The Review of Symbolic Logic
Volume 15, Number 4, December 2022

AGAINST CUMULATIVE TYPE THEORY

TIM BUTTON
University College London
and

ROBERT TRUEMAN
University of York

Abstract. Standard Type Theory, STT, tells us that $b^n(a^m)$ is well-formed iff $n = m + 1$. However, Linnebo and Rayo [23] have advocated the use of Cumulative Type Theory, CTT, which has more relaxed type-restrictions: according to CTT, $b^\beta(a^\alpha)$ is well-formed iff $\beta > \alpha$. In this paper, we set ourselves against CTT. We begin our case by arguing against Linnebo and Rayo's claim that CTT sheds new philosophical light on set theory. We then argue that, while CTT's type-restrictions are unjustifiable, the type-restrictions imposed by STT are justified by a Fregean semantics. What is more, this Fregean semantics provides us with a principled way to resist Linnebo and Rayo's Semantic Argument for CTT. We end by examining an alternative approach to cumulative types due to Florio and Jones [10]; we argue that their theory is best seen as a misleadingly formulated version of STT.

Standard Type Theory, STT, tells us that $b^n(a^m)$ is well-formed iff $n = m + 1$. However, Linnebo and Rayo [23] have advocated the use of Cumulative Type Theory, CTT, which has more relaxed type-restrictions: according to CTT, $b^\beta(a^\alpha)$ is well-formed iff $\beta > \alpha$. Other philosophers, including Williamson [44, pp. 237–238], Krämer [21, p. 527], and Florio and Jones [10], have since expressed sympathy for cumulative types.

We set ourselves against cumulative type theory. We begin our case by arguing against Linnebo and Rayo's claim that CTT sheds new philosophical light on set theory: in Section 2 we highlight some important mathematical differences between CTT and set theory, and in Section 3 we explore the philosophical consequences of these differences. Then, in Section 4, we push our case against CTT further, by arguing that the type-restrictions it imposes are unjustifiable. This marks an important difference between CTT and STT: a *Fregean* semantics justifies STT's type-restrictions (see Section 5), and this Fregean semantics also provides us with a principled way to resist Linnebo and Rayo's Semantic Argument for CTT (see Section 6). We end, in Section 7, by examining an alternative approach to cumulative types due to Florio and Jones [10]; we argue that their theory is best seen as a misleadingly formulated version of STT.

Received: April 16, 2020.
2020 *Mathematics Subject Classification*: 03B38, 03A05.
Key words and phrases: cumulative type theory, set theory, Fregean semantics, absolute generality.

doi:10.1017/S1755020321000435

§1. Formal type theories.

We start by outlining the formalisms of STT and CTT. For simplicity of exposition, in this paper we focus on *monadic* type theories. (We also only consider un-ramified type theories.)

1.1. STT.

STT has a countable infinity of types, $0 \leq n < \omega$. The type of a term is indicated with a numerical superscript: a^n is a type n term. We have constants and variables of every type. Atomic formulas are made by combining a type $n{+}1$ term with a type n term: $b^n(a^m)$ is well-formed iff $n = m + 1$. Intuitively, $b^{n+1}(a^n)$ applies a type $n{+}1$ entity to a type n entity, where an entity is of type n iff it is a value of a type n variable; however, exactly what this intuitive gloss amounts to will depend on your preferred interpretation of the types (see Sections 4 and 5).

Every type of variable can be bound by quantifiers. We here present the rules for $\forall$; the rules for $\exists$ are the obvious duals. For all types n, the following inferences are licensed, provided that (i) all expressions are well-formed and (ii) b^n does not occur in any undischarged assumptions on which $\phi(b^n)$ depends:

$$\frac{\forall x^n \phi(x^n)}{\phi(a^n)}\ \forall\mathrm{E}^n \qquad\qquad \frac{\phi(b^n)}{\forall x^n \phi(x^n)}\ \forall\mathrm{I}^n.$$

To ensure that each level of the type hierarchy is well-populated, we have the following scheme, for each type n:

> *STT-Comprehension.* $\exists z^{n+1} \forall x^n (z^{n+1}(x^n) \leftrightarrow \phi(x^n))$, whenever $\phi(x^n)$ is well-formed and does not contain z^{n+1}.

STT has the usual stock of logical devices: quantifiers, connectives, and the identity sign, $=$. The identity sign can be flanked by a pair of terms of any type, but they must be terms of the same type; so $a^m = b^n$ is well-formed iff $m = n$. Identity is governed by the following scheme, for each type n:

$$x^n = y^n \text{ iff } \forall z^{n+1}(z^{n+1}(x^n) \leftrightarrow z^{n+1}(y^n)).$$

We can treat this as an axiom scheme or an explicit definition. But, either way, $x^n = y^n$ is *typically ambiguous*: there is not a single identity relation that applies across all the types, but a different relation for each type.

1.2. CTT.

Linnebo and Rayo [23] ask us to consider an alternative, cumulative, type theory, CTT. This type theory was formally developed by Degen and Johannsen [7]. (We discuss a different approach to cumulation, due to Florio and Jones [10], in Section 7.) The basic thought behind CTT is that the entities cumulate as you ascend through the types. Let us see how this is implemented.

First, CTT relaxes STT's syntax. In STT, $b^n(a^m)$ is well-formed iff $n = m + 1$. But, if the types cumulate, then everything at level 0 reappears at level 1; so, since $c^2(a^1)$ is meaningful, $c^2(a^0)$ should be too. More generally, CTT allows that $b^\beta(a^\alpha)$ is well-formed iff $\beta > \alpha$. And note that we use 'α' and 'β' rather than 'n' and 'm' here: if the types cumulate, we will want to be able to consider transfinite types, and so we must allow ourselves a transfinite stock of type-indices. (One obvious way to do this is to stipulate that the type-indices are von Neumann's ordinals, but the only important constraint is that the type-indices be well-ordered.)[1]

[1] Cf. Linnebo and Rayo [23, p. 294, 24, pp. 178–179] on 'definite' collections of languages and alternative 'labels'. For readability, we use standard ordinal notation in this paper, but this

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Second, CTT has rather permissive inference rules for quantifiers. (Again, we only outline the rules for $\forall$.) For all types $\beta \geq \alpha$, the following inferences are licensed, provided that (i) all expressions are well-formed and (ii) b^β does not occur in any undischarged assumption on which $\phi(b^\beta)$ depends:[2]

$$\frac{\forall x^\beta \phi(x^\beta)}{\phi(a^\alpha)}\ \forall\mathrm{E}^\beta_\alpha \qquad\qquad \frac{\phi(b^\beta)}{\forall x^\alpha \phi(x^\alpha)}\ \forall\mathrm{I}^\beta_\alpha.$$

These rules are intuitively sound, given the idea of cumulation: every type α entity is a type $\beta \geq \alpha$ entity too; so if ϕ holds of every type β entity, then ϕ holds of each type α entity.

Third, to ensure that each successor-level of the type hierarchy is well-populated, CTT has a Comprehension scheme, for each type α:[3]

> *CTT-Comprehension.* $\exists z^{\alpha+1}\forall x^\alpha(z^{\alpha+1}(x^\alpha) \leftrightarrow \phi(x^\alpha))$, whenever $\phi(x^\alpha)$ is well-formed and does not contain $z^{\alpha+1}$.

Fourth, CTT has an infinitary inference rule for each limit type λ:[4]

$$\frac{\forall x^\alpha \phi(x^\alpha),\text{ for all } \alpha < \lambda}{\forall x^\lambda \phi(x^\lambda)}\ \mathit{Limit}^\lambda.$$

Intuitively, this guarantees that nothing essentially ‘new’ happens at limit types, so that any type λ entity is an entity of some type $\alpha < \lambda$.

So far, we have identified entities across types quite freely. However, Linnebo and Rayo [23, pp. 281–283] retain the rule that a strict identity claim, $x^\alpha = y^\beta$, is well-formed iff $\alpha = \beta$. To deal with cross-type identity, they explicitly define a new sign, $\equiv$, for any types α and β and where $\gamma = \max(\alpha, \beta) + 1$:[5]

$$a^\alpha \equiv b^\beta \text{ iff}_{\text{df}}\ \forall x^\gamma(x^\gamma(a^\alpha) \leftrightarrow x^\gamma(b^\beta)).$$

This definition is typically ambiguous: it defines different relations for different α and β. But all of these relations behave like identity: if $\phi(a^\alpha)$ and $\phi(b^\beta)$ are both well-formed, then $\phi(a^\alpha)$ and $a^\alpha \equiv b^\beta$ together entail $\phi(b^\beta)$.[6] Now we can prove the following theorem scheme, for all $\alpha \leq \beta$:[7]

> *Type-Raising Scheme.* $\forall x^\alpha \exists y^\beta\ x^\alpha \equiv y^\beta$

So, if $\alpha \leq \beta$, then every type α entity is a type β entity, in the sense of ‘is’ expressed by $\equiv$.

is easily eliminable; e.g., ‘$\alpha + 1$’ can be parsed as ‘the next index after α’, and ‘ω’ as ‘the first limit index’.

[2] These are the obvious natural-deduction versions of Degen and Johannsen’s [7, p. 149] sequent-calculus rules. Linnebo and Rayo [23, p. 288] are not specific on the rules they adopt, but [23, 282n20] appeal to a result from Degen and Johannsen [7, p. 150] which uses these rules.

[3] Degen and Johannsen [7, p. 149] and Linnebo and Rayo [23] offer a variant formulation, using λ-abstraction.

[4] Degen and Johannsen [7, 153] and Linnebo and Rayo [23, p. 288].

[5] Degen and Johannsen [7, p. 149] draw no distinction between $=$ and $\equiv$.

[6] For a proof, see Lemma 2 of Appendix A.

[7] Linnebo and Rayo [23, p. 288] take Type-Raising as an axiom scheme; we prove it in Lemma 1 of Appendix A.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

We also provide another (typically ambiguous) explicit definition, where $\gamma = \max(\alpha, \beta) + 1$:[8]

$$a^\alpha \varepsilon b^\beta \text{ iff}_{\text{df}} (\exists x^\gamma \equiv b^\beta) x^\gamma(a^\alpha).$$

This membership-like notion applies b^β to a^α, but is well-formed for every α and β. So $a^\alpha \varepsilon b^\beta$ allows us to simulate $b^\beta(a^\alpha)$, even when $\alpha \geq \beta$.

If we provide no further axioms, though, then ε can be ill-founded. To rule this out, we lay down two final schemes, for all α, β:[9]

Type-Founded. $\forall a^\alpha \forall b^{\beta+1}(a^\alpha \varepsilon b^{\beta+1} \rightarrow \exists x^\beta\, a^\alpha \equiv x^\beta)$
Type-Base. $\forall x^0 \forall y^\alpha\, y^\alpha \not\varepsilon x^0$

This completes the list of axioms and inference rules for CTT.

It is worth making a brief observation about syntax. In moving from STT to CTT, we are asked to relax STT's syntax: $b^\beta(a^\alpha)$ is well-formed iff $\beta > \alpha$. There is an obvious way to relax this further, whilst retaining a typed theory: allow that $b^\beta(a^\alpha)$ is well-formed for *any* α and β. However, this further relaxation would have no real effect. As just noted, CTT can simulate $b^\beta(a^\alpha)$ using the formula $a^\alpha \varepsilon b^\beta$, where the latter is defined using the more stringent type-restrictions. Consequently, we can be largely indifferent on whether to use the *stringent* type-restrictions, so that $b^\beta(a^\alpha)$ is well-formed iff $\beta > \alpha$, or the more *liberal* type-restrictions, so that $b^\beta(a^\alpha)$ is well-formed for any α and β. In what follows, we will tend to adopt the stringent type-restrictions, but we will revisit this in Section 4.

For each type-index τ, the theory CTT$^\tau$ has a countable infinity of distinct variables of every type $< \tau$, and no variables of any type $\geq \tau$. We refer to the cumulative type theories in general as 'CTT', using 'CTT$^\tau$' with the superscript when it is important to pay attention to the bound.

§2. The sets-from-types theorem.

Degen and Johannsen [7] and Linnebo and Rayo [23] note that CTT interprets an iterative set theory. In this section, we present a strengthened version of their formal results. We discuss its philosophical significance in Section 3. For ease of exposition, we will consider set theories without urelements (and similar type theories); we could accommodate urelements if we liked, but it would complicate our discussion without adding any real insight.

2.1. The interpretation.

We will focus on a 'pure' version of CTT$^\tau$, which we call CTT$^\tau_{\text{p}}$. This augments CTT$^\tau$ with principles guaranteeing that there is exactly one type 0 entity, and that coextensive entities at higher-types are identical. (For details, see Appendix B.2.) The set theory that CTT$^\tau_{\text{p}}$ can interpret is Zr, i.e., *Z*ermelo set theory

[8] Degen and Johannsen [7, p. 151] and Linnebo and Rayo [23]. *Notation*: $(\exists x^\gamma \equiv b^\beta)\phi$ abbreviates $\exists x^\gamma(x^\gamma \equiv b^\beta \wedge \phi)$; similarly, $(\forall x^\gamma \equiv b^\beta)\phi$ abbreviates $\forall x^\gamma(x^\gamma \equiv b^\beta \rightarrow \phi)$; and similarly for other two-place infix predicates.

[9] In Appendix A, we prove that Type-Founded and Type-Base are independent from the axioms given so far. Linnebo and Rayo [23, p. 289] provide a version of *Type-Base*, but no version of *Type-Founded* (though they clearly want some such principle; see [23, p. 283n.22]). Degen and Johannsen [7] tackle this slightly differently; see the start of Appendix B. With these principles, we can establish that: if $\alpha < \beta$ then $a^\alpha \varepsilon b^\beta$ iff $b^\beta(a^\alpha)$; if $\alpha \geq \beta$ and α is minimal for a^α and β is minimal for b^β, then $a^\alpha \not\varepsilon b^\beta$. (Here, we say that γ is minimal for c^γ iff $\forall x^\delta\, c^\gamma \not\equiv x^\delta$ for all $\delta < \gamma$.)

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

together with the principle that the sets are arranged into well-ordered *ranks*. This theory omits Replacement, and so is strictly weaker than ZF. (For more details, see Appendix B; note that ZF = Zr + Replacement.)

To interpret Zr with $\text{CTT}_{\text{p}}^{\tau}$, we first define a translation. For each Zr-formula ϕ, let $\phi^{(\kappa)}$ be the formula which results by replacing each '$\in$' with 'ε', each '$=$' with '$\equiv$', and superscripting each variable with κ. For example, the Axiom of Powersets$^{(\kappa)}$ is:

$$\forall a^{\kappa}\exists b^{\kappa}\forall x^{\kappa}(x^{\kappa}\ \varepsilon\ b^{\kappa} \leftrightarrow (\forall v^{\kappa}\ \varepsilon\ x^{\kappa})v^{\kappa}\ \varepsilon\ a^{\kappa}).$$

Now we can prove the following result (see Appendix B, Theorem 13):[10]

The Sets-from-Types Theorem. $\text{CTT}_{\text{p}}^{\tau} \vdash \text{Zr}^{(\kappa)}$, for any limit $\kappa > \omega$ with $\kappa + 2 < \tau$.

Otherwise put: $\text{CTT}_{\text{p}}^{\tau}$ proves the translations of all theorems of Zr.

2.2. Differences between Zr and Zr$^{(\kappa)}$.

The proof of the Sets-from-Types Theorem involves establishing a tight association between two notions: an entity's *type*, as in $\text{CTT}_{\text{p}}^{\tau}$ and $\text{Zr}^{(\kappa)}$, and a set's *rank*, as in Zr. This sort of connection leads Linnebo and Rayo [23, p. 289] to claim that 'there is no deep *mathematical* difference between the ideological hierarchy of type theory and the ontological hierarchy of set theory'.

Whether to describe them as 'deep' may be a matter of taste, but it is worth noting three mathematical differences between Zr, on the one hand, and $\text{CTT}_{\text{p}}^{\tau}$ and $\text{Zr}^{(\kappa)}$, on the other.[11] We summarize the differences in the following table:

	Zr	$\text{CTT}_{\text{p}}^{\tau}$ and $\text{Zr}^{(\kappa)}$
(1)	Ranks defined internally	Types supplied externally
(2)	Untyped variables	Typed variables
(3)	Recursively axiomatized	Arithmetically complete

We will now explain these three differences.

Concerning (1). The notion of *rank* is explicitly defined within Zr itself, much as it is within ZF.[12] By contrast, the notion of *type* is metatheoretic for both $\text{CTT}_{\text{p}}^{\tau}$ and $\text{Zr}^{(\kappa)}$. Every variable carries a type-index, and these type-indices are supplied *externally*. Indeed, when we take the very first step of describing the syntax of CTT^{τ}, we assume as given all the type-indices $< \tau$.

Concerning (2). Zr is essentially *untyped*. It has exactly one kind of variable, which ranges over all sets of all ranks. By contrast, every variable in $\text{CTT}_{\text{p}}^{\tau}$ and $\text{Zr}^{(\kappa)}$ carries a type-index, and $\text{CTT}_{\text{p}}^{\tau}$'s quantifier rules indicate that type α variables range only over entities of type $\leq \alpha$. These theories have no *untyped* variables; that is, they have no variables which range over all entities of *all* types. (Note that, despite our use of the phrase 'ranging over', this difference shows up at the level of the formal theories,

[10] This extends Degen and Johannsen's [7, sec. 4.1] results concerning Z. Linnebo and Rayo [23 cover only Z without Foundation. The bound $\kappa + 2 < \tau$ is needed as $a^{\kappa}\ \varepsilon\ b^{\kappa}$ abbreviates $\exists x^{\kappa+1}(\forall z^{\kappa+2}(z^{\kappa+2}(x^{\kappa+1}) \leftrightarrow z^{\kappa+2}(b^{\kappa})) \wedge x^{\kappa+1}(a^{\kappa}))$.

[11] Linnebo and Rayo [23, p. 284, 289] mention differences (1) and (2) themselves, but they do not mention (3).

[12] Indeed, it is definable within LT; see Appendix B.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

prior to interpretation. Indeed, none of the differences depend upon any semantic considerations.)

Concerning (3). Clearly, Zr is recursively axiomatized (see Appendix B.1). However, neither $\mathrm{CTT}_{\mathrm{p}}^{\tau}$ nor $\mathrm{Zr}^{(\kappa)}$ is recursively axiomatizable, thanks to the intrinsically infinitary Limit-rules. Indeed, Limit^{ω} makes these theories arithmetically complete, since it simulates Hilbert's ω-rule.[13]

2.3. Mathematical foundations. We will now explain why these three differences are mathematically significant. In brief: the differences show that Zr is expressively richer but deductively weaker than $\mathrm{Zr}^{(\kappa)}$; this makes Zr much more suitable as a framework for considering mathematical foundations.

Differences (1) and (2) show that Zr is *expressively richer* than $\mathrm{Zr}^{(\kappa)}$. To see this, consider how we might formulate questions about the height of a hierarchy. In the case of Zr, we might ask a specific question like: *Should we countenance a strongly inaccessible rank*? That question is formulated within the *object language* of Zr, and this is possible because Zr's untyped variables range over all the sets, whatever their rank. So, whilst Zr does not *settle* whether there are any sets of strongly inaccessible rank, it does allow us to *formulate* the claim that there are, and tells us that any such sets obey Extensionality and Separation (for example). In the case of $\mathrm{CTT}_{\mathrm{p}}$, the analogous question about the height of a type-hierarchy would be: *Should we countenance a strongly inaccessible type-index*? But this question is, of course, formulated within a *metalanguage*. After all, each $\mathrm{CTT}_{\mathrm{p}}^{\tau}$ has variables of all and only the types $< \tau$, and $\mathrm{Zr}^{(\kappa)}$ has variables of all and only the types $\leq \kappa + 2 < \tau$,[14] so neither theory allows us to formulate questions about entities of type τ; they literally lack the vocabulary for doing so.

