Journal of Philosophical Logic (2023) 52:1295–1325
https://doi.org/10.1007/s10992-023-09707-6

Internal Categoricity, Truth and Determinacy

Martin Fischer[1] · Matteo Zicchetti[2]

Received: 5 April 2022 / Accepted: 25 March 2023 / Published online: 3 August 2023

Abstract

This paper focuses on the categoricity of arithmetic and determinacy of arithmetical truth. Several 'internal' categoricity results have been discussed in the recent literature. Against the background of the philosophical position called *internalism*, we propose and investigate truth-theoretic versions of internal categoricity based on a primitive truth predicate. We argue for the compatibility of a primitive truth predicate with internalism and provide a novel argument for (and proof of) a *truth-theoretic* version of internal categoricity and internal determinacy with some positive properties.

Keywords Categoricity · Internalism · Axiomatic truth · Intolerance · Determinacy

1 Introduction

Philosophers of mathematics often have strong intuitions about arithmetic: arithmetic is usually accepted to be about a *particular* subject matter, the natural numbers. One usually speaks of *the* natural numbers indicating, informally, a certain uniqueness. Additionally, in most cases, *arithmetical truth* is accepted as *determinate* (at least) in the following sense: for any arithmetical statement, φ, either φ is true or its negation is. Dedekind's categoricity theorem for arithmetic has been employed as evidence to support the intuition of uniqueness. According to a model-theoretic, 'external' reading,[1] the categoricity theorem shows that all full models of second-order arithmetic are isomorphic.

[1] Although the external reading of Dedekind's theorem is the most frequent, there are alternative readings as explained in [19, Section 1, p.7].

✉ Matteo Zicchetti
matteo.zicchetti@gmail.com

Martin Fischer
M.Fischer@lrz.uni-muenchen.de

1 Munich Center for Mathematical Philosophy (MCMP), Ludwig-Maximilians-Universität München, München, Germany

2 Faculty of Philosophy, University of Warsaw, Warsaw, Poland

Recently, an alternative approach to the questions of uniqueness and determinacy gained some attention. This approach focuses on so-called ‘internal’ versions of categoricity by avoiding model-theoretic notions. This internal approach promises to provide a new, interesting route to answer the philosophical questions of the uniqueness of the natural numbers and of determinacy of truth. The two main approaches for our discussion are given in [22] as well as [23] and more recently in [4]. Parsons’ approach to internal categoricity is based on an open-ended conception of arithmetic and employs first-order arithmetical theories only. On the other hand, Button and Walsh prefer a second-order version based on pure second-order logic. Both of these variants have complications: whereas the first-order approach lacks expressive resources, the presuppositions of the second-order approach are unreasonably excessive.

Our purpose in this paper is to provide an improved version of internal categoricity arguments incorporating a primitive notion of truth. The main application will be on questions of determinacy as discussed by Parsons, and therefore we focus on first-order approaches. The main objective of our approach is to overcome some of the problematic issues of the first-order, Parsons-style approach. Additionally, we will argue that our version of internal categoricity circumvents some difficulties with respect to the second-order approach. The novelty of our approach lies in the introduction of a primitive truth predicate. In our *truth-theoretic* approach, the primitive, axiomatic notion of truth is going to expand the expressive resources given by first-order means, allowing to recover part of the second-order resources in a first-order setting.

The paper is structured as follows: Section 2 is divided into two parts: the first part introduces internalism, and in particular, it focuses on Parsons internalist position and strategy. The second part surveys two versions of internal categoricity (relevant for our investigation) and discusses their virtues and drawbacks: a version in pure second-order logic and a version over a first-order arithmetical theory. Section 3 discusses the issue of uniqueness and determinacy of arithmetic within the context of Parsons’ strategy. In particular, it focuses on the role played by both first and second-order approaches for the question of uniqueness and determinacy. In Section 4, we propose our approach to (and argument for) the internal categoricity of arithmetic and the determinacy of arithmetical truth. Our strategy is to adopt an axiomatic theory of truth. Section 4 aims to circumvent the drawbacks of the first-order approach (discussed in Sections 2 and 3). We do so by providing several truth-theoretic results (these are going to be our Propositions 1, 2 and 3). In the remainder of this article, we evaluate and discuss our results and address some possible questions and worries.

