedently understood predicates and re-

[28]Specifically, one might motivate traditional ontology as follows. First accept tame quantifier variance which says that there are contexts where e.g., it's true to say 'all the beers are in the fridge' though some in Australia are not. Then insist that such utterances must be understood as involving quantifier restrictions of some most natural unrestricted quantifier sense. If you accept all this, then asking Quine's question 'what is there?' while employing this favored, most natural (and unrestricted) quantifier sense, seems to yield an ontological question with a definite right answer and that has the kind of metaphysical weight we traditionally expect questions about ontology to have.

[29]For this modal machinery allows one to say that, for example, for any plurality of existing things xx, it would be *interpretationally possible* for there to be an additional object y which is distinct from all the xxs, rather than asserting the actual existence of such an object y.

[30]Here I omit many details constraining limits on which such axioms can be thus postulated (with a resonable expectation of changing truthvalues for ordinary sentences as below). For example[4], proposes that relevant posits need to be not only logically coherent, but metaphysically necessarily logically possible to satisfy without changing the extension of certain antecedently understood non-mathematical vocabulary.

[31]Restrictions on relevant vocabulary chosen to avoid semantic vocabulary that's independently prone to generate liar paradox.

> lations whose application the postulation is not empowered to change.

In this way (it claims) we can coherently describe variant languages more ontologically profligate than our own in sufficient detail to do traditional neo-Carnapian work like reducing access worries and resisting the above Quinean on-ramp to traditional metaphysics without paradox. So we don't need an interpretational possibility operator to do the general neo-Carnapian work referenced above.

More aggressively, minimalists might argue that the interpretational possibility operator is a particularly awkward candidate for stating claims about general neo-Carnapian language change as above. For one thing, the interpretational possibility operator only allows for quantifier expansion (not contraction). Additionally the interpretational possibility operator only allows for *some* of the the changes in word meaning that are in principle possible (e.g., changes to the meaning of "set" and "there is" which make the axiom of extensionality express a falsehood are not considered). And this narrowness arguably makes the interpretational possibility operator unsuitable for giving the most elegant and unified account of possible quantifier meaning change.

Relatedly, employing the interpretational possibility operator requires taking on certain metasemantic baggage (extending common theories of word meaning in a certain way), that's not required by the basic neo-carnapian idea above. For it involves supposing that our current concepts like SET or PHYSICAL OBJECT have an *interpretational essence* which draws a principled distinction between language change events that qualify as talking in terms of different *sets* vs. using the word "set" to mean something entirely new. Linnebo explicitly acknowledges that positing this extra element of word meaning is a novelty[17][32]

[32] Perhaps dependence theorists could defend this seeming controversial extra commitment by arguing we independently need something like their notion of interpretational essences to

[33].

Finally, all the main versions of dependence potentialism use quantifying in to express potentialist claims about how hierarchies of sets could be extended. So minimalist potentialists who avoid quantifying in (as discussed above) might object that using interpretational possibility to capture claims about neo-carnapian language change packs irrelevant and undesirable commitments (to facts about the reidentification of objects before and after ontologically inflationary language change) into these claims.

However, I should admit that whether the last commitment of using the interpretational possibility operator to articulate general carnapian philosophy of language is a benefit or a flaw is somewhat debatable. For in *Thin Objects* Linnebo positively advocates "thin realism" about ontology on which there are definite facts about reidentifying objects before and after ontologically inflationary language change. And such reidentification facts are easily expressed using the interpretational possibility operator, but perhaps cannot be expressed via

---

best account for coordinated expectations about how newly introduced and familiar types of objects.

Here's what I have in mind. Neocarnapians want to dispel access worries about apparent knowledge of abstract/causally wimpy/weird objects by saying the following. Mathematicians and sociologists etc. introducing any suitably coherent axioms/posits often can and do change quantifier meanings so their postulated axioms wind up expressing truths. However, in many such cases, questions involving new kinds of objects (e.g., is the square root of negative 1 identical to Julius Caesar? Is it a politician? Is it north of the Rhine?) seem to have definite right answers which are not logically necessitated by anything the mathematicians/sociologists introducing the new objects explicitly say.