Difference (3), however, shows that Zr is *deductively weaker* than $\mathrm{Zr}^{(\kappa)}$. This is obvious—one is arithmetically complete, the other is not—but let us draw out a couple of consequences. The Sets-from-Types Theorem tells us that $\mathrm{CTT}_{\mathrm{p}}^{\tau}$ interprets Zr. However, this interpretation is not *faithful*, i.e., some non-theorems of Zr become theorems of $\mathrm{Zr}^{(\kappa)}$ under interpretation; nor is the interpretation *mutual*, i.e., Zr cannot interpret $\mathrm{Zr}^{(\kappa)}$.[15]

This combination of expressive richness with deductive weakness makes Zr much more suitable as a framework for mathematical foundations than $\mathrm{Zr}^{(\kappa)}$ or $\mathrm{CTT}_{\mathrm{p}}^{\tau}$.[16] Concerning expressive strength: if our hierarchy is to serve as any kind of mathematical foundation, then questions about the height of the hierarchy will be of pressing importance; but only Zr provides a suitable framework for raising such questions. Concerning deductive weakness: any adequate foundational theory must be recursively

[13] Assuming $\kappa > \omega$ and $\kappa + 2 < \tau$. *Sketch*. Using the Sets-from-Types Theorem, use $\mathrm{CTT}_{\mathrm{p}}^{\tau}$ to develop $\mathrm{Zr}^{(\omega+\omega)}$. In $\mathrm{Zr}^{(\omega+\omega)}$, define $\mathbb{N}$ as the set of finite von Neumann ordinals, and define $+$ and $\times$ as usual. Suppose we can show $\phi(n)$ for each n; then since the type of each n is n, for each $n < \omega$ we can show $\forall x^n(x^n \varepsilon \mathbb{N} \rightarrow \phi(x^n))$; now use Limit^{ω}.

[14] See footnote 10.

[15] *Illustration*. Let $\mathrm{Con}_{\mathrm{Zr}}$ be a suitable consistency sentence for Zr. This is independent from Zr, by the second incompleteness theorem; but $\mathrm{Zr}^{(\kappa)}$ proves $\mathrm{Con}_{\mathrm{Zr}}^{(\kappa)}$, since it is arithmetically complete. The same example shows that Zr does not interpret $\mathrm{Zr}^{(\kappa)}$.

[16] *Pace* Degen and Johannsen's [7] sentiment that $\mathrm{CTT}_{\mathrm{p}}$ might serve 'as a foundation for set theory'. Note that differences (1)–(2) also underpin the philosophical discussion of Section 3.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

axiomatizable since, as Gödel [18, p. 45] put it, only recursively axiomatizable theories can leave no doubt regarding whether a putative proof *is* a proof, so that 'the highest possible degree of exactness is obtained'; but only Zr is recursively axiomatized.

2.4. Gödel on 'superfluous restrictions' in type theory. We just quoted Gödel on mathematical foundations. Having made the quoted remarks, Gödel went on to make a more famous claim:

> the system of axioms for the theory of aggregates, as presented by Zermelo, Fraenkel, and von Neumann...is nothing else but a natural generalization of the theory of types, or rather, what becomes of the theory of types if certain superfluous restrictions are removed.[17]

He continued by outlining the 'superfluous restrictions' thus:[18]

(i) '$a \in b$' is meaningful iff the type of 'b' is exactly one greater than that of 'a';
(ii) each class (of any type) can contain classes of exactly one type; and
(iii) only finite types are allowed.

Whilst explicitly disavowing exegetical aims, Linnebo and Rayo [23, pp. 273–274, 278] motivate CTT by suggesting that CTT arises from STT simply by lifting these 'superfluous restrictions'.

Certainly CTT lifts restrictions (i)–(iii). But Zr also lifts these restrictions, and in a different way. Moreover, it is this latter way which we find in Gödel's [18] lecture. On each of points (1)–(3) from Section 2.2, Gödel sides against the use of anything like $\mathrm{Zr}^{(\kappa)}$.

Concerning (1). Gödel [18, p. 47] is clear that the theory which arises by removing STT's 'superfluous restrictions' will supply its own 'types'.[19]

Concerning (2). Gödel [18, p. 49] complains that, in STT, we have to formulate 'the logical axioms for each type separately', and he states that the theory which removes STT's 'superfluous restrictions' will avoid this complaint. Such a theory will therefore employ an *untyped* variable, which can range over all entities.

Concerning (3). As already noted, Gödel [18, p. 45 and 48] insists that an adequate formalization of the foundations of mathematics must be recursively axiomatizable, and explicitly remarks that such theories are necessarily arithmetically incomplete.

Gödel, then, seems never to have envisaged theories like $\mathrm{Zr}^{(\kappa)}$ or $\mathrm{CTT}^{\tau}_{\mathrm{p}}$.[20] Rather, Gödel's suggestion was that removing STT's 'superfluous restrictions' led to ZFU, by the simple stipulation that the 'type' of x is α iff $x \in V_{\alpha+1} \setminus V_\alpha$, with these segments of the set hierarchy defined directly *within* ZFU in the (now) familiar fashion.[21] That is,

[17] Gödel [18, pp. 45–46].

[18] Gödel [18, pp. 46–47]; for discussion, see Feferman [9, p. 37] and Tait [39, p. 88].

[19] Tait [39, p. 92] emphasises this point, and Linnebo and Rayo [23, p. 289 n.28] concede it.

[20] Can we consider (or might Gödel have considered) the move from STT to Zr as involving two steps: *first*, Linnebo and Rayo's step from STT to some CTT^{τ} and $\mathrm{Zr}^{(\kappa)}$; *second* the addition of an untyped variable to $\mathrm{Zr}^{(\kappa)}$, yielding Zr? (Thanks to an anonymous referee for posing this question.) This may be a useful heuristic, but it is slightly technically infelicituous, since the result of adding an untyped variable to $\mathrm{Zr}^{(\kappa)}$ will be arithmetically complete (cf. footnote 29).

[21] Cf. Scott [36, p. 208]: 'the best way to regard Zermelo's theory is as a simplification and extension of Russell's [STT].... The simplification was to make the types *cumulative*.' Note

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Gödel simply identified a set's 'type' with (what we now call) its *rank*, and advocated the use of recursively axiomatized theories whose untyped variables range over all the sets (of all ranks).

§3. The (in)significance of the Sets-from-Types Theorem. We have noted the important mathematical differences between Zr and $\text{Zr}^{(\kappa)}$. We will now show how these differences undermine the philosophical significance of the Sets-from-Types Theorem. In broad brush strokes: Linnebo and Rayo think that the Sets-from-Types Theorem sheds important new light on set theory; we disagree, since $\text{Zr}^{(\kappa)}$ and Zr and importantly distinct.

3.1. Elsa's worries. To reconstruct Linnebo and Rayo's [23, pp. 289–294, 24, p. 178] intended use of the Sets-from-Types Theorem, we will introduce a character, Elsa. Elsa wants to use Zr to talk about the hierarchy of sets, but she has some ontological worries. Following post-Quinean orthodoxy, Elsa draws a sharp distinction between a theory's ontology and its ideology. In general, Elsa thinks that if a theory is coherent, then that is enough to guarantee the good standing of its ideology: roughly, Elsa thinks that a theory's ideology merely provides you with a way of talking about objects, and there is no standard beyond coherence by which to judge ways of talking. Now, Elsa is certain that Zr is coherent, and so she has no reservations about its ideology. But, Zr also postulates a rich ontology of sets, and Elsa insists that the mere coherence of a theory is not enough to guarantee the existence of its ontological commitments. So, Elsa worries: *What guarantees that there are enough sets*?

Linnebo and Rayo have a sequence of recommendations for Elsa. First, they will introduce Elsa to the type hierarchy, in the form of $\text{CTT}^{\tau}_{\text{p}}$, whose coherence can be assumed (at least, in this context). The question arises of how Elsa should think about ontology/ideology in the type-theoretic context. Quantification over type 0 entities is just first-order quantification; so Elsa should think that theorizing at type 0 introduces *ontological* commitments. However, Elsa can perhaps be encouraged to think that theorizing at higher types simply gives us sophisticated ways to talk about the objects at type 0, and so only introduces *ideological* commitments. If Elsa agrees to think in this way, then she will map her dichotomy between ontology and ideology onto the dichotomy between *type* 0 and *type* > 0.[22] Having done this, she will regard $\text{CTT}^{\tau}_{\text{p}}$ as ontologically unproblematic: it posits just one object (i.e., one type 0 entity). Granted, she may regard $\text{CTT}^{\tau}_{\text{p}}$ as ideologically profligate, but she thinks that its coherence guarantees the good standing of its ideology. Consequently, Elsa should

that we are talking about ZFU rather than ZrU. This is inevitable, since ZrU was not formulated until long after Gödel's lecture. However, Gödel supplied an additional argument in favour of Replacement; see footnote 24.

[22] Linnebo and Rayo [23, p. 270] claim that this is how the higher-types are widely regarded by philosophers. For the record, we think that anyone who uses type theory (cumulative or non-cumulative) should reject the idea that there is a useful ontology/ideology dichotomy to be drawn along this faultline. When Quine [31] drew his distinction between ontology and ideology, he drew it for *first-order logic*. In that setting, the distinction is clear enough: we are ontologically committed to the things we quantify over; ideological commitments are expressed by symbols in positions that cannot be quantified into. But in a type-theoretic setting, we can quantify into predicate-position. So distinctions of logical order no longer align with the quantifiable/unquantifiable distinction. See also Trueman [40, chap. 7] and Williamson [44, pp. 260–261].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

have no worries about using $\mathrm{CTT}^{\tau}_{\mathrm{p}}$. Now, via the Set-from-Types Theorem, Elsa can use $\mathrm{CTT}^{\tau}_{\mathrm{p}}$ to obtain $\mathrm{Zr}^{(\kappa)}$. So, according to Linnebo and Rayo, Elsa will have no reason to worry about using $\mathrm{Zr}^{(\kappa)}$ in place of Zr.

Having come this far, Linnebo and Rayo [23, p. 290] hope that Elsa might now be brought to share their view, that 'the two hierarchies'—the 'ideological' hierarchy of $\mathrm{Zr}^{(\kappa)}$ and the 'ontological' hierarchy of Zr—'constitute different perspectives on the same subject-matter.' But we do not need to consider that further step. We think that Elsa should balk at the line of reasoning given in the previous paragraph.

3.2. Ontology relocated. The immediate problem is that Zr and $\mathrm{Zr}^{(\kappa)}$ are importantly different theories. One of the differences, mentioned in Section 2.3, is that Elsa can ask about the height of her set-hierarchy within the object-language of Zr, whereas she can only ask about the height of a type-hierarchy within a metalanguage. But, as we will now show, this basic issue—of object language versus metalanguage—completely undermines the dialectical force of Linnebo and Rayo's line of reasoning.

Recall: Elsa wants to use Zr, but worries: *What guarantees that enough sets exist*? Linnebo and Rayo recommend that Elsa invoke the Sets-from-Types Theorem. Specifically, they encourage Elsa to fix some limit $\kappa > \omega$ with $\kappa + 2 < \tau$, then work in $\mathrm{CTT}^{\tau}_{\mathrm{p}}$ to obtain $\mathrm{Zr}^{(\kappa)}$.

Inevitably, though, this discussion of κ and τ takes place within some *metatheory* which we use to describe $\mathrm{CTT}^{\tau}_{\mathrm{p}}$. After all, as noted in Sections 2.2–2.3, $\mathrm{CTT}^{\tau}_{\mathrm{p}}$'s types are supplied *externally*. So, if Elsa is to follow Linnebo and Rayo's recommendation, she will have to countenance a suitably large index, τ, in the metatheory, so that she can both describe $\mathrm{CTT}^{\tau}_{\mathrm{p}}$ and obtain $\mathrm{Zr}^{(\kappa)}$.

At this point, though, Elsa will simply want to ask: *What guarantees that any suitable τ exists*? Such an entity would have to stand at the head of a vast sequence of type-indices. Well then: *What guarantees that enough type-indices exist*? Her ontological worries about sets have not have been addressed; they have just become worries about the ontology postulated within the metatheory.

3.3. Ideological-bootstrapping. This elementary problem undermines Linnebo and Rayo's way of dealing with Elsa. However, it is worth considering one possible line of response, via (what we call) *ideological-bootstrapping*. This idea is independently interesting, and it will buy Linnebo and Rayo some slack, but not enough slack to save their argumentative strategy.

To define $\mathrm{CTT}^{\tau}_{\mathrm{p}}$, we must be given the type-index τ. In the previous subsection, we imagined Elsa worrying about whether τ exists. But—so this line of reply runs—Elsa is mistakenly assuming here that τ must be a type 0 entity. Instead, τ could be a higher-type entity, supplied by some ideologically-rich but ontologically-innocent theory, $\mathrm{CTT}^{\sigma}_{\mathrm{p}}$. In turn, σ might be some higher-type entity, supplied by some theory $\mathrm{CTT}^{\rho}_{\mathrm{p}}$. And so on.[23]

The hope is that, somehow, considering a sequence of such theories will sooth away Elsa's ontological concerns. But, however exactly this line of response is meant to work, it will require that $\tau > \sigma > \rho > \ldots$. After all, Elsa's worries kick in as soon as the

[23] Linnebo and Rayo [24, p. 179] suggest something a little similar, though in terms of the plurally-interpreted hierarchy (see Section 4.3) and in response to a slightly different concern.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

syntax of $\mathrm{CTT}_{\mathrm{p}}^{\tau}$ is laid down; so her worries clearly cannot be addressed by starting with some theory $\mathrm{CTT}_{\mathrm{p}}^{\sigma}$ with $\sigma \geq \tau$.

This simple observation dictates the form that the attempted reply must take. We are being asked to imagine a sequence of theories, $\mathrm{CTT}_{\mathrm{p}}^{\tau_1}$, ..., $\mathrm{CTT}_{\mathrm{p}}^{\tau_n}$, as follows:

(a) τ_1 is so small that Elsa has no serious qualms about its existence.
(b) As we move along the sequence, the ideology strictly increases (i.e., $\tau_i < \tau_{i+1}$), but the earlier theory proves the existence of an entity which indexes the terms of the next theory (i.e., each $\mathrm{CTT}_{\mathrm{p}}^{\tau_i}$ proves the existence of something with order-type $\tau_{i+1} > \tau_i$).
(c) $\mathrm{CTT}_{\mathrm{p}}^{\tau_n}$ proves $\mathrm{Zr}^{(\kappa)}$, for some suitable κ.

Call this response *ideological-bootstrapping*, since ideologically weaker theories are used to define ideologically richer theories at step (b).[24]

(Note that we have assumed that the sequence of theories is *finite*. To explain why, suppose someone instead suggests this: *If Elsa has accepted the existence of an* ω-*sequence of theories* $\mathrm{CTT}_{\mathrm{p}}^{\tau_1}$, $\mathrm{CTT}_{\mathrm{p}}^{\tau_2}$, ..., *then Elsa can bootstrap her way to their limit*, $\mathrm{CTT}_{\mathrm{p}}^{\tau_\omega}$. This suggestion is spurious. If some $\mathrm{CTT}_{\mathrm{p}}^{\tau_i}$ is sufficient to introduce an entity with order-type τ_ω, then we can simply take $\mathrm{CTT}_{\mathrm{p}}^{\tau_\omega}$ as the $i+1^{\mathrm{th}}$ theory. The important case is when none of the theories $\mathrm{CTT}_{\mathrm{p}}^{\tau_i}$ suffices to introduce anything with order-type τ_ω. But in this case, Elsa will worry whether 'taking the limit' is ontologically innocent; for, by assumption, she has not found any ontologically innocent theory which supplies τ_ω.)

Ideological-bootstrapping might work in specific circumstances. For example, suppose Elsa is comfortable with the existence of $\omega + \omega + 3$, and so has no concerns with the *specification* of $\mathrm{CTT}_{\mathrm{p}}^{\omega+\omega+3}$. Invoking the Sets-from-Types Theorem, $\mathrm{CTT}_{\mathrm{p}}^{\omega+\omega+3}$ proves $\mathrm{Zr}^{(\omega+\omega)}$. This allows Elsa to simulate the set-theoretic hierarchy up to $V_{\omega+\omega}$. Living within $V_{\omega+\omega}$, Elsa can find an uncountable A well-ordered by some relation $<$.[25] Using this, Elsa can define a theory $\mathrm{CTT}_{\mathrm{p}}^{A}$, whose type indices are the members of A as ordered by $<$. Since A is uncountable, $\mathrm{CTT}_{\mathrm{p}}^{A}$ is straightforwardly richer than $\mathrm{CTT}_{\mathrm{p}}^{\omega+\omega+3}$. Moreover, using $\mathrm{CTT}_{\mathrm{p}}^{A}$, Elsa can simulate a much larger chunk of the set-theoretic hierarchy than $V_{\omega+\omega}$; living within that chunk of the hierarchy, she can find larger well-orders; these can be used to supply the indices for some further development of $\mathrm{CTT}_{\mathrm{p}}$ ... and so on. This seems like a case where ideological-bootstrapping might genuinely achieve something.

Nevertheless, there are hard limits on what ideological-bootstrapping can achieve. In the simplest case, suppose Elsa insists on starting with $\mathrm{CTT}_{\mathrm{p}}^{n}$, for some finite n, because she is uncertain whether there are infinitely many entities. Since $\mathrm{CTT}_{\mathrm{p}}^{n}$ only

[24] Gödel [18, p. 47] suggests something similar: given 'the system S_α you can... take an ordinal β greater than α which can be defined in terms of the system S_α, and by means of it state the axioms for the system β including all types less than β, and so on.' However, Gödel is not trying to defend anything like the argument of Section 3.2. As such—and unlike in the context of *ideological*-bootstrapping—Gödel need not confine himself to finite sequences of theories. For discussion of Gödel, see Feferman [8, pp. 37–38 fn. f], Incurvati [19, pp. 90–93]. Koellner [20, pp. 21–24], and Tait [39, pp. 89–93].

[25] Note that A is not a von Neumann ordinal, i.e., A is not well-ordered by $\in$. Still, the existence of some such A and $<$ follows (without Choice) from Hartog's Lemma; see Incurvati [19, p. 92] and Potter [29, p. 185].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

yields (surrogates for) finite well-orders, no amount of ideological-boostrapping from this starting point will allow Elsa to obtain any infinite well-order. So, whenever Linnebo and Rayo try to describe any theory $\mathrm{CTT}^{\tau}_{\mathrm{p}}$ such that τ is infinite, Elsa will worry whether the theory *itself* even exists.

The shape of this problem is quite general. Say that κ is a *hereditary-point* iff κ is an infinite cardinal and everything in V_κ is strictly smaller than κ (so ω is the first hereditary-point).[26] When κ is a hereditary-point, it is in principle impossible to ideologically-bootstrap your way from below κ to above κ, since every entity below level κ is strictly smaller than κ itself.

This problem is especially pertinent, given two facts about hereditary-points and ZF. First, ZF proves that there are proper-class-many hereditary-points; but, since any hereditary-point after ω would be pretty enormous, it is not unreasonable to wonder whether *any* exist; and ideological-bootstrapping cannot quiet such qualms.[27] Second, the standard models of ZF are the V_κ such that κ is strongly inaccessible; and every strongly inaccessible cardinal is a hereditary-point; so ideological-bootstrapping cannot *possibly* address any ontological worries that an Elsa-like character might have about the existence of any standard model of ZF.

The argument of Section 3.2 therefore stands essentially unchanged. Linnebo and Rayo are mistaken to think that cumulative type theories can help us to overcome ontological worries, since the very existence of the (syntactically individuated) theories themselves requires a rich ontology in the metatheory.

§4. CTT: superfluous type-restrictions. In Section 2.4, we discussed Gödel's claim that STT's type-restrictions were 'superfluous'. We should now make explicit something which we there left implicit: these type-restrictions are superfluous *given Gödel's aims*. Specifically, Gödel wanted to establish a foundational, 'formal system which avoids the logical paradoxes and retains all [of] mathematics' [18, p. 46]. Given those aims, CTT's type-restrictions are just as superfluous as STT's; it is best to follow Gödel, and work with something like Zr, with its *untyped* variables.

All of this is compatible with the idea that, given *alternative* aims, STT's or CTT's type-restrictions might not be superfluous, but deeply important. As we will show in this section, though, CTT's type-restrictions are *inevitably* 'superfluous restrictions', in the sense that any semantics for CTT also licenses the use of an *untyped* variable and allows the 'types' to be defined internally. (Cf. points (1) and (2) from Section 2.2.) So, in a slogan: CTT's type-restrictions are superfluous, on any semantics.

We will unpack the details in a moment. First, we should explain the phrase 'a semantics for CTT'. As we are using that phrase, a semantics for CTT is a general framework within which to provide models of CTT, rather than a specific model of

[26] Formally, κ is a hereditary-point iff κ is an infinite cardinal and any of these equivalent conditions hold (we leave the reader to prove the equivalences): (1) $(\forall x \in V_\kappa)|x| < \kappa$; (2) $H_\kappa = V_\kappa$, where $H_\kappa = \{x : |\mathrm{trcl}(x)| < \kappa\}$; (3) either $\kappa = \omega$ or κ is a $\beth$-fixed point, i.e., $\beth_\kappa = \kappa$; (4) $|V_\kappa| = \kappa$. Characterisation (1) formalizes the definition in the text and (2) gives the idea its name.