2 Internalism and Categoricity

Internalist approaches to categoricity aim to avoid the model-theoretic notions involved in the external formulation of Dedekind’s Theorem. When focusing on the case of arithmetic, internalist approaches try to avoid the problem of capturing the ‘full’ models by not talking about models at all and try to capture the relevant form of uniqueness in a formal framework internally.

Internalist approaches avoid the notion of a model altogether because, despite the external reading’s apparent attractiveness, its applicability to support the uniqueness thesis has been criticised by several authors for different reasons.[2] To circumvent some of the putative issues with external approaches to categoricity, philosophers have been proposing alternative, 'internal' versions of Dedekind's theorem. Although a clear and uniform demarcation of 'the' internalist position has not been provided yet, the distinction between 'internal' and 'external' approaches has some characteristic features.[3] Although internalist ideas are present already in works by Putnam and Parsons, the terminology has only been recently introduced in the discussion of internal categoricity theorems.[4] For our investigation, it is sufficient to present the *core* internalist tenet: the rejection of meta-theoretical, semantic notions, and importantly, the notions of a *model* and *satisfaction*, to understand the notion of *structure*.

Button and Walsh formulate the position of 'internalism', in contrast to what they call a 'modelist' position that uses model-theory to explicate structures. They summarise their internalist position with the slogan 'metamathematics without semantics!' [4, p. 227]. Therefore, internalism gives precedence to deductive methods: metamathematics 'should be undertaken within the logical framework of [the] very theories under investigation'. Button and Walsh carry out their proposal within the formal proof system of second-order logic, also called **CA** (to be introduced in Section 2.1).[5] Their internalism includes the rejection of semantic ascent in the form of a 'language-object satisfaction relation' [4, p. 226].

Parsons' approach shares some basic tenets, such as the rejection of model-theoretic tools in foundational investigations of arithmetic. This is motivated by the conviction that 'the idea of the natural numbers is more elementary than that of set' [23, p. 274]. Parsons aims to clarify the uniqueness claim for the natural numbers in the setting of communication:[6] he considers two mathematicians, Kurt and Michael, who both accept first-order arithmetic, and asks whether they must agree on arithmetical claims.[7] In order to argue for agreement, Parsons intends to employ a first-order categoricity theorem. Importantly, Parsons is interested in answering the question of agreement presupposing that the two mathematicians do not accept resources stronger than first-order arithmetic:

[2] For example, [4, pp. 159-60] suggest that employing the external reading of the theorem is question-begging. [24, p. 481] and [2] argue that the external reading is not immune from the so-called 'just more theory' argument. Notably, the putative issues with the external reading concern the understanding of 'full' in full second-order models. See also [23, §48] and [3, sections 3.3, 3.4] for further details and discussion of these issues. See [3, p. 291] for more pointers to the literature.

[3] This distinction is reminiscent of the distinction between internal and external questions [6]. Moreover, [2] points out that internalism is also inspired by [24]. An investigation of the 'history' of internalism would exceed the scope of this article and is also not relevant for this investigation.

[4] Core ideas of internalism are present in [24] and in [22], [23, p. 112] although Parsons focuses more specifically on the categoricity of arithmetic. Similar ideas in the context of set theory have been discussed by [17] and [20]. The term 'internal categoricity' is used by [28], but it can already be found earlier in [31].

[5] For a presentation of this system, see [4, p. 34]

[6] Parsons refers to [9] as motivation.

[7] Field describes this agreement claim as the so-called intermediate claim (IC): Any two people who accept schematic arithmetic must regard each other's theories as equivalent [12, Chapter 12, p. 357]

> What Parsons needs, then, is a categoricity theorem for arithmetic that relies on no resources beyond what's available to Kurt himself, no resources beyond those of number theory itself—what might be called a *pure* categoricity argument. [19, p. 28]

Despite the differences between the second-order and first-order approaches, we take the shared rejection of external, model-theoretic means to characterise arithmetical structures (for foundational purposes) as the starting point of our investigation. For establishing a uniqueness claim for arithmetic, we intend to stay within the logical framework under investigation, i.e. following Parsons, a first-order arithmetical framework. Of course, to have any hope to be successful in doing so, internalism must provide an understanding of an arithmetical structure not relying on external, model-theoretic tools.