So (dependence theorists might argue) there's independent reason to take our word meanings to include something new: defaults about what's preserved during ontologically inflationary language change events which introduce new kinds of objects. For example (they might say) there must be something about our *current* word meanings – something like an interpretational essence – which sets default expectations like

- new pure mathematical objects (by default) never have location properties, and never are physical objects, people etc.
- new kinds of composite objects whose parts are physical (by default) inherit location properties from their parts in such and such a way.

Again, I won't try to develop this project. I just want to mention it as a possible line of defense against the conceptual parsimony argument above.

[33] Relatedly (thanks to REDACTED in conversation for pressing me on this point), dependence potentialists might claim to better capture mixed sentences like 'there are more numbers than there are donkeys in Blackpool'. But minimalists can also handle these with explicit axioms for ur-elements—at the cost of greater complexity.

the conditional logical possibility operator (note that one cannot express them using the trick referenced above, since this just specifies possible-worlds truth conditions for *sentences* in the new language, not any de re facts).

So Linnebo might try to reverse my final criticism by arguing that we need the interpretational possibility operator to state **the best version** of broadly neo-carnapian philosophy of language, namely thin realism. For reasons of space I won't persuse this argument further here. I merely want to note that, if successful, such a reply might be able to block conceptual parsimony arguments for favoring minimalist potentialism.

7 Actualist Interpretation Spinning in the Void?

The final family of worries for dependence potentialism I want to discuss elaborates on familiar worry about whether potentialists can answer the question 'but how many sets are there actually?' in a principled way.

7.1 Revived arbitrariness worry

Recall that (unlike minimalists who can eliminate these terms from their logical regimentations of set theory) dependence potentialists posit a kind of double duty for terms like 'set' and 'element'. They take mathematicians' use of these words to normally express potentialist claims about how actualist hierarchies of sets could be extended. However, they also take our current set talk to have an actualist reference to some particular hierarchy of sets (perhaps allowing for some vagueness), which can be deployed by philosophers.

This threatens to revive a version of the arbitrariness worries for traditional actualist set theory in §2. We saw that actualists faced an arbitrariness worry about answering the question, 'why does our term 'set' apply to a hierarchy that goes up so high, but no further?'

But similarly, we can ask dependence theorists 'in virtue of what do we count as currently talking in terms of a hierarchy of sets with this height rather than another?'. And nothing in our practice of set talk obviously motivates associating our current concept/word 'set' with one particular height[34]. Thus one might worry that dependence theorists wind up simply replacing a metaphysical arbitrariness worry with a metasemantic arbitrariness worry.

Current dependence theorists say (to my knowledge) rather little about how our actual practice could pick out a unique referent for our set talk. For example, does the fact that mathematicians treat pairing as an axiom as clearly prevent them from being (in some sense) acceptably interpreted as referring to an actualist structure $V_{\omega+36}$ that doesn't satisfy pairing?

- On one hand, you might expect that acceptable precisifications of the actualist reference of our 'set' talk must satisfy ZFC, because the actualist reading of our mathematical language needs to make normal mathematical talk come out true (though potenitalist interpretations are better/more explanatory).

- On the other hand, you might say no, the powerset axiom may expresses a falsehood (on the actualist reading), but this is OK because it is true

[34]I take actual dependence theorists' claims about interpretational possibility to concern what's possible for *our current set concept* (not just what's possible for someone engaged in a hypothetical actualist hierarchy of sets practice). There could (in principle) be an explicitly actualist practice of "set" talk, which is quite different from ours and did clearly talk in terms of some definite number of layers of sets. (Imagine someone explicitly starting out stipulating that there are no sets, and then repeating Linnebo's formula for adding a layer of sets via Fregean abstraction). And one could develop a version of dependence potentialism which appeals to facts about how people speaking this different language using the letters "set" could engage in ontologically inflationary language change.

However, I take it that such a version of dependence theoretic potentialism would face a very strong version of the worry about introducing intuitively irrelevant content into the explication of mathematical claims above. For surely interpreting set theory as making claims about logical possibility of structures satisfying approximately traditional set axioms better preserves the current/intuitive meaning of set claims, than interpreting set claims as talking about possible changes to the imaginary language practice described above.