[27] Boolos [3, p. 258] had qualms about the existence of the first $\aleph$-fixed point; calling it κ, he wrote that κ is 'so big... that it calls into question the truth of any theory, one of whose assertions is the claim that there are at least κ objects'. The first hereditary-point after ω is at least as large as Boolos's κ; it is a $\beth$-fixed point, as in (3) of footnote 26, and hence an $\aleph$-fixed point.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

some CTT$^\tau$. (Compare the idea of 'the possible worlds semantics for modal language'.) So, in providing a semantics for CTT, we fix the meaning of phrases like 'a model of CTT' and 'an entity of type α'; the latter will be the sort of entity which, according to the semantics, can be the value of a type α variable.

4.1. The abstract argument for introducing untyped variables. Our argument begins with an uncontentious point: the stringently-stated rules for CTT tell us that $y^\beta(x^\alpha)$ is well-formed iff $\beta > \alpha$; but these rules are needlessly stringent, on any given semantics.

To see this, fix some semantics for CTT, and let $\beta \leq \alpha$. The formula $x^\alpha \varepsilon y^\beta$ is well-formed according to CTT. So, for any model $\mathcal{M}$ and any type α entity a^α and type β entity b^β from $\mathcal{M}$, either $\mathcal{M} \models a^\alpha \varepsilon b^\beta$ or $\mathcal{M} \models \neg a^\alpha \varepsilon b^\beta$. (Note: what exactly this comes to will depend on the details of the semantics; but we are proceeding abstractly for now and want to consider *any* semantics for CTT.) Now, as explained in Section 1.2, the formula $x^\alpha \varepsilon y^\beta$ perfectly simulates the formula $y^\beta(x^\alpha)$; that is, it perfectly simulates the notion of applying a type β entity to a type α entity. So we could have allowed $y^\beta(x^\alpha)$ to count as well-formed, even though $\beta \not> \alpha$. So, CTT's stringently-stated type-restrictions are needlessly stringent.

To be clear, this is not an *objection* to CTT's type-restrictions. We are really just repackaging a point we made in Section 1.2, and also made by Linnebo and Rayo [23, pp. 282–283], that we can liberalise CTT's stringently-stated formation rules, and allow that $y^\beta(x^\alpha)$ is well-formed for any type-indices α and β. From a purely formal point of view, this changes almost nothing. So, in what follows, we will simply *allow* that CTT counts every formula $y^\beta(x^\alpha)$ as well-formed.

Significantly, though, CTT still lacks *untyped* variables. But, for exactly the same reason, this is also needlessly stringent, on any given semantics.

To see this, fix some semantics for (liberally formulated) CTT. Now $y^\beta(x^\alpha)$ is well-formed for any α and β. So, for any model $\mathcal{M}$ and any type α entity a^α and type β entity b^β from that model, either $\mathcal{M} \models b^\beta(a^\alpha)$ or $\mathcal{M} \models \neg b^\beta(a^\alpha)$. That is, any model assigns a truth value to the application of any entity to any entity, whatever their types might happen to be. So we could have allowed the *untyped* atomic formula, $y(x)$, to count as well-formed: whatever specific values the variables take, the formula would just amount to applying some entity to some entity, which is exactly what the semantics allows.

The upshot is that any semantics for CTT also licenses the use of untyped variables. This time, though, we do have an *objection* to CTT's type-restrictions. Whereas stringently-formulated CTT can simulate any *typed*-formula $y^\beta(x^\alpha)$, via $x^\alpha \varepsilon y^\beta$, it lacks the technical resources to simulate the *untyped*-formula $y(x)$. Untyped variables have to be added by hand. But, once we have added them, we will have moved from a typed to an untyped theory; if we choose to retain 'typed' variables, then they will just behave as restricted untyped variables.

Of course, if there had been no consistent way to introduce untyped variables, then CTT's type-restrictions would have been far from superfluous. But, in this sort of a context, theories like Zr provide us with a clear method for consistently introducing untyped variables.[28] Moreover, they also provide us with a paradigm for how to

[28] We do not need all of Zr; we can make do with the subtheory LT. For details, see Appendix B and Button [5].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

define the notion of 'type' (i.e., *rank*) within the theory. So CTT's type-restrictions are genuinely superfluous.[29]

4.2. Illustration: the class semantics. The argument of the previous subsection is very abstract. To make it more concrete, in this subsection and the next, we will consider two specific semantics in detail: the class semantics, and the plural semantics. Just as our abstract argument predicts, both semantics clearly license the use of untyped variables.

(To avoid any unfortunate misunderstandings: we offer these semantics merely as illustrations. When we say that no semantics could justify the adoption of CTT's type-restrictions, we are not making an inductive inference from these two examples; that conclusion was established by the abstract argument of Section 4.1.)

We start by considering the class semantics. To define a model for CTT within this semantics, we first specify some suitable set of urelements, U. We then stipulate that the type α entities are the members of $U_{\alpha+1}$, where we define:

$$U_1 := U \cup \{\emptyset\} \qquad U_{\alpha+1} := \wp(U_\alpha) \cup U \qquad U_\beta := \bigcup_{\alpha<\beta} U_\alpha \text{ for limit } \beta.$$

Finally, we offer a general clause governing the semantics of atomic sentences:

'$b^\beta(a^\alpha)$' is true iff the referent of 'a^α' is a member of the referent of 'b^β'.

Uncontroversially, CTT is sound for the class semantics. A stringently-typed formula like '$b^2(a^0)$' will be true (in a model) iff the referent of 'a^0' is a member of the referent of 'b^2'. A liberally-typed formula like '$b^0(a^2)$' will also be true (in a model) iff the referent of 'a^2' is a member of the referent of 'b^0'; and this will inevitably be false, since the latter is guaranteed to be an urelement, i.e., an individual without members.

Our semantic clause for atomic sentences employed type restrictions. However, on the class semantics, the type-restrictions are straightforwardly superfluous. We can easily offer a similar semantic clause for untyped terms:

'$b(a)$' is true iff the referent of 'a' is a member of the referent of 'b'.

Otherwise put: there is no barrier to introducing untyped variables, whose values can be any individual or class. Of course, given the old paradoxes, we will have to take care in introducing untyped variables. However, as we have already discussed, Zr-like theories show us how to do this safely.

4.3. Illustration: the plural semantics. The class semantics concerns a class-hierarchy built from a basis of individuals. The plural semantics concerns a plural-hierarchy built from a similar basis.[30] In a little more detail, we use the phrase 'plural*' as a catch-all for whatever we find at any level in the plural hierarchy, i.e., any object,

[29] *Pedantic Objection*. Perhaps CTT$^\tau$'s externally supplied types are not wholly superfluous, since they allow us to formulate the intrinsically infinitary Limit-rules which gives CTT$^\tau$ a kind of strength which a recursive theory like Zr cannot simulate (see (3) from Section 2.2). *Pedantic Reply*. Those who *want* to lean on CTT$^\tau$'s infinitary features can incorporate them within a Zr-like setting. We will illustrate how using Zr itself. For each index $\alpha < \tau$, introduce a new constant, c_α; add to Zr each sentence 'c_α is a von Neumann ordinal'; add the sentence '$c_\alpha \in c_\beta$' iff $\alpha < \beta$; for each limit $\lambda < \tau$, add the infinitary rule: from $\phi(c_\alpha)$ for all $\alpha < \lambda$, infer $(\forall x \in c_\lambda)\phi(x)$.

[30] Rayo [33] develops this plural interpretation.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

any objects, any objectses, ..., any objects(es)$^{\alpha}$....[31] We then offer this general clause governing the semantics for atomic sentences:

> '$b^{\beta}(a^{\alpha})$' is true iff what 'b^{β}' refers to includes what 'a^{α}' refers to.[32]

So '$b^{2}(a^{0})$' is true iff what 'b^{2}' refers to includes what 'a^{0}' refers to; and '$b^{0}(a^{2})$' is true iff what 'b^{0}' refers to includes what 'a^{2}' refers to. But equally, the semantic clause applies perfectly well to untyped terms:

> '$b(a)$' is true iff what 'b' refers to includes what 'a' refers to.

Again: there is no barrier to introducing untyped variables, whose values can be any plural*.

As before, care must be taken to preserve consistency. But we know how to take care: roughly stated, we just need to do for plurals* what Zr does for classes/sets. In more detail, instead of setting up a plural*-hierarchy using type-restricted variables with *externally* supplied type-indices, we can reason about plurals* using an untyped variable, with the plurals* arranged into a cumulative hierarchy according to their rank (with 'rank' defined within the theory, using our untyped variable). And this work has been carried out carefully: Oliver and Smiley [28, chap. 15] and Florio and Linnebo [11, sec. 12.6] both present consistent plural logics featuring untyped variables. Indeed, Florio and Linnebo develop their untyped plural logic precisely by starting with the CTT on the plural semantics, and then collapsing the types in the way that we have described.

§5. STT: type-restrictions justified.

We have argued that CTT's type-restrictions are inevitably superfluous. They are unnecessary for the aim of providing a foundational theory for mathematics, and they cannot be justified semantically, since any semantics for CTT will permit the introduction of an untyped variable.

In this section, we will show that STT's type-restrictions are not similarly superfluous. We can justify the adoption of STT's type-restrictions by invoking the *Fregean semantics*. Indeed, on this semantics, a formula is intelligible iff it is well-formed in STT.

5.1. Against referentialism.

In Sections 4.2 and 4.3, we used the class and plural semantics to illustrate our objection to CTT. Both of these semantics are *referentialist*. By this we mean that both semantics treat every type of term as a type of *referring* term: every type of term performs the same semantic role—*referring*—and all that changes is what they refer to—individuals, classes/plurals*, or something else.[33]

[31] Our word 'plural*' is a 'pseudo-singular device', in the sense of Oliver and Smiley [28, chap. 15]; in natural language, it infelicitously behaves like a singular term. Florio and Linnebo [11, sec. 11.8] use 'higher plurality' here.

[32] The inclusion is *vertical* in the sense of Oliver and Smiley [28, chap. 15]. Vertical inclusion only ever holds between plurals* of different levels, and is analogous to set-membership. Vertical inclusion is to be contrasted with *horizontal* inclusion, which is analogous to subsethood: b horizontally includes a iff b vertically includes everything that a vertically includes.

[33] We are speaking as if variables *refer*. This is one way to gloss a Tarskian referentialist approach to semantics: the value of a variable (on a Tarskian valuation) can be thought of as the variable's referent (on the valuation). In certain contexts, describing variables as *referring* is misleading (see Button and Walsh [6, chap. 1]), but we do not think it will do any

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

The class and plural semantics render CTT's type-restrictions superfluous, precisely because they are referentialist. After all, if every type of term performs the same kind of semantic role as every other type of term, then every type of term can be meaningfully substituted for every other type of term. In that case, as we argued in Section 4.1, the semantics will also allows us to introduce an untyped variable. It follows, immediately, that any semantics which might justify STT's type-restrictions will have to be non-referentialist; in other words, it will have to assign different kinds of semantic role to different types of term.

Now, at one time, this might have seemed like an impossible demand. According to the old Quinean [30, pp. 66–68] orthodoxy, we can only quantify into the position of a referring term; so type theory—which allows us to bind variables of every type—must be given a referentialist semantics. Fortunately, times have changed, and philosophers are increasingly willing to accept quantification into other kinds of position.[34] In what follows, we will simply assume that the old Quinean orthodoxy is mistaken, and will present a particular non-referentialist semantics—the Fregean semantics—which justifies STT's type-restrictions.

5.2. Conceptual but referentialist semantics. The Fregean semantics is a variety of *conceptual* semantics. On a conceptual semantics, type theories are theories of predication:[35] 'a^0' is a name which refers to an object; 'b^1' is a first-level predicate which expresses a property of objects (a *type* 1 *property*);[36] 'c^2' is a second-level predicate which expresses a property of properties of objects (a *type* 2 *property*); and so on.

This way of characterising conceptual semantics is schematic, and we get different versions of the semantics when we supply different accounts of what it means for a predicate to *express* a property. On one view of predication, predicates 'express' properties in the sense that they *refer* to properties, just as names refer to objects. To illustrate, take the following sentence:

(1) Socrates pontificates

According to this view of predication, 'pontificates' refers to the property *Pontification*.[37] Clearly, combining this account of predication with the conceptual semantics yields another brand of referentialism. Every type of term is still referential; all that changes is whether it refers to an ordinary individual (like Socrates) or to something within a property-hierarchy (like *Pontification*). We then have the following semantic clause for atomic sentences:

'$b^\beta(a^\alpha)$' is true iff the referent of 'a^α' instantiates the referent of 'b^β'

harm here. If we wanted, we could say that a semantics is *referentialist* iff it treats every type of *constant* as a referring term, and then use a Robinsonian or hybrid approach to handle variables (again, see Button and Walsh [6, chap. 1]).

[34] See Boolos [3, Prior [31, chap. 3], Rayo and Yablo [34], Trueman [40], Williamson [43, pp. 458–460, 44, chap. 5], and Wright [46].

[35] This point is emphasised throughout Florio and Jones [10].

[36] We have taken the label 'conceptual semantics' from Linnebo and Rayo [23, p. 272], who use 'concept' instead of 'property'. Of course, Linnebo and Rayo are following Frege here. However, this use of 'concept' is potentially misleading; we prefer 'property', which avoids any psychological overtones.

[37] This was arguably the standard way of thinking about predication before Frege introduced his alternative (see below), and plenty of philosophers after Frege have advocated versions of it too: see Bealer [1, chap. 4], Gaskin [16, 17], Strawson [37, 38], and Wiggins [42].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

This allows us to make sense of '$b^{\beta}(a^{\alpha})$', for *any* types α and β. For example, '$b^0(a^0)$' is true iff the referent of 'a^0' instantiates the referent of 'b^0'. Now, admittedly, this formula would correspond to something slightly peculiar in natural language. If 'a^0' referred to Socrates, and 'b^0' referred to Plato, then we might try to render '$b^0(a^0)$' as:

(2) Socrates Plato

This is scarcely grammatical English. Still, for referentialists about predication, (2) is intelligible: it says that Socrates instantiates Plato. Indeed, precisely this point is made by Magidor [27], who insists that (2) is perfectly meaningful and trivially false. We are not agreeing with Magidor here, but we do think that *referentialists* about predication should agree with her. Moreover, and as in Section 4.1, *referentialists* about predication will ultimately find type-restrictions superfluous; nothing will prevent them from introducing untyped variables and insisting that '$b(a)$' is true iff the referent of 'a' instantiates the referent of 'b'.

5.3. Fregean semantics. There is, however, a *non-referentialist* version of the conceptual semantics: it is a *Fregean* semantics.

Unlike referentialists, Fregeans do not think that predicates refer to properties (not, at least, in anything like the sense that a name 'refers').[38] Rather, they think that the role of a predicate is to *say something of* an object; for example, 'pontificates' says of an object that it pontificates. This is the sense in which Fregeans think that predicates are 'incomplete', and they indicate this by writing their predicates with gaps. So rather than writing the predicate in (1) as 'pontificates', they write it as 'x pontificates', where 'x' marks a gap for a name to go. We can then say that sentence (1) is true iff 'x pontificates' says something true of the referent of 'Socrates', i.e., iff Socrates pontificates.[39]

From this Fregean perspective, (2) is not just ungrammatical, but *unintelligible*. We arrive at it by taking (1) and replacing its predicate, 'x pontificates', with a referring name, 'Plato'. Names and predicates are made to work together, but two names cannot work together in the same way. It is not within a name's remit to *say* anything of an object; names just refer to objects. And that is why (2) is a meaningless string: neither name says anything of the referent of the other (let alone something true or false).

Now consider the following sentence:

(3) Someone pontificates

This sentence is not made by combining a predicate with a name. Instead, it is made by combining two predicates, 'x pontificates' and 'Someone Y'. Crucially, though, these are two different types of predicates: 'x pontificates' is a *first-level* predicate, meaning that 'x' marks a gap for a name; 'Someone Y' is a second-level predicate, meaning that 'Y' marks a gap for a first-level predicate. Just as first-level predicates play a

[38] For discussion of the very different sense in which predicates could be said to refer, see Trueman [40, chaps. 4–6].

[39] This account of predication is what we take to be suggested by Frege's (e.g., [13, 14, 15, sec. 31]) discussions of predication; however, we do not want to commit to any exegetical claims here. It is worth noting that the gap between our Fregeans and the referentialists about predication need not be as large as it initially appears. Even if referentialists think of words like 'pontificates' as referring terms, on a par with names like 'Socrates', concatenation behaves like a Fregean predicate: 'xy' says of a pair of objects that the former instantiates the latter. This point is originally due to Frege [14, pp. 192–193], and is further developed by Trueman [40, secs. 3.4 and 8.4].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

different kind of semantic role from the names they can take as input, second-level predicates play a different kind of semantic role from the first-level predicates that they can take as input. We might describe this role thus: a second-level predicate says something of things said of objects. This means that (3) is true/false iff ‘Someone Y’ says something true/false of what ‘x pontificates’ says of objects. Specifically: ‘Someone Y’ says something true of what ‘x pontificates’ says of objects iff ‘x pontificates’ says something true of someone; and it says something false of what ‘x pontificates’ says of objects iff ‘x pontificates’ says something false of everyone.

Again, from this Fregean perspective, it is easy to see that we cannot meaningfully substitute a name for the first-level predicate in (3). Attempting to do this would yield:

(4) Someone Plato

This string is not just ungrammatical, but meaningless. To be meaningful, the input to ‘Someone Y’ must be the kind of expression that says something of objects. But ‘Plato’ *refers to* an object, rather than *saying anything of* objects (let alone something true of someone or false of everyone). So, if we try to plug ‘Plato’ into the argument-place of ‘Someone Y’, we end up with garbage.[40]

The crucial point is that, on the Fregean semantics, different types of term play different types of semantic role: ‘a^0’ is a name which *refers* to an object; ‘b^1’ is a first-level predicate which *says something of* objects; ‘c^2’ is a second-level predicate which *says something of things said of objects*; and so on. And rather than having a single semantic clause which applies to all atomic sentences, we have different clauses for different types of predication:

‘$b^1(a^0)$’ is true iff ‘b^1’ says something true of the referent of ‘a^0’
‘$c^2(b^1)$’ is true iff ‘c^2’ says something true of what ‘b^1’ says of objects
...

These semantic clauses allow us to make sense of ‘$b^n(a^m)$’ iff $n = m + 1$. This is how the Fregean semantics justifies STT’s type-restrictions: a formula is intelligible on the Fregean semantics iff it is well-formed in STT.

For the same reason, the Fregean semantics also prohibits the introduction of untyped variables. Untyped variables would need to be able to take any entity of any type as their values. But, on the Fregean semantics, there is no one sense in which different types of entity could be the ‘value’ of a variable; the sense in which an object is the value of a type 0 variable is incommensurable with the sense in which a type 1 property is the value of a type 1 variable.

To be clear, we are not trying to argue here that anyone should adopt the Fregean semantics.[41] Our point here is just that STT’s type-restrictions, unlike CTT’s, are justified by at least one semantics.[42]

[40] Whitehead and Russell [41, Introduction, chap. II, sec. 4] present a similar argument (in their distinctive terminology).

[41] For an extended argument to that effect, see Trueman [40].

[42] We have considered two conceptual semantics: referentialist and Fregean. Wright [45], MacBride [25], Liebesman [22], and Rieppel [35] offer a third approach, which attempts to provide a *middle-way* between referentialism and Fregeanism. They agree with referentialists that ‘x pontificates’ denotes *Pontification*, but they agree with Fregeans that ‘x pontificates’ says of objects that they teach. Given the latter point, they agree that first-level predicates play a different kind of semantic role from names; so they agree with Fregeans that ‘$c^2(a^0)$’ is

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

5.4. 'Cumulative types' as ambiguous. We have just argued that the Fregean semantics prohibits the introduction of untyped variables. But what it cannot prohibit, of course, is the introduction of *ambiguous* variables, which sometimes behave as one type, and sometimes behave as another. And in fact, this provides the Fregeans with one way of starting to make sense of CTT. Specifically, they can treat a^0 as an ambiguous term: in $b^1(a^0)$, it behave as a *name*, and so refers to an object; in $c^2(a^0)$, it behaves as a *first-level predicate*, and so says something of an object.

If that is how we are to read formulas like $c^2(a^0)$, though, then they no longer represent any departure from STT. Working in STT, we can introduce an injective type-raising function, $\uparrow$, from objects to type 1 properties; so a^0 is an object, but $\uparrow a^0$ is a type 1 property (We also lay down rules to ensure that $\uparrow a^0$ behaves as a suitable surrogate for 'the a^1 such that $a^1 \equiv a^0$'; for details, see §C.) To avoid ambiguity, we can then rewrite $c^2(a^0)$ as $c^2(\uparrow a^0)$, which is now well-formed according to STT's type-restrictions.