Internalism intends to make sense of arithmetical structures without invoking semantic, language-object relations, such as satisfaction. According to a (second-order version of) internalism, being an arithmetical interpretation is understood employing a triple of elements with a one-place property, a first-order element, and a one-place function. This is expressed via an object-linguistic statement of the respective language

$$\mathsf{PA}(N, Sc, 0). \tag{1}$$

In the second-order case, $\mathsf{PA}(N, Sc, 0)$ is the conjunction of the second-order axioms of arithmetic. It is formulated in the pure language of second-order logic, so that 'N','0','Sc' are parameters of the right kind replacing the constants in a standard axiomatisation of $\mathbf{PA^2}$ as for example stated in [26, p. 4] or [4, p. 29]. In the context of the pure second-order logic internalist approach, the claim that (1) holds means that it is provable in the deductive system of second-order logic **CA**.

We refer to these statements, such as (1), as internal interpretations, whereas [4] call them 'internal structures', in order to contrast them with the 'external' understanding of structures. Externalism spells out what it means to be an arithmetical interpretation by providing a model $\mathfrak{M} = (M, \mathsf{S}^{\mathfrak{M}}, \mathsf{0}^{\mathfrak{M}})$ for the signature $\{\mathsf{S}, \mathsf{0}\}$, satisfying $\mathbf{PA^2}$:

$$(M, \mathsf{S}^{\mathfrak{M}}, \mathsf{0}^{\mathfrak{M}}) \models \mathbf{PA}^2 \tag{2}$$

That is, the external approach spells out arithmetical structures employing a satisfaction relation between the triple $(M, \mathsf{S}^{\mathfrak{M}}, \mathsf{0}^{\mathfrak{M}})$ and $\mathbf{PA^2}$. Here, satisfaction is understood as a semantic, language-object relation. Following this understanding of structure, we can see that the externalist has a strong, semantic understanding of what an arithmetical interpretation is.

It is crucial to see that (1) and (2) are different: in (1) there is no reference to an external semantic satisfaction relation, as it is completely expressed object-linguistically. This is also made explicit by Button and Walsh:

> In speaking informally of 'an arithmetical structure', here, the internalist is not aiming to draw attention to some specific object which stands in a language-object satisfaction relation to some theory. She is not engaging in semantic ascent. She is simply saying something in a second-order object language, along the

following lines: some property (a second-order entity), some (first-order) object, and some function (a second-order entity) collectively behave arithmetically. [4, p. 226]

Analogously, when considering an arithmetical sentence φ, the external approach spells out the claim that φ is satisfied in some arithmetical structure $(M, \mathsf{S}^{\mathfrak{M}}, \mathsf{0}^{\mathfrak{M}})$ as an external semantic language-world relation, in the following manner:

$$(M, \mathsf{S}^{\mathfrak{M}}, \mathsf{0}^{\mathfrak{M}}) \models \varphi \tag{3}$$

Internalism aims at internalising (3) as a deductive consequence of **CA**:

$$\mathsf{P}A(N, Sc, 0) \rightarrow \varphi(N, Sc, 0), \tag{4}$$

where $\varphi(N, Sc, 0)$ is the relativisation of φ by the parameters 'N','Sc', '0', i.e. all the quantifiers in φ, a sentence of the language of second-order arithmetic, are restricted to N, and the parameters uniformly replace the constants in φ. In other words, an external result (and understanding thereof) involves 'leaving' a theory's object-language to consider its semantics in some model-theoretic meta-language. On the other hand, an internal result is proved deductively and within the object language.

Although the internalist's motivation and intentions remain, in the first-order case, there are some complications due to the limited expressive resources of first-order logic. For example, in the first-order case, we do not have a finite axiomatisation of Peano arithmetic, but only a schematic one. This is going to make a difference for the formulation of the uniqueness claim.