Hence, it seems to me that dependence theorists must say that something about our actual *current* practice (and reference magnetic natural kinds and other etc.) provides "set" with a determinate extension — to whatever extent that it has one.

on the potentialist reading (which explains/justifies what mathematicians say). Given Linnebo and Studd's talk of acceptably adding a single layer to whatever iterative hierarchy one is currently talking in terms of, I'd guess they'd prefer this option. But, isn't it counterintutive to claim mathematicians who treat pairing as clearly true are acceptably interpreted as referring to an actualist structure $V_{\omega+36}$ that doesn't satisfy pairing?

Thus dependence theorists seem forced to posit a fact about how tall a hierarchy we are currently talking in terms of, which is neither

- motivated by something in our conception of the sets or mathematical practice (e.g., axioms we accept)
- a plausible natural kind in the Stalnakerian sense (as we might say that the notion of 'all possible ways of choosing' is/the intended interpretation of second order quantification is)[35]

Admittedly, dependence potentialists have some options for answering this arbitrariness worry that actualists lack (since they don't need the actualist interpretation for set talk to make all normally assertable set claims come out true). For example, they could say that the current actualist reference of our set talk is completely indeterminate (if they are willing to accept the controversial claim that there can be vagueness about what objects exist, not just about how properties apply[22])[36]. Or they could say that there's a brute metasemantic fact/law that determines what hierarchy of sets we are talking about at any

[35] In principle, an actualist could say that there's a notion of absolute infinity (the intended height of the hierarchy of sets) that's a referential natural kind, just as others would make this claim for a favored notion of full second order quantification or conditional logical possibility. However, I don't know of anyone who actually takes this route. And appeal to such an intrinsically favored notion of absolute infinity seems deeply contrary to the spirit of potentialism, hence not available to the dependence potentialist.

[36] But this complete indeterminacy would seem to be quite unlike more familiar indeterminacy about the boundaries of the concept red, or which electrons are part of a cat's tail at a given moment. Hence one would like any dependence theoretic potentialist who takes this route to say more about what kind of indeterminacy they have in mind and how it fits into a larger philosophy of language and metaphysics.

given moment, without our guidance from anything in our current practice, in a mysterious way. For example, maybe the actualist reference of ‘set’ is always to a structure V_n, where n is the number of seconds that have passed since the iterative hierarchy conception of sets was first considered or the number of times some mathematician has uttered a sentence containing the word ‘set’. However, both moves seem a bit drastic.[37]

7.2 Violating metasemantic intuitions

In addition to the above problem shared with actualists, dependence theorists may also face a worry about violating intuitive metasemantic principles connecting use to meaning.

The basic worry is this. According to the dependence theorist, all mathematicians’ set-theoretic talk is implicitly modalized (expressing potentialist claims). But if this is the case, how can this modalized usage do anything to settle the interpretation of ‘set’ or ‘every’ in unmodalized (non potentialist) contexts?

Relatedly, one might see dependence theorists’ proposal as violating the following (admittedly controversial) claim about literal and metaphorical meaning.

> Live vs. Dead Metaphors principle: A term can only have a certain literal (as opposed to metaphorical) meaning in some language

[37]Minimalist potentialists avoid this arbitrariness worry. For their paraphrases can entirely eliminate terms like ‘set’ and ‘element’, so they don’t have to admit questions like ‘how many sets are there actually?’ as meaningful.

One *can* still ask the minimalist potentialist ‘How tall is the largest actually existing structure satisfying the ZFC axioms?’. However, there is no danger of needing to invoke an arbitrary stopping point (or massive indeterminacy) when answering this question. For the minimalist can simply say the answer to this question will reflect how many non-set objects our current language talks in terms of (what kinds of objects non-set mathematical objects, physical objects, sociological objects etc), and how many of each of these objects there are. For example, if we employ familiar practices for talking in terms of real numbers, then uncountably many objects actually exist and can take part in the largest model of ZFC. Also note that questions about what models there actually are forced upon us for reasons independent of set-theoretic potentialism, while questions about what sets there are only forced upon us once we’ve adopted set-theoretic potentialism. (Thanks to an anonymous reviewer for suggesting this point).