This idea can be extended across all finite types. The result is $\text{STT}_\uparrow$, which augments STT with a theory of type-raising functions, like $\uparrow$, whilst retaining STT's type-restrictions. We can then prove the following strong result: CTT^ω and $\text{STT}_\uparrow$ are *definitionally equivalent* (where CTT^ω is the fragment of CTT which uses all and only finite type indices; for details, see Appendix C).

There is, however, an important limitation to this equivalence result. Since entities do not really cumulate in $\text{STT}_\uparrow$, $\text{STT}_\uparrow$ cannot accommodate transfinite types, and so cannot recapture any transfinite uses of CTT. This is significant, because Linnebo and Rayo's main argument for CTT invokes transfinite types (see Section 6). For this reason, Linnebo and Rayo must have intended CTT to be taken at face-value, rather than as a disguised form of $\text{STT}_\uparrow$. Unfortunately for them, though, nothing could justify CTT's type-restrictions, taken at face-value; that was the lesson of Section 4.

§6. The semantic argument.

We have established an important difference between CTT and STT: nothing could justify CTT's type-restrictions, but the Fregean semantics can justify STT's type-restrictions. In this section, we will respond to Linnebo and Rayo's Semantic Argument for CTT. This argument is designed to show that STT is semantically unstable, and that restoring stability pushes us to CTT. We will not present any new objections to CTT in this section; our aim is simply to explain how an advocate of the Fregean semantics should reply to Linnebo and Rayo.

6.1. Naïve Optimism and Naïve Union. Linnebo and Rayo introduce us to two notions:

- A *β-order language* is a language which contains variables of all (and only) the types $\alpha < \beta$.[43]
- A *generalized semantic theory* for a language is 'a theory of all possible interpretations the language might take' [23, p. 275]. In particular, a generalized

unintelligible. However, unlike Fregeans, they cannot embrace STT: according to the middle-way, every type 1 property is also a type 0 object, but STT-Comprehension straightforwardly entails that there are strictly more type 1 properties than objects. Moreover, one of us, Trueman [40, chaps. 4 and 8], has also argued at length that this middle-way is philosophically incoherent.

[43] It can also contain type $\alpha+1$ constants, for any $\alpha < \beta$.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

semantic theory for a β-order language provides an interpretation which allows any type α entity to be the value of a variable x^α, for each $\alpha < \beta$.[44]

These notions are connected by two formal results Linnebo and Rayo [23, Appendix B]:

> *Blocker Theorem*. No language can provide a generalized semantic theory for itself.
>
> *Enabler Theorem*. For any β, let $\beta^* = \beta + 2$ if β is a limit and $\beta^* = \beta + 1$ otherwise; then a β^*-order language can provide a generalized semantic theory for a β-order language.

The Blocker Theorem holds by familiar, liar-like reasoning. Moreover, as Florio and Shapiro [12, pp. 162–163] note, it shows that these two principles are jointly inconsistent:

> *Naïve Optimism*. Any language can be given a generalized semantic theory.
>
> *Naïve Union*. For any languages, there is a union language, which combines all the expressions of those languages.

To see the problem: by Naïve Union, there is a language, $\mathcal{U}$, which is the union of all languages; by Naïve Optimism, $\mathcal{U}$ can be given a generalized semantic theory in some language $\mathcal{G}$; by the Blocker Theorem, $\mathcal{G}$ is not a sub-language of $\mathcal{U}$; but this contradicts the fact that $\mathcal{U}$ is the union of all languages, including $\mathcal{G}$.

6.2. Linnebo and Rayo's Semantic Argument. Linnebo and Rayo avoid contradiction by restricting Naïve Union as follows:

> *Limited Union*. For any limit λ, if there is a β-order language for every $\beta < \lambda$, then there is also a λ-order language.[45]

Having restricted Naïve Union in this way, Linnebo and Rayo's [23, pp. 275–281] *Semantic Argument* for CTT now gets going. Here is a very brief summary. Suppose we start with an ordinary first-order language. By Naïve Optimism, this language has a generalized semantic theory. By the Blocker Theorem, this generalized semantic theory cannot be given in a first-order language; but, by the Enabler Theorem, it can be given in a second-order language. Naïve Optimism now requires that this second-order language has a generalized semantic theory; as before, the Blocker and Enabler

[44] This is very slightly different from what Linnebo and Rayo [23, p. 300] actually say: they consider interpretations of constants (see footnote 43). The particular requirement on generalized semantic theories is an application of the principle that for each α, it is possible to quantify unrestrictedly over all entities of type α. (Linnebo and Rayo [23, p. 274] only state this principle for type 0, but their argument requires that the principle apply to all types; Florio and Linnebo [11, sec. 11.5] explicitly commit themselves to the fully general principle.) We discuss the broader concept of *absolute generality* in Section 7.

[45] Linnebo and Rayo [23, p. 276]. Note that they also [23, p. 294], [24, p. 176] consider a second, slightly differently restricted principle: *For any 'definite totality' of languages, there is a union language*. For our purposes, there is no significant difference between these formulations. Linnebo and Rayo [24, 179–180] treat 'definite totality' as an unanalysed notion. However, the function of this notion is as follows: given any 'definite totality' of languages, we can comprehend a limit-index, λ, which acts as an upper bound of the orders on the languages among that 'definite totality'. (This notion of an 'upper bound' makes sense, since every CTT-like language has well-ordered indices.) So, once we recall that we have only insisted that our type-indices be *well-ordered*, not that they *be ordinals*, the two principles come to the same thing.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

theorems will lead us to provide this semantics in a third-order language. This process repeats, running through every finite order. At this point, Limited Union kicks in, giving us an ω-order language which combines all of the finite orders into a single language. To present a generalized semantic theory for this language, Naïve Optimism and the Blocker and Enabler Theorems push us up to an $\omega+2$-order language. And there is now no stopping us: Naïve Optimism, Limited Union and the two theorems keep pushing us to countenance languages of higher and higher orders. Moreover, when we supply the semantics for variables of some limit type λ, the only plausible option is to allow them to take all entities of all types $< \lambda$ as values. And this requires that at least some of our types be *cumulative*.

6.3. Rebutting the semantic argument. We agree with the following conditional: *if* we accept both Naïve Optimism and Limited Union, *then* there is good reason to embrace CTT. Our response is to reject Naïve Optimism (and to insist on Naïve Union). However, we will show that our stance is more principled that Linnebo and Rayo's.

Linnebo and Rayo [23, p. 280] motivate Limited Union as follows: whenever you are 'prepared to countenance languages of order β for every $\beta < \lambda$', you 'should also countenance languages of order λ', since 'they would be made up entirely of vocabulary that had been previously deemed legitimate'. This line of reasoning is compelling. However, it clearly generalizes, to provide a motivation for Naïve Union. After all: whenever you are prepared to countenance some languages, you should also countenance their union, for that union would be made up entirely of vocabulary that had been previously deemed legitimate. In short: the only motivation Linnebo and Rayo offer for Limited Union is *really* a motivation for Naïve Union.

Of course, Naïve Union is inconsistent with Naïve Optimism. So, if there were a stellar argument in favour of Naïve Optimism, we could see the retreat from Naïve Union to Limited Union as a simple instance of the heuristic that, on encountering a contradiction, we should aim to get as close as we can to what we initially wanted, without falling into inconsistency.[46] Regrettably, though, Linnebo and Rayo provide no argument for Naïve Optimism. So, *prima facie*, an equally good instance of that heuristic would be to accept Naïve Union and instead restrict Naïve Optimism. This threatens to leave us with a deadlock, between those who want to restrict Naïve Union (and so embrace CTT), and those who want to restrict Naïve Optimism (and so might reject CTT).

Fortunately, the argument of Section 5 provides a principled way to break the deadlock: if we are working with a Fregean semantics for the types, then we should restrict Naïve Optimism. Specifically, we should replace Naïve Optimism with the following:

> *Finite Optimism*. Any language of any finite order can be given a generalized semantic theory.

To be clear: the motivation for this restriction is not simply to avoid contradiction. (As far as restoring formal consistency goes, Finite Optimism is serious overkill.) Rather, Finite Optimism expresses the exact amount of optimism which is even *coherent* on the

[46] Cf. Maddy [26, p. 485, 492ff.] on the rules of thumb '*one step back from disaster*' and '*maximize*'; and cf. Linnebo and Rayo [23, p. 274, esp. fn. 8] on the rule of thumb: 'Because we can.'

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Fregean semantics. Since Fregean types cannot cumulate, we cannot make any sense of the idea of an $\omega{+}2$-order language on the Fregean semantics. Finite Optimism and Naïve Union push us to countenance an ω-order language, like STT itself, but we are pushed no further. Otherwise put: STT is the *principled* limit on Fregean types.

§7. Partially cumulative types.

In this paper, we have critically discussed CTT, which is the approach to cumulative types favoured by Linnebo and Rayo. In this final section, we will discuss an alternative approach to cumulative types, due to Florio and Jones [10, sec. 5].

CTT is cumulative in two senses: first, $b^\beta(a^\alpha)$ is well-formed whenever $\beta > \alpha$; second, the values of x^β include all of the values of x^α, whenever $\beta \geq \alpha$. Florio and Jones' cumulative type theory—call it FJT—is cumulative only in the first of these senses. Indeed, for them, no type α entity is a type β entity, when $\alpha \neq \beta$. As we will see, this difference between CTT and FJT is a double-edged sword: on the one hand, it provides Florio and Jones with the means to defend FJT from the argument we offered against CTT in Section 4; on the other hand, it leaves so little distance between FJT and STT, that FJT is best seen as a misleadingly formulated version of STT.

7.1. FJT.

Since entities do not cumulate up the types in FJT, its quantifier rules must be more restrictive than CTT's (see Section 1.2). Indeed, FJT has exactly the same quantifier rules as STT (see Section 1.1). Consequently, in FJT, you cannot generalize about *everything* that has a type 2 property by writing $\forall x^1(a^2(x^1) \rightarrow \phi(x^1))$.[47] In FJT, that formula generalizes over every type 1 property that has a^2, but it says nothing about any objects that have it. To cover everything that might have a^2, we must conjoin that formula with $\forall x^0(a^2(x^0) \rightarrow \phi(x^0))$. Indeed, to generalize over everything that might have a type n property, we will need n conjuncts. This is illustrated by Florio and Jones [11, p. 55] version of Comprehension:

> *FJT-Comprehension.* $\exists z^n \bigwedge_{i<n} \forall x^i(z^n(x^i) \leftrightarrow \phi_i(x^i))$, for each $n > 0$, whenever each $\phi_i(x^i)$ is well-formed and does not contain z^n.

The various ϕ_is need have nothing in common, so this is an instance of FJT-Comprehension:

$$\exists z^2(\forall x^1(z^2(x^1) \leftrightarrow x^1 = x^1) \wedge \forall x^0(z^2(x^0) \leftrightarrow x^0 \neq x^0)).$$

As Florio and Jones [10, p. 61] observe, this entails $\forall x^0 \forall y^1 \; x^0 \not\equiv y^1$, where $\equiv$ is defined as before. More generally, in FJT, if $n \neq m$ then $\forall x^n \forall y^m \; x^n \not\equiv y^m$. So FJT contradicts CTT's Type-Raising principle (see Section 1.2).

7.2. FJT is finitary.

In formulating FJT-Comprehension, we have reverted to using natural numbers as type indices, rather than allowing that types might be transfinite (contrast the formulation of CTT-Comprehension in Section 1.2). We have done this for a simple reason: formulating FJT-Comprehension for a transfinite type, β, would

[47] Throughout this section, we assume a conceptual semantics, and so speak of type $n > 0$ entities as *properties*. Florio and Jones [10, pp. 45–47] offer FJT as a theory of predication, and we also think that STT is best understood as a theory of predication.

require infinitary conjunction:

$$\exists z^{\beta} \bigwedge_{\alpha<\beta} \forall x^{\alpha}(z^{\beta}(x^{\alpha}) \leftrightarrow \phi_{\alpha}(x^{\alpha})).$$

But FJT does not allow for infinitary conjunction. Consequently, FJT cannot comprehend any *transfinite* types.[48]

Much of our discussion of CTT focussed on the Sets-from-Types Theorem (see Section 2–3). However, due to its finitary nature, FJT cannot establish any similar result. Indeed, if we add surrogates for purity and extensionality to FJT, the resulting theory is *decidable*.[49]

7.3. Interpreting FJT's types. Having discussed the Sets-from-Types Theorem, we then argued that CTT's type-restrictions cannot be justified semantically (see Section 4). We began with Linnebo and Rayo [23, pp. 282–283] observation that, even if we stuck with the stringent formation rules for CTT, we could always apply b^{β} to a^{α} in CTT with the formula $a^{\alpha} \mathrel{\varepsilon} b^{\beta}$, which is defined as follows (where $\gamma = \max(\alpha, \beta) + 1$):

$$a^{\alpha} \mathrel{\varepsilon} b^{\beta} \text{ iff}_{\text{df}} (\exists x^{\gamma} \equiv b^{\beta})x^{\gamma}(a^{\alpha}).$$

We then argued that, since every type of entity can be applied to every type of entity in CTT, there can be no barrier to introducing untyped variables.

This line of argument is not straightforwardly applicable to FJT. Since entities do not cumulate up the types in FJT, b^n is not identical to any entity of type $k \neq n$. So, as Florio and Jones [10, sec. 7] stress, it is doubtful whether $a^m \mathrel{\varepsilon} b^n$, i.e., $(\exists x^k \equiv b^n)x^k(a^m)$ with $k = \max(m, n) + 1$, provides us with a way of applying b^n to a^m in FJT.

Nonetheless, we are still left with the question of how to justify the type-restrictions imposed by FJT. Florio and Jones [10, pp. 45–47] explicitly intend to provide FJT with some version of the conceptual semantics, but it is unclear which version they could have in mind. The referentialist version that we discussed in Section 5.2 licenses the use of an untyped variable; the Fregean version that we discussed in Section 5.3 justifies STT's type-restrictions; so it seems that neither of these versions of the conceptual semantics could serve their purpose.

In fact, appearances are somewhat misleading here. It is true that, when FJT is taken at face value, the Fregean semantics cannot justify its type-restrictions. However, it turns out that the Fregean semantics can make good sense of FJT, if its terms are interpreted as being systematically ambiguous, in the following way: in '$c^2(b^1)$', 'c^2' expresses a type 2 property, but in '$c^2(a^0)$', it expresses a type 1 property. (Compare the interpretation of CTT in $\text{STT}_{\uparrow}$ of Section 5.4.)[50]

[48] At least: Florio and Jones nowhere discuss infinitary conjunction, and only ever use natural numbers as type indices.

[49] The surrogate for extensionality is the scheme, for all $n > 0$: $\forall x^n \forall y^n (\bigwedge_{i<n} \forall z^i (x^n(z^i) \leftrightarrow y^n(z^i)) \to x^n = y^n))$; the surrogate for purity is Type-Purity (see Appendix B.2). To see that the resulting theory is decidable, note two facts: (i) all its variables are explicitly typed and (ii) for each n, it proves that there are exactly $h(n)$ type n entities, where $h(0) = 1$ and $h(n+1) = 2^{h(0)+\cdots+h(n)}$; it follows that every quantifier provably has a fixed finite range.

[50] Eagle-eyed readers will notice a slight difference between this and Section 5.4. When dealing with CTT, we read $c^2(a^0)$ as $c^2(\uparrow a^0)$, since CTT licenses Type-Raising, which projects entities upwards through the levels of the type hierarchy. By contrast, FJT contradicts Type-Raising; and FJT-Comprehension effectively projects entities downwards.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

This ambiguity can easily be handled by augmenting STT with a theory of type-lowering relations. We start by introducing a type-lowering relation, $\triangleright$, from type 2 to type 1. We then read '$c^2(b^1)$' verbatim, but treat '$c^2(a^0)$' as shorthand for '$\forall x^1(c^2 \triangleright x^1 \rightarrow x^1(a^0))$'. This latter formula is perfectly well-formed according to STT's type-constraints, and the idea can be extended across all types. The resulting theory is STT$_\triangleright$. We can then prove that FJT and STT$_\triangleright$ are definitionally equivalent. (For details and proof, see Appendix D.)

We think that FJT is best understood as a (somewhat misleading) formulation of STT$_\triangleright$. To begin with, there is no obvious reason to resist this interpretation of FJT. Linnebo and Rayo had a clear technical reason for refusing to interpret CTT via STT$_\uparrow$: the major selling point of CTT was meant to be its ability to accommodate transfinite types (see Section 6). But, as we saw in Section 7.2, FJT is just as limited to finite types as STT. So FJT, like STT, cannot go beyond *Finite Optimism*.

Not only is there no reason for Florio and Jones to resist the interpretation of FJT as STT$_\triangleright$, there is good reason for them adopt it. Their [10] main aim is to argue that cumulative type theories can accommodate *absolute generality*. However, as we will now show, FJT can accommodate absolute generality iff it is taken as a mere notational variant of STT$_\triangleright$.

7.4. STT accommodates absolute generality. We start by explaining how STT accommodates absolute generality.

In traditional set-theoretic semantics, domains are taken to be sets. In STT, we can think of them as properties. For example, we can think of a domain of objects as a type 1 property, d^1, and we can say that x^0 is in that domain iff $d^1(x^0)$. As Williamson [43] clearly explains, there is a real advantage to thinking of domains in this type-theoretic way. There is no set of all objects, and so if we think of domains as sets, unrestricted quantification over all objects is impossible. But STT straightforwardly supplies a type 1 property, U^1, held by all objects, i.e.:[51]

$$\forall x^0 U^1(x^0).$$

(Nothing special is signified by our use of a capitalized 'U' here; it simply aids readability.)

Whilst U^1 includes all the objects, one might worry that it is still *restricted*, since it includes no type 1 properties. But, in the context of STT, this worry is toothless; no sense can be made of this idea in STT. To regard U^1 as *restricted*, we would have to be able to make sense of the idea of a more inclusive domain, which contains both objects and properties.[52] But that is incoherent in STT. To say 'd contains both objects and properties' is to say $\exists x^0 \exists y^1(d(x^0) \wedge d(y^1))$, which is just ungrammatical in STT. For $d(x^0)$ to be grammatical, d must be type 1; for $d(y^1)$ to be grammatical, d must be type 2; but every term has a unique type.

Suppose, then, we introduce suitably typed domains, d^1 and d^2. In STT, these domains are *incommensurable*, to use Williamson's [43, p. 458] phrase. This does not mean that d^1 and d^2 have different members; it means that we cannot even *express* the idea that they have the same (or different) members. We might put this by saying that,

[51] Via $\exists z^1 \forall x^0(z^1(x^0) \leftrightarrow x^0 = x^0)$, which is an instance of STT-Comprehension and CTT- and FJT-Comprehension.

[52] We are not saying that there would have to *be* a more inclusive domain, only that it would have to *make sense* to say that there is. If ϕ make sense, then so must $\neg\phi$.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

in STT, we cannot articulate a univocal notion of *Thing* or *Entity* which applies to both objects and properties. (We can still talk about 'type 0 entities', 'type 1 entities', etc., but we cannot think of 'entity' as a recurring categorematic component in these constructions.) So, if a first-order quantifier quantifies over all objects, then it quantifies over absolutely every *thing* it makes sense to imagine that it might quantify over.

We can put the same point slightly differently by drawing on Florio and Jones' [10, p. 49] explication of *unrestrictedness*: 'an unrestricted domain is a domain such that true universal quantification over it precludes there from being absolutely any counterexamples whatsoever.'[53] This informal explication can be converted into a formal definition in STT. To say that 'everything melted' is true over the domain *things in the freezer* is just to say that $\forall x(x$ is in the freezer $\rightarrow x$ melted$)$. More generally, to say that $\forall x^{n-1} y^n(x^{n-1})$ is *true over* domain d^n is just to say $\forall x^{n-1}(d^n(x^{n-1}) \rightarrow y^n(x^{n-1}))$. To say that there are absolutely no counterexamples to this restricted generalization is to say that the generalization still holds good even when we lift the restriction, and return to $\forall x^{n-1} y^n(x^{n-1})$. And finally, to say that there are absolutely no counterexamples to *any* true quantification over d^n is just to generalize over all y^n. Assembling this, we obtain, for all $n > 0$:

(1) d^n is unrestricted iff$_{\text{df}}$

$$\forall y^n(\forall x^{n-1}(d^n(x^{n-1}) \rightarrow y^n(x^{n-1})) \rightarrow \forall x^{n-1} y^n(x^{n-1})).$$

This definition is adequate because, in STT, only generalizations of the form $\forall x^{n-1} y^n(x^{n-1})$ can be true over d^n. And U^1, as introduced at the start of this subsection, is unrestricted according to (1): since $\forall x^0 U^1(x^0)$, if $\forall x^0(U^1(x^0) \rightarrow y^1(x^0))$, it immediately follows that $\forall x^0 y^1(x^0)$. More generally, within STT, it is obvious that d^n is unrestricted iff $\forall x^{n-1} d^n(x^{n-1})$.