2.1 Categoricity in Pure Second-Order Logic

Button and Walsh [4, p. 228] prove a version of internal categoricity in second-order logic: they take as the base theory the deductive system **CA** of pure second order logic with unrestricted second-order comprehension as a logical principle:

$$\exists X \forall x (x \in X \leftrightarrow \varphi(x)), \tag{COMP}$$

i.e. COMP holds for any φ in which X does not occur free, formulated in the empty signature.[8] That is, φ only contains connectives, two types of quantifiers, variables, parameters, and the identity sign.[9] Within the context provided by internalism, to be an arithmetical interpretation is expressed as $\mathsf{PA}(X, f, z)$, where X is a relation variable, f is a function variable, and z is an individual variable. $\mathsf{PA}(X, f, z)$ is the conjunction of the following axioms:

$$z \in X \wedge \forall x (x \in X \rightarrow (\exists! y (y \in X \wedge f(x) = y))) \tag{PA:1}$$

$$\forall x (x \in X \rightarrow f(x) \neq z) \tag{PA:2}$$

[8] For an exposition of **CA** see [4, p. 34]

[9] See [4, p. 224]

$$\forall x, y(x \in X \wedge y \in X \rightarrow (f(x) = f(y) \rightarrow x = y)) \quad \text{(PA:3)}$$

$$\forall Z \subseteq X(z \in Z \wedge \forall y(y \in X \rightarrow (y \in Z \rightarrow f(y) \in Z))) \rightarrow X = Z) \quad \text{(PA:4)}$$

Within **CA**, Button and Walsh formulate the relevant internal isomorphism between internal arithmetical structures with $\mathsf{ISO}(X, N, f, z, M, g, w)$, which is the conjunction of the following:

$$\forall x \forall y((x, y) \in X \rightarrow (x \in N \wedge y \in M)) \quad \text{(iso:1)}$$

$$\forall x \in N\, \exists! y \in M((x, y) \in X) \quad \text{(iso:2)}$$

$$\forall y \in M\, \exists! x \in N((x, y) \in X) \quad \text{(iso:3)}$$

$$(z, w) \in X \wedge \forall x, y((x, y) \in X \rightarrow (f(x), g(y)) \in X) \quad \text{(iso:4)}$$

Informally, (iso:1)–(iso:4) is understood as the claim that 'X is an internal isomorphism between the internal structures N, f, z and M, g, w'. With this notion of internal isomorphism at hand, Button and Walsh provide the following result:[10]

Theorem 1 [Button and Walsh] ***CA*** *proves*

$$\forall N f z \forall M g\, w(\mathsf{PA}(N, f, z) \wedge \mathsf{PA}(M, g, w) \rightarrow \exists X\, \mathsf{ISO}(X, N, f, z, M, g, w))$$

The idea of the proof is to consider all binary relations that relate the two 'zero elements' of any two internal interpretations, and that are 'closed under' the 'successor' functions in the relevant way. This is captured by the property of hereditariness defined by the second-order formula $H(Y)$:

$$H(Y) \leftrightarrow ((z, w) \in Y \wedge (\forall x \in N)(\forall y \in M)[(x, y) \in Y \rightarrow (f(x), g(y)) \in Y]),$$

The existence of the least element, i.e. the intersection, of all hereditary relations is guaranteed by the comprehension principle of **CA**:

$$\exists X \forall x \forall y((x, y) \in X \leftrightarrow \forall Z[H(Z) \rightarrow (x, y) \in Z]), \quad (\dagger)$$

This X in (†) satisfies (iso:1), (iso:2), (iso:3) and (iso:4). We want to draw the attention of the reader to the fact that the instance of comprehension employed for (†) is impredicative due to the universal second-order quantifier on the right-hand side. This issue of impredicativity will be relevant later on.

A virtue of Theorem 1 is its generality: one quantifies over all internal arithmetical structures and shows that all such internal structures are isomorphic. There is another sense in which this theorem is quite general, for it does not depend on any specific arithmetical language. Let us be explicit: although, as one can see, $\mathsf{PA}(X, f, z)$ reminds us of a finite axiomatisation of second-order arithmetic $\mathbf{PA^2}$, we should note that in these descriptions (of the internal arithmeticalstructures) of the form $\mathsf{PA}(X, f, z)$, no