> if competent speakers of that language sometimes do (or at least can) use the term in line with the literal meaning. Once a metaphor is so dead that people would have no idea how to assess whether something falls under the (old) literal extension of the term (including by appeal to relevant experts), you no longer have a metaphor; the current literal meaning of the word just is the old metaphorical meaning.

An analogous metasemantic principle about actualist reference has some appeal.

> Live vs. Dead Actualist Extension Principle: A term can only take on an actualist meaning and extension in a language if competent speakers of that language sometimes do (or at least can) use the term in line with the literal meaning (i.e., committing themselves to what's true on the actualist reading of the term, and reasoning in a sufficiently accurate/reliable way to be charitably interpreted as so doing).

Yet dependence theorists' attributions of actual meaning and reference seem to violate this principle. For (on the dependence theorist view) set theorists always use set talk in line with the potentialist translation, not the supposed actualist meaning (which, e.g., may have the sets stopping at a successor stage so the powerset axiom fails). We don't have any second practice for evaluating claims about what sets literally/actually exist, which could motivate attributing an actualist extension to 'set' (and be interpreted as reliably speaking the truth about this notion). Thus claiming that their actual uses of the terms 'set' and 'element' secure an actualist reference (whether precise or vague) can seem inadequately motivated [38][39].

7.3 Studd on Mathematicians' Actualist Reference

In *Everything, More or Less: A Defence of Generality Relativism* [23] Studd makes a sketch of how mathematicians' behavior could *change* the range of acceptable actualist interpretations of their set talk. This is the most explicit proposal I'm aware of, for how our actual use is supposed to determine an actualist reference for our set talk. So one might hope it would provide a way of answering the arbitrariness and undermotivation by use worries above.

First [23], Studd imagines people who start out speaking a language Q that 'talks in terms of' a certain hierarchy of sets and *knowingly* attempt to and develop a new language E which talks in terms of extra sets. This splinter group could (Studd suggests) achieve their ends by accepting certain schemas for reasoning from claims in the old language Q to claims in the new language E, and vice versa, including the following.

$$Q : things(vv) \Rightarrow E : thing(\{vv\})$$

$$Q : things(vv), Q : v \prec vv \Rightarrow E : v \in \{vv\}$$

$$Q : things(vv), E : v \in \{vv\} \Rightarrow Q : v \prec vv$$

[38] Admittedly, Linnebo does say that *dependence potentialist philosophers* sometimes use set and $\in$ with their actualist meaning. But (contra the live vs. dead actualist extension principle above) such philosophers seem to know or say very little about what the actual height of the hierarchy of sets is – or how we could even begin to determine it. One doesn't have to accept any kind of bold verificationism to find this extreme disconnect between supposedly determine meaning facts and all assessment practices troubling (especially if claims about certain possible heights for the hierarchy of sets being reference magnetic natural kinds are rejected).

[39] Perhaps this worry can be somewhat softened by emphasizing that potentialist paraphrases aim at Carnapian explication rather than unveiling the deep structure of our exact current concepts, but I'm not sure.

Intuitively, these schemas embody the idea that each plurality vv of objects quantified over in the old language Q is supposed to form a set in the new language. So, Studd suggests, charitable interpretation requires taking the quantifiers in the new language E to range over strictly more objects than quantifiers in their original Q. By adopting such inference rules, our splinter group could start talking in terms of new objects.

Clearly, we do not embrace any such schemas. However, Studd proposes that unknowingly accepting inconsistent axioms of set theory (including the ones below) can give rise to similar kind of expansionary quantifier meaning change:

$$things(vv) \Rightarrow thing(\{vv\})$$

$$things(vv), v \prec vv \Rightarrow v \in \{vv\}$$

$$things(vv), v \in \{vv\} \Rightarrow v \prec vv$$

The above inference principles are inconsistent in a familiar Russellian way[40]. So Studd suggests it's charitable to interpret speakers who accept these principles as undergoing (unwitting) language change analogous to the switch from Q to E envisaged above, rather than saying something inconsistent.