7.5. Absolute generality in FJT. We have seen that STT can accommodate absolute generality. So, if we read FJT as a (misleadingly formulated) notational variant of STT$_{\triangleright}$, then FJT can equally accommodate absolute generality. But, as we will now show, FJT cannot accommodate absolute generality if it is taken at face-value.

To establish this, we will assume in what follows that FJT is to be taken at face-value, so that $d^2(y^1)$ and $d^2(x^0)$ apply the very same type 2 property to y^1 and x^0. (That assumption will remain in force until we explicitly lift it in Section 7.7.) So understood, FJT allows type 2 properties to serve as domains containing both objects and type 1 properties. In fact, FJT delivers a domain, U^2, which contains all type 1 properties and all objects, i.e., such that:[54]

$$\forall x^1 U^2(x^1) \wedge \forall x^0 U^2(x^0)$$

But now first-order quantification becomes a form of *restricted* quantification: in a clear sense, U^1 is a restriction of U^2, since U^2 contains everything in U^1, and more besides.[55]

[53] This explication has an obvious shortcoming: it employs unrestricted quantification itself, in talking about 'absolutely any counterexamples'. However, this shortcoming is shared by every account of unrestricted quantification. Moreover, anyone who already understands unrestricted quantification should agree with Florio and Jones' explication.

[54] Via the FJT-Comprehension instance: $\exists z^2(\forall x^1(z^2(x^1) \leftrightarrow x^1 = x^1) \wedge \forall x^0(z^2(x^0) \leftrightarrow x^0 = x^0))$.

[55] Formally: $\forall x^0(U^1(x^0) \rightarrow U^2(x^0))$, but $\exists y^1(U^2(y^1) \wedge \forall x^0(U^1(x^0) \rightarrow x^0 \not\equiv y^1))$.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

The point here is that FJT *does* treat objects and type 1 properties as a species of a single genus. Indeed, for each $n > 0$, we can think of *Thing*^{n} as the property U^n such that $\bigwedge_{m<n} \forall x^m U^n(x^m)$. So in FJT, it makes sense, *and is true*, to say that first-order quantifiers quantify over some things but not others.[56]

Again, we can make the same point in terms of Florio and Jones' idea that d^n is unrestricted iff there are absolutely no counterexamples to any universal generalization which is true over d^n. Applied to FJT, this does not quite yield a simple *definition* of unrestrictedness,[57] but it does yield a schematic necessary condition for unrestrictedness: if d^n is unrestricted, and y^m is true of everything in d^n that it can be meaningfully applied to, then y^m is true of absolutely everything it can be meaningfully applied to. Formalizing this intuitive idea, we obtain the following, for all $m, n > 0$:

(2) d^n is unrestricted $\rightarrow$

$$\forall y^m (\bigwedge_{i<\min(m,n)} \forall x^i (d^n(x^i) \rightarrow y^m(x^i)) \rightarrow \bigwedge_{i<m} \forall x^i y^m(x^i)).$$

This makes U^1 restricted, since FJT yields an H^2 which applies to every object but to no type 1 property, i.e., such that:[58]

$$\forall x^1 \neg H^2(x^1) \wedge \forall x^0 H^2(x^0).$$

Clearly $\forall x^0(U^1(x^0) \rightarrow H^2(x^0))$, but $\neg\forall x^1 H^2(x^1)$; so U^1 is restricted by (2).[59] A similar argument shows that every domain of every type is restricted in FJT.[60] (And the same style of argument shows that no domain is unrestricted in CTT.)[61]

7.6. Florio and Jones on (R=U). Our argument that every domain is restricted in FJT was based on Florio and Jones' own explication of *unrestrictedness*. But they thought that FJT could accommodate absolute generality. In this subsection, we will lay out their reasoning, and explain why it was mistaken.

Alongside their explication of *unrestrictedness*, Florio and Jones [10, p. 51] introduce a further notion: a domain is *Russellian* for a generalization $\forall v F v$ iff it coincides with the *range of significance* of the predicate F, i.e., the range of things that F can be meaningfully applied to. They then propose [10, pp. 51–53]:

(R=U) A domain is Russellian iff it is unrestricted.

[56] Krämer [21] presents a very similar argument, but directed against CTT rather than FJT.

[57] This is because we can ask whether d^n is m-unrestricted for any $m > 0$; see (4) in Section 7.6.

[58] Via the FJT-Comprehension Instance: $\exists z^2(\forall x^0(z^2(x^0) \leftrightarrow x^0 = x^0) \wedge \forall x^1(z^2(x^1) \leftrightarrow x^1 \neq x^1))$.

[59] This informal argument crucially assumes that FJT is taken at face value. Take the idea that 'y^2 is true of everything in d^1 that it can meaningfully be applied to' can be glossed in FJT as $\forall x^0(d^1(x^0) \rightarrow y^2(x^0))$. Under interpretation into STT$_\triangleright$, this formula becomes $\forall x^0(d^1(x^0) \rightarrow \forall z^1(y^2 \triangleright z^1 \rightarrow z^1(x^0)))$. This no longer says anything about whether y^2 itself is true of everything in d^1.

[60] Assuming that the type hierarchy does not have a terminal level.

[61] When CTT is taken at face value, and again assuming that the type hierarchy does not have a terminal level. In detail: first, we observe that if $d^{\beta+1}$ is unrestricted, then $\forall y^{\gamma+1}(\forall x^\alpha(d^{\beta+1}(x^\alpha) \rightarrow y^{\gamma+1}(x^\alpha)) \rightarrow \forall x^\gamma y^{\gamma+1}(x^\gamma))$, for all $\alpha \leq \min(\beta, \gamma)$. Via CTT-Comprehension, we obtain an $H^{\beta+2}$ such that $\forall x^{\beta+1}(H^{\beta+2}(x^{\beta+1}) \leftrightarrow \exists x^\beta\ x^{\beta+1} \equiv x^\beta)$. Then any $d^{\beta+1}$ is restricted, since $\forall x^\beta(d^{\beta+1}(x^\beta) \rightarrow H^{\beta+2}(x^\beta))$ but $\neg\forall x^{\beta+1} H^{\beta+2}(x^{\beta+1})$.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Here is the idea behind (R=U): a counterexample to $\forall v F v$ would be something of which F is false; but F does not say *anything* (whether true or false) of the things which fall outside of its range of significance; so if $\forall v F v$ is true over d, and d is Russellian for $\forall v F v$, then there cannot be any counterexamples to $\forall v F v$; so d is unrestricted for $\forall v F v$.

Florio and Jones [10, pp. 57–58] attempt to use (R=U) as follows. The domain U^1 is Russellian for the generalization $\forall x^0 a^1(x^0)$: after all, type 1 terms express type 1 properties, and type 1 properties apply meaningfully only to objects.[62] So if we read $\forall x^0 a^1(x^0)$ as a quantification over U^1, then by (R=U) it is *unrestricted*. Whilst U^1 is a strict sub-domain of U^2, none of the extra entities in U^2 fall within a^1's range of significance.

Our basic problem with (R=U) is quite simple: there is a fundamental mismatch between the R and the U. *Unrestrictedness* is normally understood in absolute terms: either a domain is absolutely unrestricted, or it is not. By contrast, Florio and Jones' notion of *Russellianness* is a relative matter: a domain is not just Russellian *full stop*; it is only ever Russellian *for* a generalization $\forall v F v$. This basic problem can be overcome in STT, but not in FJT.

In STT, a property d^n can (meaningfully) be a domain for, and only for, generalizations of the form $\forall x^{n-1} y^n(x^{n-1})$. After all, if we attempt to relativize the generalization $\forall y^m(x^i)$ to d^n, obtaining $\forall x^i(d^n(x^i) \rightarrow y^m(x^i))$, then the result is grammatical in STT iff $n = m = i + 1$. Consequently, the relativity involved in *Russellianness* can be safely ignored: it would not even make sense to say that d^n is Russellian for $\forall x^{m-1} y^m(x^{m-1})$ when $n \neq m$. Indeed, since the range of significance of any type n property in STT is always exactly the type n–1 entities, we can say that d^n is Russellian iff$_{\text{df}}$ $\forall x^{n-1} d^n(x^{n-1})$. Using (1) from Section 7.4, we can then prove (R=U) for STT.

In FJT, by contrast, a property d^n can (meaningfully) be a domain for generalizations $\forall x^i y^m(x^i)$ with $n \neq m$, so we cannot simply ignore the relativity in *Russellianess*. Let us, then, try to accommodate it. Officially, a domain is supposed to be Russellian for a *generalization*. However, since the range of significance of any type m property in FJT is always exactly the type $k < m$ entities, all that really matters is the type of the predicate used in the generalization. This leads to an explicitly relativized notion of *Russellianness* as follows:

(3) d^n is m-Russellian iff$_{\text{df}}$ all and only the type $k < m$ entities have d^n
i.e., $\left(\bigwedge_{k<m} \forall y^k \bigvee_{i<n} (\exists x^i \equiv y^k) d^n(x^i)\right) \wedge \left(\bigwedge_{i<n} \forall x^i (d^n(x^i) \rightarrow \bigvee_{k<m} \exists y^k x^i \equiv y^k)\right)$.

The first conjunct captures the idea that every type $k < m$ entity has d^n; it says that every type $k < m$ entity is an entity in d^n. The second conjunct captures the idea that only the type $k < m$ entities have d^n; it says that every (type $i < n$) entity in d^n is a type $k < m$ entity.[63] This definition allows us (meaningfully) to ask whether d^n is m-Russellian, for any n and m. Furthermore, if $n < m$, then d^n is not m-Russellian. In particular, U^1 is

[62] This explains why Florio and Jones abandoned Linnebo and Rayo's CTT, in favour of a theory which invalidates Type-Raising: in CTT, every type of entity can be applied to every type of entity (using ε if necessary), and so the range of significance of a^1 includes all entities of all types.

[63] Florio and Jones [10, p. 56fn.15] have some doubts about whether $\equiv$ expresses cross-type identity in FJT. If these doubts are justified, then our formal definition of *m-Russellian* will have to be revised as follows:

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

not 2-Russellian. However, U^1 is 1-Russellian. So, in FJT, *Russellianness* is significantly relativized.

To make sense of (R=U) in FJT, then, Florio and Jones must relativize the notion of *unrestrictedness*, so that it matches the relativity in *Russellianness*. Tacitly, they do exactly this, describing domains as unrestricted *for* certain generalizations [10, e.g., pp. 52–53]. Florio and Jones do not define this relative sense of 'unrestricted', but we can easily provide a definition on their behalf. To say that d^n is unrestricted with regard to type m is, presumably, to say this: if y^m is true of everything in d^n that it can be meaningfully applied to, then it is true of absolutely everything it can be meaningfully applied to. Formalizing this, we obtain the following, for all $m, n > 0$:

(4) d^n is m-unrestricted iff$_{\text{df}}$
$$\forall y^m\left(\bigwedge_{i<\min(m,n)} \forall x^i (d^n(x^i) \rightarrow y^m(x^i)) \rightarrow \bigwedge_{i<m} \forall x^i y^m(x^i)\right).$$

Indeed, this just turns (2), which is a schematic necessary condition on unrelativized unrestrictedness, into a definition of relativized m-unrestrictedness.

We can now understand (R=U) thus: a domain is m-Russellian iff it is m-unrestricted. But so understood, (R=U) is false: U^2 is 1-unrestricted but not 1-Russellian, with U^2 as given in Section 7.5. Moreover, we do not need any principle like (R=U) to determine whether a given domain is m-unrestricted; we can just use definition (4). For example, it is clear from (4) that U^1 is 1-unrestricted but 2-restricted. More generally, d^n is m-unrestricted iff both $n \geq m$ and $\bigwedge_{i<m} \forall x^i d^n(x^i)$.

The only remaining question is whether the salient notion of *unrestrictedness* in FJT is the absolute notion governed by (2), or the relative notion defined by (4). We think it is completely clear that the relevant notion is the absolute one. After all, the debate here is about *absolute generality*. It would be false advertising to enter that debate, promising to vindicate unrestricted quantification, and then only deliver *relatively* unrestricted quantification. To emphasise this point, return to the example of U^1: evidently, U^1 is 1-unrestricted but 2-restricted, as defined by (4). Precisely because U^1 is 2-restricted, though, there is a clear sense in which U^1 is restricted *simpliciter*. In particular, with H^2 as given at the end of Section 7.5, everything which is U^1 is H^2, i.e., $\forall x^0(U^1(x^0) \rightarrow H^2(x^0))$, but some entities are not H^2, in that $\exists x^1 \neg H^2(x^1)$.

Indeed, this is exactly where Florio and Jones [10, p. 57] go wrong. They recognise that you can find a type 1 entity not in U^1, but say: 'it does not entail that F is meaningfully predicable of that entity', where $\forall x^0 F(x^0)$ is the generalization under consideration. However, FJT has *precisely* that entailment when F's type is > 1, as in the case of $F = H^2$.

$$(3_*)\ d^n \text{ is } m\text{-Russellian}_* \text{ iff}_{\text{df}} \left(\bigwedge_{k<m} \forall y^k d^n(y^k)\right) \wedge \left(\bigwedge_{m\leq k<n} \forall y^k \neg d^n(y^k)\right).$$

If $n < m$, then 'd^n is m-Russellian$_*$' is ill-formed rather than false. Nevertheless, our key points about *Russellianness* still go through. First, Russellianness$_*$ is significantly relativized, since H^2 is 1-Russellian$_*$ but not 2-Russellian$_*$, with H^2 as given at the end of Section 7.5. Second, (R=U) is false, since U^2 is 1-unrestricted but not 1-Russellian$_*$, with U^2 as given in Section 7.5.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

7.7. *FJT: the case for ambiguity.*

It might be helpful to end our discussion of FJT by summarizing our case for reading it as a mere notational variant of $\mathrm{STT}_{\triangleright}$.

First. We see no reason not to read FJT in this way. Linnedo and Rayo could not read CTT^{ω} as a notational variant of $\mathrm{STT}_{\uparrow}$, because they wanted to extend CTT^{ω} into the transfinite. But FJT is as finitary as $\mathrm{STT}_{\uparrow}$.

Second. If we take FJT at face-value, then it is unclear how we should interpret it. Florio and Jones explicitly intended to give FJT a conceptual semantics, but we know of no version of that semantics which could justify FJT's type-restrictions, taken at face-value.

Third. If we take FJT at face-value, then it cannot accommodate absolute generality. However, if we read FJT as a notational variant of $\mathrm{STT}_{\triangleright}$, then it can supply absolutely unrestricted domains.

§8. Conclusion.

In this paper, we have argued for four main claims:

(a) CTT cannot be used to close the gap between an ideological hierarchy of types and an ontological hierarchy of sets (Section 2 and 3).
(b) CTT's type-restrictions are superfluous, on any semantics (Section 4).
(c) STT's type-restrictions can be justified by a Fregean semantics, which also provides us with a way to resist Linnebo and Rayo's Semantic Argument in favour of CTT (Sections 5 and 6).
(d) FJT is best understood as a misleading formulation of $\mathrm{STT}_{\triangleright}$ (Section 7).

We start with (a). The Sets-from-Types Theorem allows us to simulate Zr within CTT. But deep mathematical differences remain between Zr and and $\mathrm{Zr}^{(\kappa)}$, rendering $\mathrm{Zr}^{(\kappa)}$ unsuitable as a framework for mathematical foundations. Furthermore, the Sets-from-Types Theorem cannot allay any ontological worries we might have about set theory: CTT's type-indices are supplied externally, and so the Sets-from-Types Theorem merely shunts our ontological worries into the metalanguage.

Next is (b). CTT is a remarkably relaxed type theory: it allows us to apply every type of entity to every type of entity. But it still retains the constraint that all of its variables are typed and, in CTT, that type-restriction is superfluous. Once every type of entity can be applied to every type of entity, there can be no barrier to introducing untyped variables.

We come now to (c). The strict type-restrictions imposed by STT can be justified by the Fregean semantics. On this semantics, different types of term play fundamentally different types of semantic role, so that they cannot be meaningfully intersubstituted. Moreover, this semantics yields a principled reason to reject *Naïve Optimism*, a crucial premise in Linnebo and Rayo's Semantic Argument.

We end with (d). Florio and Jones' FJT was meant to be a *partially* cumulative type theory, but we argue it is best understood as a notational variant of $\mathrm{STT}_{\triangleright}$: taking FJT at face-value leaves it unable to accommodate absolute generality; whereas $\mathrm{STT}_{\triangleright}$—which is definitionally equivalent to FJT—provides absolutely unrestricted domains of quantification.

§A. Elementary facts about CTT.

The remainder of this paper comprises technical appendices, covering the formal results mentioned in the main text. We will start with some elementary observations about CTT. As mentioned in Section 1.2, for each ordinal τ, we have a theory CTT^{τ}.[64] Recall that we have explicitly defined $\equiv$ and ε, for any types α and β and where $\gamma = \max(\alpha, \beta) + 1$:

$$a^{\alpha} \equiv b^{\beta} \text{ iff}_{\text{df}}\ \forall x^{\gamma}(x^{\gamma}(a^{\alpha}) \leftrightarrow x^{\gamma}(b^{\beta}))$$
$$a^{\alpha}\ \varepsilon\ b^{\beta} \text{ iff}_{\text{df}}\ (\exists x^{\gamma} \equiv b^{\beta})x^{\gamma}(a^{\alpha}).$$

In what follows, we will frequently invoke the following simple facts about $\equiv$ and ε; crudely, they allow us to move seamlessly between different type-levels:

Lemma 1. *If $\alpha \leq \beta$ and $\beta + 1 < \tau$, then CTT^{τ} proves: $\forall a^{\alpha} \exists b^{\beta}\ \ a^{\alpha} \equiv b^{\beta}$*

Proof. By $\forall \mathrm{I}^{\beta+1}_{\beta+1}$, we have $\forall x^{\beta+1}(x^{\beta+1}(a^{\alpha}) \leftrightarrow x^{\beta+1}(a^{\alpha}))$, i.e., $a^{\alpha} \equiv a^{\alpha}$; so $\forall a^{\alpha} \exists b^{\beta}\ a^{\alpha} \equiv b^{\beta}$ by $\exists \mathrm{I}^{\beta}_{\alpha}$ followed by $\forall \mathrm{I}^{\alpha}_{\alpha}$. □

Lemma 2. *For any ϕ, and any α, β, γ with $\max(\alpha, \beta, \gamma) + 2 < \tau$, CTT^{τ} proves:*

(1) *if $a^{\alpha} \equiv b^{\beta}$ and $\phi(a^{\alpha})$, then $\phi(b^{\beta})$, when this is well-formed*
(2) *if $a^{\alpha} \equiv b^{\beta} \equiv c^{\gamma}$, then $a^{\alpha} \equiv c^{\gamma}$*
(3) *if $a^{\alpha} \equiv b^{\beta}$ and $a^{\alpha}\ \varepsilon\ c^{\gamma}$, then $b^{\beta}\ \varepsilon\ c^{\gamma}$*
(4) *if $a^{\alpha} \equiv b^{\beta}$ and $c^{\gamma}\ \varepsilon\ a^{\alpha}$, then $c^{\gamma}\ \varepsilon\ b^{\beta}$*

Proof. (1) Suppose $a^{\alpha} \equiv b^{\beta}$ and $\phi(a^{\alpha})$. Let $\delta = \max(\alpha, \beta)$; by CTT-Comprehension there is some $c^{\delta+1}$ such that $\forall x^{\delta}(c^{\delta+1}(x^{\delta}) \leftrightarrow \phi(x^{\delta}))$. Since $\phi(a^{\alpha})$, by $\forall \mathrm{E}^{\delta}_{\alpha}$ we have that $c^{\delta+1}(a^{\alpha})$. Since $a^{\alpha} \equiv b^{\beta}$, i.e., $\forall z^{\delta+1}(z^{\delta+1}(a^{\alpha}) \leftrightarrow z^{\delta+1}(b^{\beta}))$, by $\forall \mathrm{E}^{\delta+1}_{\delta+1}$ we have that $c^{\delta+1}(b^{\beta})$. Now $\phi(b^{\beta})$ by $\forall \mathrm{E}^{\delta}_{\beta}$.