[10] For the statement of the theorem, see [4, p. 228]: and the details of the proof see [4, Section 10.B]. The proof is basically an internalisation of the proof given by [25, Theorem 4.10] and is quite similar to the proof of the internal categoricity of second-order arithmetic provided by [30, Theorem 1]

fixed arithmetical vocabulary is either mentioned or used; these internal arithmetical structures are formulated in the language of pure second-order logic and therefore characterised by purely 'logical' properties modulo **CA**. As Button and Walsh point out, the focus on deductions in pure logic might lead to an unwanted interpretation of internalism as a form of if-thenism. Button and Walsh explicitly deny such a reading:

> That impression is simply *wrong*. Unlike if-thenists, internalists *affirm* PA_{int} [internalised Peano arithmetic] unconditionally. [4, p. 236][11]

Additionally, philosophers have expressed worries concerning the use of impredicative comprehension in the second-order approach. In particular, as also [19, p. 29] point out, Theorem 1 is not satisfactory from the perspective of Parsons' project. Moreover, [19] suggest that the pure second-order approach is not transparent in an informative sense: the 'bridge principles' between the internal arithmetical structures needed for the wanted categoricity are 'hidden' within the strong comprehension principles of **CA**. It is important to note that unrestricted comprehension is built into the logic to obtain the result in the form of a logical theorem. Finally, one might think, following [20], that the machinery of second-order logic with impredicative comprehension might be too 'extravagant'. The following section briefly surveys and discusses a first-order version of the categoricity theorem. As we will show, the first-order approach has its disadvantages. Therefore, one of our main aims of Section 4 is to provide a more acceptable, less extravagant and predicative version of Theorem 1, which – we believe – is more acceptable from the perspective of Parsons' project, that we will discuss in the next section. Additionally, we will argue that our approach improves on the first-order categoricity theorem.

2.2 Categoricity in First-Order Arithmetic

According to Parsons, 'Dedekind's theorem is essentially first-order' [23, §49, p. 281]. The idea behind his claim is that for the construction of the isomorphism it is sufficient to rely on a suitable form of recursion. The standard form of recursion for a fixed vocabulary is not sufficient, and therefore, in order to allow for induction and recursion to be applicable to a wider range, Parsons and others have opted for an open-ended understanding of schematic theories.[12] Such an open-ended conception of arithmetic does not settle on a fixed first-order arithmetical theory, and allows for suitable language expansions by 'definite' predicates that are then allowed within the induction schema.

It is illustrative to briefly present Parsons' argument in [23]. Suppose (with Parsons) that our language contains two arithmetical internal structures represented by the triples $(\mathsf{N}^i, \mathsf{S}^i, 0^i)$ and $(\mathsf{N}^j, \mathsf{S}^j, 0^j)$, where $\mathsf{N}^i, \mathsf{N}^j$ aresortal predicates, $\mathsf{S}^i, \mathsf{S}^j$ unary function

[11] Concerning the question of whether internalism is an if-thenism, there seems to be tension in the use of pure **CA** and the acceptance of PA_{int}. This question deserves separate attention and exceeds the scope of this article.

[12] For such approaches involving an open-ended conception, see [17, 20] and [22, 23].

symbols and 0^i and 0^j constants. The idea is now to define a function f from N^i into N^j by recursion on 0^i and S^i such that the following holds:

$$f(0^i) = 0^j; \quad \forall x(\mathsf{N}^i(x) \rightarrow [f(\mathsf{S}^i(x)) = \mathsf{S}^j(f(x))]) \tag{$*$}$$

By relying on open-ended induction for both domains one can prove the injectivity and surjectivity of f.[13] Parsons is not fully explicit in his formulation. Some of the technical details have been made explicit in [19] and [29]. For our discussion we present a more detailed sketch.

One considers a first order language $\mathcal{L}_A^{ij}$ (A for arithmetical) with a signature $(\mathsf{N}^i, \mathsf{S}^i, 0^i, +^i, \times^i, \mathsf{N}^j, \mathsf{S}^j, 0^j, +^j, \times^j)$ containing two primitive arithmetical vocabularies. The theory $\mathbf{PA}^i \cup \mathbf{PA}^j$ is then Peano arithmetic in the language $\mathcal{L}_A^{ij}$, where full induction is formulated for formulas φ of the mixed language $\mathcal{L}_A^{ij}$ for both number properties N^i as well as N^j, i.e.