Studd puts this proposal forward as, "[the] basis for an idealized account of universe expansion applicable to the ordinary English speaker." However, I want to raise two worries.

[40] They let you infer that, for any plurality of things vv, there's a set $\{vv\}$ whose elements are exactly the objects v in this plurality vv (written $v \prec vv$). But accepting the existence of this set (together with normal plural comprehension principles saying that, for any ϕ, there's a plurality vv of the objects such that ϕv) lets you derive the existence of the Russell set and hence contradiction.

First, contemporary people who are aware of Russell's paradox seemingly don't have the paradoxical inference dispositions above. So it's not clear how the considerations in Studd's second story could apply to interpreting us[41]

Second, there's an obvious question about when/how often speakers are supposed to go through language change events Studd proposes. Suppose I have the inconsistent inference dispositions Studd mentions and don't think about set theory for an hour. How many times should a charitable interpreter take my language to have changed during that time? Insofar as standing dispositions to make certain inferences (or to regard failure to make accept them when suggested as irrational) drive the above charitable interpretation, it is hard to see how one could give any non-arbitrary answer to this question.

Thus I don't think Studd's story provides much help with the problem of apparent commitment to very arbitrary facts about how the current actualist extension of our term 'set' gets fixed, or massive indeterminacy. Nor does it dispell uneasiness about positing an actualist reference for our set talk that

[41] An anonymous reviewer helpfully suggested that Studd's story just requires, "a disposition to take arbitrary pluralities of currently available objects to provide a basis for introducing talk about yet more sets", and that we do have such a disposition. Yet I think this suggestion raises a bit of an interpretive dilemma.

On one hand, the idea might be that ordinary mathematicians will allow we could think of the hierarchy of sets (in our current sense) as being extended by some proper classes, and then redefine "set" to apply to that larger plurality – to both sets (in our original sense) and proper classes. In this case, we don't have any paradoxical dispositions that could drive an expanded interpretation of the existential quantifier. But adopting progressively more expansive interpretations of actual mathematicians set talk for because of *this* disposition would seem to be treating an instance of a general phenomenon (acceptance that we could explicitly redefine the word "set" to apply to more objects, just as we could redefine the term "apple") quite differently from others in an unmotivated way. Accordingly appeal this disposition of actual mathematicians doesn't seem capable of motivating Studd's interpretive extreme measures.

On the other hand, the intended claim might be that mathematicians are disposed say things that would be *paradoxical* if we interpret 'set' as having a persisting actualist reference. In this case, we'd have a clear distinctive feature of set-theoretic talk that could motivate special treatment (like the attribution of unwitting quantifier meaning change Studd suggests). But this version of the claim is implausible. Mathematicians post-Russell's paradox (unlike the people in Studd's second thought experiment) have settled practices which carefully avoid letting one derive claims that would imply contradiction if we take the extension of 'set' to remain fixed. Unlike accepters of naive set theory, they reliably refuse to go from identifying arbitrary pluralities of objects to belief in a set that collects exactly this plurality of objects. For example, they are disposed to accept that there's a set of even numbers, but not that there's a set of all sets, or a set of all singletons etc.

seems to be jobless and spinning in the void.

8 Conclusion

In this paper, I've presented a three kinds of argument for favoring minimalist over dependence theoretic versions of potentialist set theory (and considered possible strategies for responding to them). I don't claim these arguments settle the family feud between minimalist and dependence theoretic potentialisms. I only hope to have advanced the debate (and presented the background, state of play and larger philosophical significance of this element of potentialist inside baseball in a way that might interest a larger audience)[42] .

References

[1] Neil Barton. Forthcoming.

[2] Sharon Berry. Chalmers, Quantifier Variance and Mathematicians' Freedom. In Alessandro Torza, editor, *Quantifiers, Quantifiers, and Quantifiers. Themes in Logic, Metaphysics and Language. (Synthese Library vol 373)*, pages 191–219. Springer, 2015.