(2)–(4) We leave these to the reader. They are not completely immediate consequences of (1), since the definitions of $\equiv$ and ε are typically ambiguous. □

Lemma 3. *If $\max(\alpha, \beta) + 2 < \tau$, then CTT^{τ} proves: $a^{\alpha}\ \varepsilon\ b^{\beta+1} \leftrightarrow (\exists x^{\beta} \equiv a^{\alpha})b^{\beta+1}(x^{\beta})$*

Proof. By Type-Founded and Lemmas 1 and 2. □

It is worth emphasising that Type-Founded and Type-Basis are independent from CTT's other axioms. To show this, we begin by building an ill-founded set-theoretic structure, $\mathcal{A}$. Let $\mathbf{a} = \{\emptyset, \mathbf{a}\}$; now define:

$$A_1 := \mathbf{a} \qquad\qquad A_{n+1} := \wp(A_n) \qquad\qquad A := \bigcup_{n<\omega} A_n.$$

So $A_2 = \{\emptyset, \{\emptyset\}, \{\mathbf{a}\}, \mathbf{a}\}$. Let $\mathcal{A}$ be the structure whose domain is A and which interprets $\in$ verbatim; evidently, $\mathcal{A}$ is ill-founded. Using a slight tweak of the class semantics of Section 4.2, we now create a model, $\mathcal{M}$, of CTT^{ω} without Type-Founded. We start by defining a ranking function $\rho : A \longrightarrow \mathbb{N}$ on $\mathcal{A}$ as follows:

$$\rho(\emptyset) = 0 \qquad\qquad \rho(\mathbf{a}) = 1 \qquad\qquad \rho(c) = n \text{ iff } c \in A_n \setminus A_{n-1.}$$

So $\rho(\{\emptyset\}) = 2$. Now we stipulate that $\mathcal{M}$'s type n entities are all those $c \in A$ such that $\rho(c) \leq n$, and applications are stipulated to hold as follows, for all $m < n$ and all $b, c \in A$:

$$\mathcal{M} \models c^{n}(b^{m}) \text{ iff } b \in c.$$

[64] Throughout the appendices, we will assume that all type-indices are ordinals; nothing turns on this, but it makes the technicalities more familiar.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

It is easy to confirm that $\mathcal{M}$ models CTT$^{\omega}$ without Type-Founded. But, by construction, $c^m \equiv c^n$ whenever $\min(m, n) \geq \rho(c)$. So $\mathbf{a}^2(\mathbf{a}^1)$ with $\mathbf{a}^2 \equiv \mathbf{a}^1$, and hence $\mathbf{a}^1 \mathrel{\varepsilon} \mathbf{a}^1$. So $\mathcal{M}$ violates Type-Founded. Admittedly, Type-Basis holds in $\mathcal{M}$, but we can violate it with a similar construction: start with a Quine atom $\mathbf{b} = \{\mathbf{b}\}$; let $B_0 = \mathbf{b}$ and $B_{n+1} = \wp(B_n)$; define $\rho(c) = n$ iff $c \in B_n \setminus B_{n-1}$; and note that $\mathbf{b}^0 \mathrel{\varepsilon} \mathbf{b}^0$.

§B. Obtaining Zr in CTT$_\mathrm{p}$. In Section 2, we stated the Sets-from-Types Theorem. In this appendix, we prove that result. We also introduce the interpreting theory, CTT$_\mathrm{p}$, and the interpreted theory, Zr, and discuss how CTT$_\mathrm{p}$ deals with Replacement.

B.1. The theory Zr. The set theory which we simulate is Zr. We can think of Zr as arising by adding to Z the principle that the sets are arranged in well-ordered levels; Zr is therefore strictly stronger than Z and strictly weaker than ZF.[65] We follow Button's [5] formulation of Zr, starting with a core of definitions:

Definition 4. *Say that h is a history, written $Hist(h)$, iff $(\forall a \in h)\forall x(x \in a \leftrightarrow (\exists c \in h)x \subseteq c \in a)$. Say that s is a level, written $Lev(s)$, iff $\exists h(Hist(h) \land \forall x(x \in s \leftrightarrow \exists c(x \subseteq c \in h)))$.*[66]

Using these definitions, we can consider some axioms:

Extensionality $\forall a \forall b(\forall x(x \in a \leftrightarrow x \in b) \rightarrow a = b)$
Separation $\forall a \exists b \forall x(x \in b \leftrightarrow (\phi(x) \land x \in a))$, for every ϕ not containing b
Stratification $\forall a(\exists s \supseteq a)Lev(s)$
Endless $\forall a \exists b\ a \in b$
Infinity $\exists a(\exists x\ x \in a \land (\forall x \in a)\exists y(x \in y \in a))$

The theory LT has, as axioms, Extensionality, all instances of Separation, and Stratification, which serves as a principle of foundation. The theory Zr adds Endless and Infinity to LT. In what follows, these next two results will be extremely useful:[67]

Lemma 5. *Extensionality + Separation proves: if $Lev(s)$, then $s = \{x : \exists r(Lev(r) \land x \subseteq r \in s)\}$.*

Theorem 6. *Extensionality + Separation proves: the levels are well-ordered by $\in$, i.e.:*

(1) $\exists s(Lev(s) \land \phi(s)) \rightarrow \exists s(Lev(s) \land \phi(s) \land \neg(\exists r \in s)(Lev(r) \land \phi(r)))$
(2) $\forall s \forall t((Lev(s) \land Lev(t)) \rightarrow (s \in t \lor s = t \lor t \in s))$

This last result allows us to define the rank of a, written $\mathrm{rank}(a)$, within LT, in terms of the $\in$-least level with a as a subset.

B.2. The theory CTT$_\mathrm{p}$. The theory CTT$_\mathrm{p}$ extends CTT with two new principles.[68] First, we add a version of 'extensionality', for all $\alpha \leq \beta$:

$$\textit{Type-Ext}.\ \forall a^{\alpha+1} \forall b^{\beta+1}([\forall x^{\alpha}(a^{\alpha+1}(x^{\alpha}) \rightarrow b^{\beta+1}(x^{\alpha})) \land \\ \forall x^{\beta}(b^{\beta+1}(x^{\beta}) \rightarrow (\exists y^{\alpha} \equiv x^{\beta})a^{\alpha+1}(y^{\alpha}))] \rightarrow a^{\alpha+1} \equiv b^{\beta+1}).$$

[65] Zr is equivalent to Potter's [29] theory **Z**; this is strictly stronger than Zermelo's Z.
[66] *Notation*: we let '$x \subseteq c \in h$' abbreviate '$(x \subseteq c \land c \in h)$'; similarly for other infix predicates.
[67] See Button [5, sec. 3] for proofs.
[68] Compare these with Degen and Johannsen [7, 149 *Ext*, 153 *Nullity*].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Second, to achieve 'purity', we add an axiom stating that there is exactly one object:

Type-Purity. $\forall x^0 \forall y^0\ x^0 = y^0$

Note that, modulo CTT's other axioms, Type-Founded follows from Type-Ext and Type-Purity.

To begin our simulation of Zr within CTT_p, we will show that CTT_p^τ proves Extensionality$^{(\kappa)}$ and Separation$^{(\kappa)}$.

LEMMA 7. *$CTT_p^\tau \vdash$ Extensionality$^{(\kappa)}$, whenever $\kappa + 2 < \tau$.*

Proof. Suppose κ is a limit (the proof is easier when κ is a successor). Without loss of generality, fix $\alpha \leq \beta < \kappa$ and suppose $\forall x^\kappa(x^\kappa \mathrel{\varepsilon} a^\alpha \leftrightarrow x^\kappa \mathrel{\varepsilon} b^\beta)$.

Using Lemma 1, find $a^{\alpha+1} \equiv a^\alpha$ and $b^{\beta+1} \equiv b^\beta$. Suppose $a^{\alpha+1}(x^\alpha)$. So $x^\alpha \mathrel{\varepsilon} a^{\alpha+1}$ by Lemma 3; by Lemma 1 there is $x^\kappa \equiv x^\alpha$, and $x^\kappa \mathrel{\varepsilon} a^\alpha$ by Lemma 2; so $x^\kappa \mathrel{\varepsilon} b^\beta$, and now $b^{\beta+1}(x^\alpha)$ by Lemmas 2 and 3. Similar reasoning shows: if $b^{\beta+1}(x^\beta)$ then $(\exists y^\alpha \equiv x^\beta)a^{\alpha+1}(y^\alpha)$. By Type-Ext, $a^{\alpha+1} \equiv b^{\beta+1}$; hence $a^\alpha \equiv b^\beta$ by Lemma 2. Generalizing, for any $\alpha, \beta < \kappa$:

$$\forall x^\kappa(x^\kappa \mathrel{\varepsilon} a^\alpha \leftrightarrow x^\kappa \mathrel{\varepsilon} b^\beta) \rightarrow a^\alpha \equiv b^\beta.$$

Now Extensionality$^{(\kappa)}$ holds, using Limit$^\kappa$ twice. □

LEMMA 8. *$CTT_p^\tau \vdash$ Separation$^{(\kappa)}$, whenever $\kappa + 2 < \tau$.*

Proof. Suppose κ is a limit (the proof is easier when κ is a successor). Fix $\alpha < \kappa$ and ϕ such that $\phi(x^\kappa)$ is well-formed. Fix a^α and find $a^{\alpha+1} \equiv a^\alpha$ by Lemma 1. Using CTT-Comprehension, fix $b^{\alpha+1}$ such that:

$$\forall x^\alpha(b^{\alpha+1}(x^\alpha) \leftrightarrow (\forall x^\kappa \equiv x^\alpha)(\phi(x^\kappa) \land x^\kappa \mathrel{\varepsilon} a^{\alpha+1})).$$

Suppose $z^\kappa \mathrel{\varepsilon} b^{\alpha+1}$; by Lemma 3 there is $z^\alpha \equiv z^\kappa$ such that $b^{\alpha+1}(z^\alpha)$; so using the biconditional, $\phi(z^\kappa) \land z^\kappa \mathrel{\varepsilon} a^{\alpha+1}$. Conversely, suppose $\phi(z^\kappa) \land z^\kappa \mathrel{\varepsilon} a^{\alpha+1}$; by Type-Founded there is $z^\alpha \equiv z^\kappa$, and $(\forall x^\kappa \equiv z^\alpha)(\phi(x^\kappa) \land x^\kappa \mathrel{\varepsilon} a^{\alpha+1})$ by Lemma 2; so $b^{\alpha+1}(z^\alpha)$, and $z^\kappa \mathrel{\varepsilon} b^{\alpha+1}$ by Lemma 3. Summarizing: $z^\kappa \mathrel{\varepsilon} b^{\alpha+1} \leftrightarrow (\phi(z^\kappa) \land z^\kappa \mathrel{\varepsilon} a^{\alpha+1})$. By Lemma 1 there is $b^\kappa \equiv b^{\alpha+1}$. Generalizing and using Lemma 2, for any $\alpha < \kappa$:

$$\forall a^\alpha \exists b^\kappa \forall z^\kappa(z^\kappa \mathrel{\varepsilon} b^\kappa \leftrightarrow (\phi(z^\kappa) \land z^\kappa \mathrel{\varepsilon} a^\alpha)).$$

Now Separation$^{(\kappa)}$ follows by the Limit$^\kappa$-rule. □

Consequently, CTT_p^τ proves Lemma 5$^{(\kappa)}$ and Theorem 6$^{(\kappa)}$. The latter result states that the *Lev*$^{(\kappa)}$s are well-ordered by ε. Here, '*Lev*$^{(\kappa)}$' is the obvious translation of the definition of '*Lev*'; we also call these levels$^{(\kappa)}$. In what follows, we also write things like $x^\kappa \subseteq^{(\kappa)} y^\kappa$ for $(\forall v^\kappa \mathrel{\varepsilon} x^\kappa)v^\kappa \mathrel{\varepsilon} y^\kappa$.

Our next goal is to show that CTT_p^τ simulates our set-theoretic principle of foundation, i.e., Stratification. We first need a small subsidiary lemma, which says (roughly) that any subset of a low-typed entity is itself low-typed:

LEMMA 9. *$CTT_p^\tau \vdash \forall a^\alpha(\forall b^\kappa \subseteq^{(\kappa)} a^\alpha)\exists x^\alpha\, b^\kappa \equiv x^\alpha$, whenever $\kappa + 2 < \tau$.*

Proof. Suppose α and κ are limits (the proof is easier otherwise). Let $\beta < \kappa$, and fix $b^\beta \subseteq^{(\kappa)} a^\alpha$; it suffices to show that $\exists x^\alpha\, b^\beta \equiv x^\alpha$, since we can then use Limit$^\kappa$ to establish the result.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

If $\beta \leq \alpha$, Lemma 1 immediately tells us that $\exists x^\alpha\ b^\beta \equiv x^\alpha$. Suppose instead that $\beta > \alpha$. Fix $\gamma < \alpha$, and suppose there is some $a^{\gamma+1} \equiv a^\alpha$. Using Lemma 1, let $b^{\beta+1} \equiv b^\beta$. By CTT-Comprehension, there is $c^{\gamma+1}$ such that:

$$\forall v^\gamma (c^{\gamma+1}(v^\gamma) \leftrightarrow b^{\beta+1}(v^\gamma)).$$

Using Lemmas 1–3: if $b^{\beta+1}(v^\beta)$, then $v^\beta \mathrel{\varepsilon} a^{\gamma+1}$ since $b^\beta \subseteq^{(\kappa)} a^{\gamma+1}$, so that there is $y^\gamma \equiv v^\beta$; now $b^{\beta+1}(y^\gamma)$, so that $c^{\gamma+1}(y^\gamma)$. Generalizing, $\forall v^\beta(b^{\beta+1}(v^\beta) \to (\exists y^\gamma \equiv v^\beta)c^{\gamma+1}(y^\gamma))$. By Type-Ext, $c^{\gamma+1} \equiv b^{\beta+1} \equiv b^\beta$. Summarizing all this, we have established the following conditional, for each $\gamma < \alpha$:

$$\exists x^{\gamma+1}\ a^\alpha \equiv x^{\gamma+1} \to \exists x^{\gamma+1}\ b^\beta \equiv x^{\gamma+1}.$$

Now, for reductio, suppose that $\forall x^\alpha\ b^\beta \not\equiv x^\alpha$. Then $\forall x^{\gamma+1}\ b^\beta \not\equiv x^{\gamma+1}$ for all $\gamma < \alpha$. So, by the relevant conditional, $\forall x^{\gamma+1}\ a^\alpha \not\equiv x^{\gamma+1}$. By the Limit$^\alpha$-rule, $\forall x^\alpha\ a^\alpha \not\equiv x^\alpha$, a contradiction. Discharging the reductio, $\exists x^\alpha\ b^\beta \equiv x^\alpha$, as required. □

Lemma 10. *$CTT^\tau_p \vdash Stratification^{(\kappa)}$, whenever $\kappa + 2 < \tau$.*

Proof. We will show that, for each $\beta \leq \kappa$, CTT$^\tau_{\mathrm{p}}$ proves $\forall a^\beta(\exists s^\beta \supseteq^{(\kappa)} a^\beta)Lev^{(\kappa)}(s^\beta)$. This is an induction on β in the metatheory, where our induction hypothesis is that for each $\alpha < \beta$ we have established $\forall a^\alpha(\exists s^\alpha \supseteq^{(\kappa)} a^\alpha)Lev^{(\kappa)}(s^\alpha)$

Induction case when $\beta = 0$. By Type-Founded, $\forall x^\kappa\ x^\kappa \not\mathrel{\varepsilon} a^0$. So $Hist^\kappa(a^0)$ and $Lev^{(\kappa)}(a^0)$, vacuously. So $\forall a^0(\exists s^0 \supseteq^{(\kappa)} a^0)Lev^{(\kappa)}(s^0)$.

Induction case when β is a limit. Applying Lemmas 1–2 to our induction hypothesis, we have $\forall a^\alpha(\exists s^\beta \supseteq^{(\kappa)} a^\alpha)Lev^{(\kappa)}(s^\beta)$. Now $\forall a^\beta(\exists s^\beta \supseteq^{(\kappa)} a^\beta)Lev^{(\kappa)}(s^\beta)$, by the Limit$^\beta$-rule.

Induction case when $\beta = \alpha + 1$. Using CTT-Comprehension twice, find h^β and s^β such that

$$\forall x^\alpha(h^\beta(x^\alpha) \leftrightarrow Lev^\kappa(x^\alpha)) \tag{1}$$

$$\forall x^\alpha\ \ s^\beta(x^\alpha). \tag{2}$$

Combining these with the induction hypothesis, we obtain:

$$\forall x^\alpha(s^\beta(x^\alpha) \leftrightarrow (\exists c^\alpha \supseteq^{(\kappa)} x^\alpha)h^\beta(c^\alpha)).$$

Hence, by Lemmas 1–3 and 9:

$$\forall x^\kappa(x^\kappa \mathrel{\varepsilon} s^\beta \leftrightarrow \exists c^\kappa(x^\kappa \subseteq^{(\kappa)} c^\kappa \mathrel{\varepsilon} h^\beta)). \tag{3}$$

Next, applying $\forall\mathrm{E}^\kappa_\alpha$ and $\forall\mathrm{I}^\alpha_\alpha$ to Lemma $5^{(\kappa)}$ gives:

$$\forall a^\alpha(Lev^{(\kappa)}(a^\alpha) \to \forall x^\kappa(x^\kappa \mathrel{\varepsilon} a^\alpha \leftrightarrow \exists c^\kappa(Lev^{(\kappa)}(c^\kappa) \wedge x^\kappa \subseteq^{(\kappa)} c^\kappa \mathrel{\varepsilon} a^\alpha))).$$

So, by (1) and Lemmas 1–3:

$$(\forall a^\kappa \mathrel{\varepsilon} h^\beta)\forall x^\kappa(x^\kappa \mathrel{\varepsilon} a^\kappa \leftrightarrow (\exists c^\kappa \mathrel{\varepsilon} h^\beta)x^\kappa \subseteq^{(\kappa)} c^\kappa \mathrel{\varepsilon} a^\kappa),$$

i.e., h^β is a history$^{(\kappa)}$. So s^β is a level$^{(\kappa)}$, by (3). Moreover, for any a^β, we have $a^\beta \subseteq^{(\kappa)} s^\beta$ by (2) and Lemmas 1–3. So $\forall a^\beta(\exists s^\beta \supseteq^{(\kappa)} a^\alpha)Lev^{(\kappa)}(s^\beta)$. □

We have now established $\mathrm{LT}^{(\kappa)}$. To obtain $\mathrm{Zr}^{(\kappa)}$, we need just two straightforward results, which we leave to the reader (they hold using CTT-Comprehension, Lemmas 1–3, and the Limit-rule).

Lemma 11. *$CTT_p^\tau \vdash Endless^{(\kappa)}$, whenever $\kappa + 2 < \tau$ and κ is a limit.*

Lemma 12. *$CTT_p^\tau \vdash Infinity^{(\kappa)}$, whenever $\kappa > \omega$ and $\kappa + 2 < \tau$.*

Assembling Lemmas 7–12, we have the Sets-from-Types Theorem:

Theorem 13. *$CTT_p^\tau \vdash Zr^{(\kappa)}$ for any limit $\kappa > \omega$ with $\kappa + 2 < \tau$.*

B.3. Replacement, and semantic considerations. We mentioned that Zr sits strictly between Z and ZF. Specifically, Zr does not include Replacement. To settle the status of Replacement with regard to $\mathrm{CTT_p}$,[69] we will move from proof theory to semantics, linking models of $\mathrm{CTT_p}$ with models of LT. (Recall that LT is the subtheory of Zr whose axioms are Extensionality, Separation and Stratification.)

In considering models of LT, we restrict our attention to transitive models.[70] Recall that a structure $\mathcal{A}$ in the signature of set theory is *transitive* iff both $(\forall x \in A) x \subseteq A$, and $(\forall a \in A)(\forall b \in A)(a \in b \leftrightarrow a \in^{\mathcal{A}} b)$. So membership and subsethood are absolute for transitive models. Also recall that being a (von Neumann) ordinal is absolute for transitive models,[71] and so is the notion of a set's (ordinal) rank. (Recall from Appendix B.1 that we can define a set's rank within LT, and hence within Zr.) Where $\mathcal{A}$ is a transitive model of LT, let $Ord(\mathcal{A})$ be the least ordinal not in A itself.

Whilst we consider only transitive models of LT, we will entertain non-standard models. A transitive model $\mathcal{A} \vDash \mathrm{LT}$ is *standard* iff for any $\alpha < Ord(\mathcal{A})$, every subset of $\{x \in A : \mathrm{rank}(x) \leq \alpha\}$ is itself in $\mathcal{A}$.

Given any model $\mathcal{M} \vDash \mathrm{CTT}_\mathrm{p}^\tau$ with $\kappa + 2 < \tau$, we can easily turn it into a transitive set-theoretic model, $\mathbf{L}^\kappa\mathcal{M}$, as follows: let $\mathbf{L}^\kappa\mathcal{M}$'s domain comprise all the type κ entities from $\mathcal{M}$; and let $\mathbf{L}^\kappa\mathcal{M} \vDash a \in b$ iff $\mathcal{M} \vDash a \mathrel{\varepsilon} b$.

Lemma 14. *When $\kappa + 2 < \tau$: if $\mathcal{M} \vDash CTT_p^\tau$, then $\mathbf{L}^\kappa\mathcal{M}$ is isomorphic to a unique transitive model of LT.*

Proof. By Lemmas 7–10, $\mathbf{L}^\kappa\mathcal{M} \vDash \mathrm{LT}$. The type indices are well-ordered. By the Limit-rule, Type-Founded, Type-Purity and Lemma 7, $\mathbf{L}^\kappa\mathcal{M}$'s membership relation is extensional and well-founded. Now use Mostowski's Collapsing Lemma. □

We can also move in the opposite direction, from transitive models of LT to models of $\mathrm{CTT_p}$. In effect, we follow the class-semantics of Section 4.2, but tweaked to ban urelements and to allow for non-standard models of $\mathrm{CTT_p}$, where a model of $\mathrm{CTT_p}$ is *standard* iff for any entities of any type α (other than the greatest) in the model, some type $\alpha+1$ property in the model applies exactly to those entities.[72] Still, the basic plan

[69] This addresses Linnebo and Rayo [23, p. 289 n.28].

[70] All the set-theoretic facts needed in this ensuing discussion of transitive models can be found in Button and Walsh [6, chap. 8]. *Notation*: We use calligraphic fonts for structures, and italics for their underlying domains; so A is the domain of $\mathcal{A}$. The definition of a transitive model is given in the model theory; so we use '$\in$', here, in the model theory, and use '$\in^{\mathcal{A}}$' for $\mathcal{A}$'s interpretation of $\in$'.

[71] Whenever we speak of ordinals in this subsection, we mean von Neumann ordinals.

[72] See Linnebo and Rayo [23, p. 279n.13].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

is simple: start with a transitive model of LT; treat entities of different rank as being of different types; and read membership as 'application'.

Unfortunately, there is a small wrinkle in implementing this plan, thanks to an irritating mismatch between the types of $\mathrm{CTT_p}$ and a set's rank. To illustrate: $\mathrm{CTT_p}$'s Limit-rule means that every type ω entity is of some finite type, but the ordinal ω has rank ω. To deal with this wrinkle, we define a function which (in effect) tells us how to map from ranks to types:

$$\alpha^* = \begin{cases} \alpha & \text{if } \alpha < \omega \\ \alpha + 1 & \text{if } \alpha \geq \omega. \end{cases}$$

We can now implement our plan. Where $\mathcal{A}$ is a transitive model of LT, define $\mathbf{C}\mathcal{A}$ as follows. Its denizens are just the members of A, and if $\mathrm{rank}(x) = \alpha$ then x is treated as a type β entity for all $\alpha^* \leq \beta < \mathit{Ord}(\mathcal{A})^*$. Then, we stipulate that $\mathbf{C}\mathcal{A} \models y^\alpha(x^\alpha)$ iff $\mathcal{A} \models x \in y$.

Lemma 15. *Let $\mathcal{A} \models LT$ be transitive. Then $\mathbf{C}\mathcal{A} \models CTT_p^{Ord(\mathcal{A})^*}$. Moreover, $\mathcal{A}$ is standard iff $\mathbf{C}\mathcal{A}$ is standard.*

Proof sketch. The quantifier-rules and Limit-rules are obviously sound. When $\equiv$ and ε are well-defined $\mathrm{CTT}^{Ord(\mathcal{A})^*}$-expressions,[73] distinct entities a^α and b^β are distinguished by $\{a^\alpha\}^{\alpha+1}$, so $\equiv^{\mathbf{C}\mathcal{A}}$ is identity and $\varepsilon^{\mathbf{C}\mathcal{A}}$ is membership. Type-Base and Type-Purity now hold, as $\mathcal{A}$ has exactly one rank-0 object, and it is empty. Type-Ext follows from Extensionality and simple reasoning about ranks. For CTT-Comprehension, fix ϕ and α with $\alpha + 1 < \mathit{Ord}(\mathcal{A})^*$; we will show that:

$$\mathbf{C}\mathcal{A} \models \exists z^{\alpha+1} \forall x^\alpha (z^{\alpha+1}(x^\alpha) \leftrightarrow \phi(x^\alpha)).$$

Let $\beta^* = \alpha + 1$; note that $\beta \in A$, as $\mathcal{A}$ is transitive. Fix $s \in A$ such that $\mathcal{A}$ thinks that s is the $\in$-least level with β as a subset. By Separation on s in $\mathcal{A}$, there is some $c \in A$ of rank $\leq \beta$ which serves as a witnessing value for $z^{\alpha+1}$ when regarded as an entity of type $\alpha + 1$. Finally, the remark about standardness is immediate from the construction. □

We now have the means to move between transitive models of LT and models of $\mathrm{CTT_p}$. Recalling that LT is strictly weaker than ZF, we can now settle the status of Replacement, in $\mathrm{CTT_p}$, by using some well-known facts concerning models of ZF:

Theorem 16. *Fix $\kappa > \omega$ such that $\kappa + 2 < \tau$:*

(1) *If κ is strongly inaccessible, every model of CTT_p^τ satisfies $ZF^{(\kappa)}$.*

(2) *If κ is not strongly inaccessible, there are models of CTT_p^τ which violate $ZF^{(\kappa)}$, and any model of CTT_p^τ which satisfies $ZF^{(\kappa)}$ is non-standard.*

Proof. (1) Let κ be strongly inaccessible with $\mathcal{M} \models \mathrm{CTT_p^\tau}$. Using Theorem 13 and Lemma 14, obtain a transitive model $\mathcal{A} \cong \mathbf{L}^\kappa\mathcal{M} \models \mathrm{Zr}$. To show that $\mathcal{M} \models \mathrm{ZF}^{(\kappa)}$, it suffices to show that $\mathcal{A} \models$ Replacement. So: fix $a \in A$ and suppose $\mathcal{A} \models (\forall x \in a)\exists!\, y\phi(x, y)$. Working outside $\mathcal{A}$, let

$$\beta = \sup\{\mathrm{rank}(c) : \mathcal{A} \models (\exists x \in a)\phi(x, c)\}.$$

[73] Recall from footnote 10 that $x^\alpha \ \varepsilon \ y^\beta$ is a CTT^τ-formula iff $\max(\alpha, \beta) + 2 < \tau$.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Since $\mathcal{A}$ is transitive and κ is strongly inaccessible, $\beta \in A$. So also $\{c : \mathcal{A} \models (\exists x \in a)\phi(x, c)\} \in A$, by Separation in $\mathcal{A}$ on what $\mathcal{A}$ thinks is a level with β as a subset.

(2) Here are two general facts about the V_α hierarchy:

$$V_\kappa \models \text{ZF} \text{ iff } \kappa \text{ is strongly inaccessible}$$
$$V_\kappa \models \text{LT} \text{ iff } \kappa > 0.$$

So suppose $\kappa > \omega$ is not strongly inaccessible. Fix σ such that $\sigma^* \geq \tau$. Now, V_σ is a transitive model of LT and $Ord(V_\sigma) = \sigma$. Using Lemma 15, obtain $\mathbf{C}V_\sigma \models \text{CTT}^\tau_{\text{p}}$. By construction, $\mathbf{L}^\kappa\mathbf{C}V_\sigma = V_\kappa$. Since $V_\kappa \not\models$ Replacement, also $\mathbf{C}V_\sigma \not\models \text{Replacement}^{(\kappa)}$.

For the second clause: suppose $\mathcal{M} \models \text{CTT}^\tau_{\text{p}}$ and $\mathcal{M} \models \text{ZF}^{(\kappa)}$, with κ not strongly inaccessible. Use Lemma 14 to obtain a unique transitive model $\mathcal{A} \cong \mathbf{L}^\kappa\mathcal{M}$. Since $\mathcal{A} \models$ ZF but $Ord(\mathcal{A})$ is not strongly inaccessible, $\mathcal{A}$ is non-standard. □

§C. Definitional equivalence for CTT$^\omega$. In Section 5.4, we stated that CTT$^\omega$ is definitionally equivalent to STT$_\uparrow$. In this appendix, we define STT$_\uparrow$, and prove the equivalence.

STT$_\uparrow$ augments STT with a new function symbol, $\uparrow$, for each n, which takes a type n entity as input and outputs a type $n+1$ entity.[74] So, for example, $b^2(\uparrow a^0)$ and $c^5(\uparrow\uparrow b^2)$ are well-formed. STT$_\uparrow$ retains STT-Comprehension; this holds for formulas containing $\uparrow$. STT$_\uparrow$ then has axioms ensuring that $\uparrow$ is injective, preserves property-possession, and delivers well-foundedness:

Up-Inject. $\forall x^n \forall y^n(\uparrow x^n = \uparrow y^n \rightarrow x^n = y^n)$
Up-Possess. $\forall x^n \forall y^{n+1}(\uparrow y^{n+1}(\uparrow x^n) \leftrightarrow y^{n+1}(x^n))$
Up-Founded. $\forall x^{n+1} \forall y^{n+1}(\uparrow y^{n+1}(x^{n+1}) \rightarrow \exists z^n\, x^{n+1} = \uparrow z^n)$
Up-Base. $\forall x^0 \forall y^0 \neg\uparrow y^0(x^0)$

For readability, where $m > n$, we write $\uparrow_m a^n$ for the result of applying $m - n$ instances of $\uparrow$ to a^n, yielding a type m entity; so $\uparrow_4 a^0$ abbreviates $\uparrow\uparrow\uparrow\uparrow a^0$, and $c^5(\uparrow_4 b^2)$ abbreviates $c^5(\uparrow\uparrow b^2)$. Simple induction, which we leave to the reader, shows that STT$_\uparrow$ proves generalizations of our new axioms; specifically, for each $m > n$:

- $\forall x^n \forall y^n(\uparrow_m x^n = \uparrow_m y^n \rightarrow x^n = y^n)$
- $\forall x^n \forall y^{n+1}(\uparrow_{m+1} y^{n+1}(\uparrow_m x^n) \leftrightarrow y^{n+1}(x^n))$
- $\forall x^m \forall y^{n+1}(\uparrow_{m+1} y^{n+1}(x^m) \rightarrow \exists z^n\, x^m = \uparrow_m z^n)$
- $\forall x^n \forall y^0 \neg\uparrow_{n+1} y^0(x^n)$.

To prove that STT$_\uparrow$ and CTT$^\omega$ are definitionally equivalent (Theorem 23), we first define an interpretation, I, from CTT$^\omega$ to STT$_\uparrow$. This preserves the interpretation of all logical symbols, including =; its only non-trivial action is as follows:[75]

$$[y^n(x^m)]^I := y^n(\uparrow_{n-1} x^m).$$

Observe that if $n = m + 1$, then $[y^n(x^m)]^I$ is just $y^n(x^m)$. Here is a very simple fact about the relationship between $\uparrow$ and the interpretations of ε and $\equiv$, which holds just by unpacking some definitions (the proof is left to the reader):

[74] As with the signs $=$, $\equiv$ or ε, we are using the same symbol (in a typically ambiguous way) for each type level.

[75] So: $[x^n = y^n]^I := x^n = y^n$; $[\phi \wedge \psi]^I := (\phi^I \wedge \psi^I)$; $[\neg\phi]^I := \neg\phi^I$; and $[\forall x^n \phi]^I := \forall x^n \phi^I$.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

LEMMA 17. *Where* $i = \max(m, n)$, *$STT_{\uparrow}$ proves*:

(1) $[x^m \equiv y^n]^I \leftrightarrow \uparrow_i x^m = \uparrow_i y^n$
(2) $[x^m \mathrel{\varepsilon} y^n]^I \leftrightarrow \uparrow_{i+1} y^n(\uparrow_i x^m)$

We now have one substantial result:

LEMMA 18. *The I-interpretations of* $\forall E^n_m$ *and* $\forall I^n_m$ *are admissible in* $STT_{\uparrow}$

Proof. We start with $\forall \mathrm{E}^n_m$. Suppose that both $\phi(x^n)$ and $\phi(a^m)$ are well-formed in CTT$^\omega$. Working in STT$_\uparrow$, suppose $\forall x^n \phi^I(x^n)$. If $m = n$, then $\phi^I(a^m)$ follows by ordinary $\forall$E in STT$_\uparrow$. So consider the case when $m < n$. The variable x^n cannot occur in any identity-claim, e.g. $x^n = c^n$, since $a^m = c^n$ is ill-formed in CTT$^\omega$; so ϕ must have this kind of shape (illustratively):

$$\psi(x^n(v^i), \dots, y^k(x^n), \dots)$$

with $i < m < n < k$; note that $i < m$, since $\phi(a^m)$ is well-formed in CTT$^\omega$. Now ϕ^I is:

$$\psi(x^n(\uparrow_{n-1} v^i), \dots, y^k(\uparrow_{k-1} x^n), \dots).$$

Using STT$_\uparrow$'s rule $\forall \mathrm{E}^n$, we can infer $\phi^I(\uparrow_n a^m)$, i.e.:

$$\psi(\uparrow_n a^m(\uparrow_{n-1} v^i), \dots, y^k(\uparrow_{k-1}\uparrow_n a^m), \dots).$$

Simplifying, and using generalized Up-Possess, we obtain:

$$\psi(a^m(\uparrow_{m-1} v^i), \dots, y^k(\uparrow_{k-1} a^m), \dots),$$

which is precisely $\phi^I(a^m)$, as required.

The admissibility of $\forall \mathrm{I}^n_m$ under interpretation follows straightforwardly. Given $\phi^I(b^n)$, with b^n suitably arbitrary: infer $\forall x^n \phi^I(x^n)$ using STT$_\uparrow$'s rule $\forall \mathrm{I}^n$; with a^m suitably arbitrary, infer $\phi^I(a^m)$ using $\forall \mathrm{I}^n_m$ under interpretation; finally, infer $\forall x^m \phi^I(x^m)$ using STT$_\uparrow$'s rule $\forall \mathrm{I}^m$. □

It is now easy to prove that I is an interpretation:

LEMMA 19. $I : CTT^\omega \longrightarrow STT_{\uparrow}$ *is an interpretation*.

Proof. We simply check all inference rules and axioms. Lemma 18 deals with the quantifier-rules, and no Limit-rules apply since we are considering CTT$^\omega$.

CTT$^\omega$-Comprehension. If $\phi(x^n)$ is an CTT$^\omega$-formula, then $\phi^I(x^n)$ is an STT$_\uparrow$-formula; now use STT$_\uparrow$-Comprehension.

Type-Founded. Suppose $[a^m \mathrel{\varepsilon} b^{n+1}]^I$ i.e. $\uparrow_{i+1} b^{n+1}(\uparrow_i a^m)$ with $i = \max(m, n+1)$ by Lemma 17.2. By generalized Up-Founded, there is z^n such that $\uparrow_i a^m = \uparrow_i z^n$, i.e., $[a^m \equiv z^n]^I$ by Lemma 17.1.

Type-Base. By generalized Up-Base, $\neg \uparrow_{n+1} y^0(x^n)$; so $[x^n \mathrel{\not\varepsilon} y^0]^I$ by Lemma 17.2. □

We now switch to working in CTT$^\omega$. It will help if we allow ourselves the use of a definite description operator, ι, within CTT$^\omega$. (This is harmless since, by standard Russellian techniques, this can always be eliminated from any formula.)[76] Now, by

[76] So we are relying on the fact that CTT$^\omega$ augmented with this device is definitionally equivalent to CTT$^\omega$. Clearly, it is. Still, for details of how to handle function symbols more austerely, see e.g., Button and Walsh [6, sec. 5.5, esp. fn.23].

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Type-Raising in CTT^{ω}, for any type n and each x^n there is a unique x^{n+1} such that $x^n \equiv x^{n+1}$; we will denote this in CTT^{ω} using $\Uparrow x^n$, i.e., $\Uparrow x^n := (\iota x^{n+1} \,.\, x^n \equiv x^{n+1})$. As before, we write $\Uparrow_n a^m$ for the result of applying $n - m$ instances of $\Uparrow$ to a^m, yielding a type n entity. We now define an interpretation, J, from $\mathrm{STT}_{\uparrow}$ to CTT^{ω}, with these actions:[77]

$$[y^{n+1}(x^n)]^J := y^{n+1}(x^n)$$
$$[\uparrow x^n]^J := \Uparrow x^n.$$

Lemma 20. *$J : STT_{\uparrow} \longrightarrow CTT^{\omega}$ is an interpretation.*

Proof. CTT^{ω}-Comprehension immediately licenses $\mathrm{STT}_{\uparrow}$-Comprehension. For Up-Inject, suppose $\Uparrow x^n = \Uparrow y^n$, i.e., $(\iota x^{n+1} \,.\, x^n \equiv x^{n+1}) = (\iota y^{n+1} \,.\, y^n \equiv y^{n+1})$; so $x^n \equiv y^n$ by Lemma 2, and hence $x^n = y^n$. Similarly, Up-Possess holds by Lemma 2. And Up-Founded and Up-Base hold via Type-Founded and Type-Base. □

It only remains to show that I and J together yield a definitional equivalence.

Lemma 21. *CTT^{ω} proves this scheme: $[[y^n(x^m)]^I]^J \leftrightarrow y^n(x^m)$.*

Proof. Note that $[[y^n(x^m)]^I]^J$ iff $[y^n(\uparrow_{n-1} x^m)]^J$ iff $y^n(\Uparrow_{n-1} x^m)$ iff $y^n(x^m)$, using Lemma 2 for the final biconditional. □

Lemma 22. *$STT_{\uparrow}$ proves these schemes: $[[y^{n+1}(x^n)]^J]^I \leftrightarrow y^{n+1}(x^n)$ and $[[\uparrow x^n]^J]^I = \uparrow x^n$.*

Proof. The first scheme is trivial. For the second:

$$\begin{aligned}
[[\uparrow x^n]^J]^I = [\Uparrow x^n]^I &= (\iota x^{n+1} \,.\, x^n \equiv x^{n+1})^I \\
&= (\iota x^{n+1} \,.\, \forall z^{n+2}(z^{n+2}(x^n) \leftrightarrow z^{n+2}(x^{n+1})))^I \\
&= (\iota x^{n+1} \,.\, \forall z^{n+2}(z^{n+2}(\uparrow x^n) \leftrightarrow z^{n+2}(x^{n+1}))) \\
&= (\iota x^{n+1} \,.\, \uparrow x^n = x^{n+1}) \\
&= \uparrow x^n.
\end{aligned}$$

□

Assembling Lemmas 19–22, we obtain:

Theorem 23. *$STT_{\uparrow}$ and CTT^{ω} are definitionally equivalent*

§D. Definitional equivalence for FJT. In Section 7, we stated that FJT is definitionally equivalent to $\mathrm{STT}_{\triangleright}$. In this appendix, we define $\mathrm{STT}_{\triangleright}$ and prove the equivalence.

The guiding idea is to simulate FJT by using a version of STT with this sort of behaviour: for all types $1 < m < n$, each type n entity a^n projects downwards to some type m entity $\downarrow_m a^n$; we can then simulate an application $a^n(x^{m-1})$ by instead considering $\downarrow_m a^n(x^{m-1})$. However, there is a small snag: we are treating $\downarrow_m$ as *functional*; but, if we assume no version of extensionality, then we will have no way to decide whether a^2 should project downwards to b^1 or c^1, if b^1 and c^1 are coextensional. The snag can be avoided by using a relational (rather than function) version of downward-projection. What follows spells this out rigorously.

[77] And: $[x^n = y^n]^J := x^n = y^n$; $[\phi \wedge \psi]^J := (\phi^J \wedge \psi^J)$; $[\neg\phi]^J := \neg\phi^J$; and $[\forall x^n \phi]^J := \forall x^n \phi^J$.