[3] Sharon Berry. Modal Structuralism Simplified. *Canadian Journal of Philosophy*, 48(2):200–222, 2018.

[4] Sharon Berry. *A Logical Foundation for Potentialist Set Theory.* Cambridge University Press, Cambridge, 2022.

[42]For example, one thing I could imagine shifting the scales in favor of dependence theory is a certain kind of progress regarding the liar paradox. It is generally agreed that paradoxes concerning the height of the hierarchy of sets and the liar paradox seem to have a lot in common, so it would be appealing to treat them similarly, if possible. Linnebo and Studd see a close connection between their proposals and Kripke's approach to the liar paradox (with determinate truth percolating up from some sentences to other sentences). So if Kripkian approaches to truth could be both vindicated and shown to have a uniquely close relationship to dependence theoretic potentialism, this would improve the appeal of dependence potentialism. But, to my knowledge, no such results have been achieved.

[5] George Boolos. Nominalist Platonism. *Philosophical Review*, 94(3):327–344, 1985.

[6] Hartry Field. *Realism, Mathematics & Modality*. Blackwell Oxford, 1989.

[7] Hartry H. Field. *Saving Truth from Paradox*. Oxford University Press, March 2008. Published: Paperback.

[8] Kit Fine. Our Knowledge of Mathematical Objects. In Tamar Szabó Gendler and John Hawthorne, editors, *Oxford Studies in Epistemology*, pages 89–109. Oxford University Press, 2005.

[9] Kit Fine. Relatively Unrestricted Quantification. In Agustín Rayo and Gabriel Uzquiano, editors, *Absolute Generality*, pages 20–44. Oxford University Press, 2006.

[10] Mario Gómez-Torrente. A Note on Formality and Logical Consequence. *Journal of Philosophical Logic*, 29(5):529–539, October 2000.

[11] William H. Hanson. Actuality, Necessity, and Logical Truth. *Philosophical Studies*, 130(3):437–459, 2006.

[12] Geoffrey Hellman. *Mathematics Without Numbers*. Oxford University Press, USA, 1994.

[13] Geoffrey Hellman. Structuralism Without Structures. *Philosophia Mathematica*, 4(2):100–123, 1996.

[14] Øystein Linnebo. Pluralities and Sets. *Journal of Philosophy*, 107(3):144–164, 2010.

[15] Øystein Linnebo. The Potential Hierarchy of Sets. *Review of Symbolic Logic*, 6(2):205–228, 2013.

[16] Øystein Linnebo. Putnam on mathematics as modal logic. In Geoffrey Hellman and Roy T. Cook, editors, *Putnam on Mathematics and Logic*. Springer Verlag, Berlin, 2018.

[17] Øystein Linnebo. *Thin Objects*. Oxford: Oxford University Press, 2018.

[18] Charles Parsons. What is the Iterative Conception of Set? In Robert E. Butts and Jaakko Hintikka, editors, *Logic, Foundations of Mathematics, and Computability Theory: Part One of the Proceedings of the Fifth International Congress of Logic, Methodology and Philosophy of Science, London, Ontario, Canada-1975*, pages 335–367. Springer Netherlands, Dordrecht, 1977.

[19] Charles Parsons. *Mathematical Thought and Its Objects*. Cambridge University Press, December 2007. Published: Hardcover.

[20] Hilary Putnam. Mathematics Without Foundations. *Journal of Philosophy*, 64(1):5–22, 1967.

[21] Hilary Putnam. Models and Reality. *Journal of Symbolic Logic*, 45(3):464–482, 1980.

[22] Theodore Sider. Against Vague Existence. *Philosophical Studies*, 114(1-2):135–146, 2003.

[23] James P. Studd. *Everything, More or Less: A Defence of Generality Relativism*. Oxford University Press, 2019.

[24] Hao Wang. A Logical Journey. From Gödel to Philosophy. *Philosophy*, 73(285):495–504, 1998.

[25] Timothy Williamson. Epistemological Consequences of Frege Puzzles. *Philosophical Topics*, 49(2):287–320, 2021.

[26] Ludwig Wittgenstein. *Philosophical Investigations*. Prentice Hall, 1999.