We define STT$_\rhd$ by augmenting STT as follows. For each $n > 0$, we have a relational constant, $\rhd$, expressing the downward-projecting relation from a type $n+1$ entity to a type n entity. So we write e.g. $e^4 \rhd d^3$ or $b^1 \rhd a^0$;[78] when convenient, we may write $d^3 \lhd e^4$ or $a^0 \lhd b^1$ instead. We introduce some useful abbreviations:

$$\begin{aligned} a^n \approx b^n &\text{ iff}_{\text{df}}\ \forall x^{n-1}(a^n(x^{n-1}) \leftrightarrow b^n(x^{n-1})), \text{when } n > 0 \\ a^n \blacktriangledown b^n &\text{ iff}_{\text{df}}\ \forall x^{n-1}(a^n \rhd x^{n-1} \leftrightarrow b^n \rhd x^{n-1}), \text{when } n > 1 \\ a^1 \blacktriangledown b^1 &\text{ iff}_{\text{df}}\ a^1 = a^1. \end{aligned}$$

So $a^n \approx b^n$ tells us that a^n and b^n are coextensive, and $a^n \blacktriangledown b^n$ tells us that a^n and b^n project downwards to exactly the same entities. The special stipulation for $a^1 \blacktriangledown b^1$ is needed as STT$_\rhd$ has *no* relational constant $\rhd$ expressing a relation from a type 1 entity to a type 0 entity, and so will hold vacuously. We concatenate chains of conjunctions; so we may write e.g., $a^2 \lhd d^3 \rhd b^2 \approx c^2$ in place of $(d^3 \rhd a^2 \land d^3 \rhd b^2 \land b^2 \approx c^2)$. STT$_\rhd$ retains the STT-Comprehension scheme for type 1 entities, i.e., $\exists z^1 \forall x^0(z^1(x^0) \leftrightarrow \phi(x^0))$; but for each $n > 1$, it has an augmented scheme:[79]

> STT$_\rhd$-*Comprehension*. $\forall y^n(\exists z^{n+1} \rhd y^n)\forall x^n(z^{n+1}(x^n) \leftrightarrow \phi(x^n))$, for any formula $\phi(x^n)$ not containing z^{n+1}.

For each $n > 0$, STT$_\rhd$ also has these axioms:

> *Down*$_\exists$. $\forall z^{n+1}\exists x^n\ z^{n+1} \rhd x^n$
> *Down*$_{Sim}$. $\forall z^{n+1}\forall x^n\forall y^n(x^n \lhd z^{n+1} \rhd y^n \to x^n \approx y^n \blacktriangledown x^n)$
> *Down*$_{Max}$. $\forall z^{n+1}\forall x^n\forall y^n(z^{n+1} \rhd x^n \approx y^n \blacktriangledown x^n \to z^{n+1} \rhd y^n)$

So Down$_\exists$ says that all entities of types ≥ 2 project downwards; Down$_{\text{Sim}}$ says that if z^{n+1} projects to two entities x^n and y^n, then x^n and y^n apply and project to exactly the same entities; and Down$_{\text{Max}}$ says that if z^{n+1} projects to some entity x^n, then z^{n+1} also projects to any y^n which applies and projects to exactly the same entities as x^n. These axioms ensure that $\rhd$-chains are always equivalent, in a strong sense which is brought out by these next two lemmas:

LEMMA 24. *STT$_\rhd$ proves this scheme (with $n > 0$). If $a^{n+1} \rhd x^n$ and $b^{n+1} \rhd x^n$ for some x^n, then $a^{n+1} \blacktriangledown b^{n+1}$.*

Proof. Suppose $a^{n+1} \rhd x^n$ and $b^{n+1} \rhd x^n$. If $a^{n+1} \rhd y^n$, then $x^n \approx y^n \blacktriangledown x^n$ by Down$_{\text{Sim}}$, so $b^{n+1} \rhd y^n$ by Down$_{\text{Max}}$; similarly if $b^{n+1} \rhd y^n$ then $a^{n+1} \rhd y^n$. □

LEMMA 25. *STT$_\rhd$ proves this scheme (with $n > 0$). Given any $\rhd$-chains*:

$$\begin{aligned} a^{n+1} \rhd a^n \rhd a^{n-1} \rhd \cdots \rhd a^1 \\ b^n \rhd b^{n-1} \rhd \cdots \rhd b^1. \end{aligned}$$

(1) *If $a^{n+1} \rhd b^n$, then $\bigwedge_{1 \leq i \leq n} a^i \approx b^i \blacktriangledown a^i$.*

[78] As with the signs $=$, $\equiv$ or ε, we are using the same symbol (in a typically ambiguous way) for each type level.

[79] It follows that some models of (plain vanilla) STT cannot be turned into models of STT$_\rhd$ just by assigning some meaning to '$\rhd$'. *Example*: it is consistent with STT that there are exactly four type 2 entities; whereas STT$_\rhd$ (and FJT) prove that there are at least eight type 2 entities.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

(2) *If there is* m *such that* $1 \leq m \leq n$ *and* $b^m \blacktriangledown a^m$ *and* $\bigwedge_{m \leq i \leq n} a^i \approx b^i$, *then* $\bigwedge_{m \leq i \leq n} b^i \blacktriangledown a^i$ *and* $a^{n+1} \rhd b^n$.

Proof. (1) From Down_{Sim}, by induction.

(2) By assumption, $a^{m+1} \rhd a^m \approx b^m \blacktriangledown a^m$, so $a^{m+1} \rhd b^m$ by Down_{Max}; so $b^{m+1} \blacktriangledown a^{m+1}$ by Lemma 24. This establishes a base case; the rest follows by induction. Now $a^{n+1} \rhd b^n$ by Down_{Max}. □

We will use these results to prove that $\text{STT}_{\rhd}$ and FJT are definitionally equivalent (Theorem 31). We first define an interpretation, I, to take us from FJT to $\text{STT}_{\rhd}$:[80]

$$[y^n(x^m)]^I := \forall y^{n-1} \forall y^{n-2} \cdots \forall y^{m+1}(y^n \rhd y^{n-1} \rhd y^{n-2} \rhd \cdots \rhd y^{m+1} \rightarrow y^{m+1}(x^m)).$$

Note that if $m + 1 = n$, then $[y^n(x^m)]^I$ is just $y^n(x^m)$.

Lemma 26. *I : FJT $\longrightarrow$ $\text{STT}_{\rhd}$ is an interpretation*

Proof. For all $0 \leq i < n$, let ϕ_i be FJT-formulas not containing z^n or z^j or y^j for any $0 \leq j < n$. (No generality is lost here, as we can relabel variables as necessary.) By multiple successive applications of $\text{STT}_{\rhd}$-Comprehension, there are $z^1 \lhd z^2 \lhd \cdots \lhd z^n$ such that $\forall x^i(z^{i+1}(x^i) \leftrightarrow \phi_i^I(x^i))$, for each $0 \leq i < n$. By Lemma 25.1, for each $0 \leq i < n$, we have:

$$\forall x^i(\forall y^{n-1} \cdots \forall y^{i+1}(z^n \rhd y^{n-1} \rhd \cdots \rhd y^{i+1} \rightarrow y^{i+1}(x^i)) \leftrightarrow \phi_i^I(x^i))$$
$$\text{i.e., } \forall x^i([z^n(x^i)]^I \leftrightarrow \phi_i^I(x^i)).$$

Conjoining these biconditionals and applying $\exists$I, we obtain $[\exists z^n \bigwedge_{i<n} \forall x^i(z^n(x^i) \leftrightarrow \phi_i(x^i))]^I$, i.e., an arbitrary instance of [FJT-Comprehension]^I. □

We now switch to working in FJT. We introduce another abbreviation, for a bounded version of coextensiveness, whenever $k \leq \min(m, n)$:

$$a^m \approxeq_k b^n \text{ iff}_{\text{df}} \bigwedge_{i<k} \forall x^i(a^m(x^i) \leftrightarrow b^n(x^i)).$$

Note the bound is $i < k$. We now define an interpretation, J, from $\text{STT}_{\rhd}$ to FJT:[81]

$$[y^{n+1}(x^n)]^J := y^{n+1}(x^n)$$
$$[y^{n+1} \rhd x^n]^J := y^{n+1} \approxeq_n x^n.$$

Lemma 27. *FJT proves the following schemes, where well-formed*:

(1) *If* $a^m \approxeq_k c^l$ *and* $b^n \approxeq_k c^l$, *then* $a^m \approxeq_k b^n$.
(2) $a^m \approxeq_k b^n$ *iff* $\forall x^k(a^m \approxeq_k x^k \leftrightarrow b^n \approxeq_k x^k)$.
(3) $a^n \approxeq_{n-1} b^n$ *iff* $[a^n \blacktriangledown b^n]^J$, *noting here that we must have* $n > 1$.

Proof. (1) Trivial.

(2) *Left-to-right.* Suppose $a^m \approxeq_k b^n$; if $a^m \approxeq_k c^k$, then $b^n \approxeq_k c^k$ by (1), and conversely. *Right-to-left.* Suppose $\forall x^k(a^m \approxeq_k x^k \leftrightarrow b^n \approxeq_k v^k)$; by FJT-Comprehension, there is some $c^k \approxeq_k a^m$; so $b^n \approxeq_k c^k$, and now $a^n \approxeq_k b^n$ by (1).

(3) Using (2), since $[a^n \blacktriangledown b^n]^J$ is $\forall x^{n-1}(a^n \approxeq_{n-1} x^{n-1} \leftrightarrow b^n \approxeq_n x^{n-1})$. □

[80] We choose variables to avoid clashes; I's other actions are trivial.

[81] We choose variables to avoid clashes in $[y^{n+1} \rhd x^n]^J$; J's other actions are trivial.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Lemma 28. *$J : STT_{\rhd} \longrightarrow FJT$ is an interpretation.*

Proof. For Down$_{\exists}$. Fix z^{n+1}; by FJT-Comprehension there is some $x^n \approxeq_n z^{n+1}$, i.e., $[z^{n+1} \rhd x^n]^J$.

For Down$_{Sim}$. Suppose $[x^n \lhd z^{n+1} \rhd y^n]^J$, i.e., $x^n \approxeq_n z^{n+1} \approxeq_n x^n$; so $x^n \approxeq_n y^n$ by Lemma 27.1. In particular, $x^n \approx y^n$, so $[x^n \approx y^n]^J$. Moreover, if $n > 1$ then $x^n \approxeq_{n-1} y^n$, so that $[x^n \blacktriangledown y^n]^J$ by Lemma 27.3; if $n = 1$ then $[x^n \blacktriangledown y^n]^J$ vacuously.

For Down$_{Max}$. Suppose $[z^{n+1} \rhd x^n \approx y^n \blacktriangledown x^n]^J$, i.e., $z^{n+1} \approxeq_n x^n \approx y^n \approxeq_{n-1} x^n$, using Lemma 27.3. So $y^n \approxeq_n x^n$, and hence $z^{n+1} \approxeq_n y^n$ by Lemma 27.1, i.e., $[z^{n+1} \rhd y^n]^J$.

For STT$_{\rhd}$*-Comprehension.* Let ϕ be any STT$_{\rhd}$-formula not containing z^{n+1} (but which may contain y^n). Fix y^n; by FJT-Comprehension, there is z^{n+1} such that:

$$\forall x^n(z^{n+1}(x^n) \leftrightarrow \phi^J(x^n)) \wedge \bigwedge_{i<n} \forall x^i(z^{n+1}(x^i) \leftrightarrow y^n(v^i))$$

$$\text{i.e., } \forall x^n(z^{n+1}(x^n) \leftrightarrow \phi^J(x^n)) \wedge z^{n+1} \approxeq_n y^n$$

$$\text{i.e., } [\forall x^n(z^{n+1}(x^n) \leftrightarrow \phi(x^n))^J \wedge z^{n+1} \rhd y^n]^J$$

So we have arbitrary instances of [STT$_{\rhd}$-Comprehension]^J. □

It only remains to show that I and J characterise a definitional equivalence.

Lemma 29. *FJT proves this scheme*: $[[y^n(x^m)]^I]^J \leftrightarrow y^n(x^m)$.

Proof. Note that the following are equivalent:

(1) $[[y^n(x^m)]^I]^J$
(2) $[\forall y^{n-1} \dots \forall y^{m+1}(y^n \rhd y^{n-1} \rhd \dots \rhd y^{m+1} \rightarrow y^{m+1}(x^m))]^J$
(3) $\forall y^{n-1} \dots \forall y^{m+1}(y^n \approxeq_{n-1} y^{n-1} \approxeq_{n-2} \dots \approxeq_{m+1} y^{m+1} \rightarrow y^{m+1}(x^m))$
(4) $y^n(x^m)$

The last equivalence uses Lemma 27.1, and repeated instances of FJT-Comprehension to provide a chain $y^n \approxeq_{n-1} a^{n-1} \approxeq_{n-2} \dots \approxeq_{m+1} a^{m+1}$. □

Lemma 30. *$STT_{\rhd}$ proves these schemes*: $[[y^{n+1}(x^n)]^J]^I \leftrightarrow y^{n+1}(x^n)$ *and* $[[y^{n+1} \rhd x^n]^J]^I \leftrightarrow y^{n+1} \rhd x^n$.

Proof. The first scheme is trivial. For the second, note that the following are equivalent:

(1) $[[y^{n+1} \rhd x^n]^J]^I$
(2) $[\bigwedge_{i<n} \forall v^i(y^{n+1}(v^i) \leftrightarrow x^n(v^i))]^I$
(3) $\bigwedge_{i<n} \forall v^i(\forall y^n \forall y^{n-1} \dots \forall y^{i+1}(y^{n+1} \rhd y^n \rhd y^{n-1} \rhd \dots \rhd y^{i+1} \rightarrow y^{i+1}(v^i)) \leftrightarrow$
$\forall x^{n-1} \dots \forall x^{i+1}(x^n \rhd x^{n-1} \rhd \dots \rhd x^{i+1} \rightarrow x^{i+1}(v^i)))$
(4) $y^{n+1} \rhd x^n$

For the last equivalence, first note that repeated use of Down$_{\exists}$ gives us chains:

$$y^{n+1} \rhd a^n \rhd a^{n-1} \rhd \dots \rhd a^1$$
$$x^n \rhd b^{n-1} \rhd \dots \rhd b^1$$

Using Lemma 25.1 twice, (3) is equivalent to:

(3′) $a^n \approx x^n \wedge a^{n-1} \approx b^{n-1} \wedge \dots \wedge a^1 \approx b^1$

Now Lemma 25.1 yields (4) ⇒ (3′), and Lemma 25.2 gives (3′) ⇒ (4). □

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

Assembling Lemmas 26–30, we obtain:

THEOREM 31. *$STT_{\triangleright}$ and FJT are definitionally equivalent*

Acknowledgements. Thanks to Neil Barton, Salvatore Florio, Peter Fritz, Luca Incurvati, Stephan Krämer, Øystein Linnebo, Nicholas Jones, Agustín Rayo, Thomas Schindler, Lukas Skiba, and an anonymous referee for *Review of Symbolic Logic*.

BIBLIOGRAPHY

[1] Bealer, G. (1982). *Quality and Concept*. Oxford: Oxford University Press.

[2] Beaney, M., editor (1997). *The Frege Reader*. Oxford: Blackwell.

[3] Boolos, G. (1985). Nominalist platonism. *The Philosophical Review*, **94**, 327–44.

[4] ———. (2000). Must we believe in set theory? In Sher, G., and Tieszen, R., editors. *Between Logic and Intuition*: *Essays in Honor of Charles Parsons*. Cambridge: Cambridge University Press, pp. 257–68.

[5] Button, T. (forthcoming). Level theory, part 1: Axiomatizing the bare idea of a cumulative hierarchy of sets. *Bulletin of Symbolic Logic*.

[6] Button, T., & Walsh, S. (2018). *Philosophy and Model Theory*. Oxford: Oxford University Press.

[7] Degen, W., & Johannsen, J. (2000). Cumulative higher-order logic as a foundation for set theory. *Mathematical Logic Quarterly*, **46**(2), 147–170.

[8] Feferman, S., editor (1995a). *Kurt Gödel*: *Collected Works*, Vol. 3. Oxford: Oxford University Press.

[9] ———. , editor (1995b). Note to Gödel (1933). In Feferman, S., editor. *Kurt Gödel*: *Collected Works*, Vol. 3. Oxford: Oxford University Press, pp. 36–44.

[10] Florio, S., & Jones, N. K. (2021). Unrestricted quantification and the structure of type theory. *Philosophy and Phenomenological Research*, **102**, 44–64.

[11] Florio, S., & Linnebo, Ø. (2021). *The Many and the One*: *A Philosophical Study of Plural Logic*. Oxford: Oxford University Press.

[12] Florio, S., & Shapiro, S. (2014). Set theory, type theory, and absolute generality. *Mind*, **123**(489), 157–174.

[13] Frege, G. (1891). Function and concept. In Beaney, M., editor (1997). *The Frege Reader*. Oxford: Blackwell, pp. 130–148.

[14] ———. (1892). On concept and object. In Beaney, M., editor (1997). *The Frege Reader*. Oxford: Blackwell, pp. 181–193.

[15] ———. (1893). *Die Grundgesetze der Arithmetik*, Vol. I. Jena: Pohle.

[16] Gaskin, R. (1995). Bradley's regress, the copula and the unity of the proposition. *The Philosophical Quarterly*, **45**, 161–180.

[17] ———. (2008). *The Unity of the Proposition*. Oxford: Oxford University Press.

[18] Gödel, K. (1933). The present situation in the foundations of mathematics. In Feferman, S., editor (1995a). *Kurt Gödel*: *Collected Works*, Vol. 3. Oxford: Oxford University Press, pp. 45–53.

[19] Incurvati, L. (2020). *Conceptions of Set and the Foundations of Mathematics*. Cambridge: Cambridge University Press.

[20] Koellner, Peter (2003). The Search for New Axioms. Ph.D. Thesis. Cambridge, MA: MIT.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

[21] Krämer, S. (2017). Everything, and then some. *Mind*, **126**(502), 499–528.

[22] Liebesman, D. (2015). Predication as ascription. *Mind*, **124**, 517–569.

[23] Linnebo, Ø., & Rayo, A. (2012). Hierarchies ontological and ideological. *Mind*, **121**(482), 269–308.

[24] ———. (2014). Reply to Florio and Shapiro. *Mind*, **123**(489), 175–181.

[25] MacBride, F. (2011). Impure reference: A way around the concept horse paradox. *Philosophical Perspectives*, **25**, 297–312.

[26] Maddy, P. (1988). Believing the axioms. I. *The Journal of Symbolic Logic*, **53**(2), 481–511.

[27] Magidor, O. (2009). The last dogma of type confusions. *Proceedings of the Aristotelian Society*, **109**, 1–29.

[28] Oliver, A., & Smiley, T. (2016). *Plural Logic* (second edition). Oxford: Oxford University Press.

[29] Potter, M. (2004). *Set Theory and its Philosophy*. Oxford: Oxford University Press.

[30] Prior, A. (1971). *Objects of Thought*. Oxford: Oxford University Press.

[31] Quine, W. v. O. (1970). *Philosophy of Logic*. Englewood Cliffs, NJ: Prentice-Hall.

[32] ———. (1951). Ontology and ideology. *Philosophical Studies*, **2**, 11–15.

[33] Rayo, A. (2006). Beyond plurals. In Rayo, A., and Uzquiano, G., editors. *Absolute Generality*. Oxford: Oxford University Press, pp. 220–254.

[34] Rayo, A., & Yablo, S. (2001). Nominalism through de-nominalization. *Noûs*, **35**, 74–92.

[35] Rieppel, M. (2016). Being something: Properties and predicative quantification. *Mind*, **125**(499), 643–689.

[36] Scott, D. (1974). Axiomatizing set theory. In Jech, T., editor. *Axiomatic Set Theory II*. Proceedings of the Symposium in Pure Mathematics of the American Mathematical Society, July–August 1967. Providence, RI: American Mathematical Society, pp. 207–214.

[37] Strawson, P. (1974). *Subject and Predicate in Logic and Grammar*. London: Methuen & Co.

[38] ———. (1987). Concepts and properties or predication and copulation. *The Philosophical Quarterly*, **37**, 402–406.

[39] Tait, W. W. (2001). Gödel's unpublished papers on the foundations of mathematics. *Philosophia Mathematica*, **9**, 87–126.

[40] Trueman, R. (2021). *Properties and Propositions*: *The Metaphysics of Higher-Order Logic*. Cambridge: Cambridge University Press.

[41] Whitehead, A., & Russell, B. (1910). *Principia Mathematica* (first edition), Vol. 1. Cambridge: Cambridge University Press.

[42] Wiggins, D. (1984). The sense and reference of predicates: A running repair of Frege's doctrine and a plea for the copula. *The Philosophical Quarterly*, **34**(136), 311–328.

[43] Williamson, T. (2003). Everything. *Philosophical Perspectives*, **17**(1), 415–465.

[44] ———. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

[45] Wright, C. (1998). Why Frege does not deserve his grain of salt. In Hale, B., and Wright, C., editors. *The Reason's Proper Study*. Oxford: Oxford University Press, pp. 72–90.

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press

[46] ———. (2007). On quantifying into predicate position: Steps towards a new(tralist) perspective. In Potter, M., editor. *Mathematical Knowledge*. Oxford: Oxford University Press, pp. 150–174.

UNIVERSITY COLLEGE LONDON
LONDON, UK
E-mail: tim.button@ucl.ac.uk

UNIVERSITY OF YORK
YORK, UK
E-mail: rob.trueman@york.ac.uk

https://doi.org/10.1017/S1755020321000435 Published online by Cambridge University Press