$$\varphi(0^i) \wedge \forall x(\mathsf{N}^i(x) \rightarrow (\varphi(x) \rightarrow \varphi(\mathsf{S}^i(x)))) \rightarrow \forall x(\mathsf{N}^i(x) \rightarrow \varphi(x)) \tag{IND^i}$$

$$\varphi(0^j) \wedge \forall x(\mathsf{N}^j(x) \rightarrow (\varphi(x) \rightarrow \varphi(\mathsf{S}^j(x)))) \rightarrow \forall x(\mathsf{N}^j(x) \rightarrow \varphi(x)) \tag{IND^j}$$

In contrast to the second-order case the additional function symbols for addition and multiplication are necessary. Moreover, the existence of the isomorphism is established by a witnessing formula rather than an explicit existential claim. $\mathsf{ISO}_{i \cong j}(\chi)$ is the conjunction of the following:

$$\forall x \forall y(\chi(x, y) \rightarrow (\mathsf{N}^i(x) \wedge \mathsf{N}^j(y))) \tag{iso:1}$$
$$\forall x(\mathsf{N}^i(x) \rightarrow \exists! y(\mathsf{N}^j(y) \wedge \chi(x, y))) \tag{iso:2}$$
$$\forall y(\mathsf{N}^j(y) \rightarrow \exists! x(\mathsf{N}^i(x) \wedge \chi(x, y))) \tag{iso:3}$$
$$\chi(0^i, 0^j) \wedge \forall x, y(\chi(x, y) \rightarrow \chi(\mathsf{S}^i(x), \mathsf{S}^j(y))) \tag{iso:4}$$
$$\forall x, y, z, x', y'(\chi(x, x') \wedge \chi(y, y') \wedge \chi(x +^i y, z) \rightarrow z = x' +^j y') \tag{iso:5}$$
$$\forall x, y, z, x', y'(\chi(x, x') \wedge \chi(y, y') \wedge \chi(x \times^i y, z) \rightarrow z = x' \times^j y') \tag{iso:6}$$

In order to establish the isomorphism, a function $f : \mathsf{N}^i \rightarrow \mathsf{N}^j$ satisfying ($*$) is used. With this, one can establish in $\mathbf{PA}^i \cup \mathbf{PA}^j$ the injectivity of f by (IND^i) and its surjectivity by (IND^j). The properties of the isomorphism are then easy to prove. Then, one can state the categoricity theorem as follows:

Theorem 2 (Maddy, Väänänen) *There is a formula χ, such that*

$$\mathbf{PA} \cup \mathbf{PA}^i \cup \mathbf{PA}^j \vdash \mathsf{ISO}_{i \cong j}(\chi)$$

The theorem is formulated with an additional version of **PA**. For the proof sketch to be complete, it only remains to establish the existence of the function f by recursion.

[13] For Parsons' sketch see [23, pp. 281-2]

Naturally one would like to use formulas from the expanded arithmetical language $\mathcal{L}_A^{ij}$ to witness the existence of the functions by primitive recursion as for example in [13, Theorem 1.54]. The difficulty in doing so is that a coding of finite sequences for objects from both domains is required, and unfortunately, such a coding is not directly available.[14] The problem is that arithmetical functions are usually employed in a standard coding of finite sequences of natural numbers. However, whereas we have functions $+^i$, $\times^i$ on the numbers in N^i and $+^j$, $\times^j$ on N^j, we have no function reaching between the two 'domains', i.e. allowing for inputs from both. To circumvent this issue, one can for instance allow for an additional overarching system of natural numbers **PA** with signature $(0, 1, +, \times)$, as in [19, Theorem 9]. Informally, the idea there is to enable the 'bridging' coding by assuming that both domains are already included in an overarching domain of natural numbers N, i.e. $N^i \subseteq N$ and $N^j \subseteq N$.[15] With this assumption in place, one can use a standard coding of the overarching theory **PA** to code finite sequences of numbers of both sorts. In order to establish the existence of the witnessing formula for the function f defined by primitive recursion, Σ_1-induction for the expanded language, i.e. containing all three arithmetical vocabularies, suffices.[16]

Since first-order **PA** is not finitely axiomatisable, Theorem 2 cannot be stated as a logical theorem: the first-order version cannot state the isomorphism as a conditional statement, with the internal structures in the antecedent; the assumptions must be stated in the form of theories, such as $\mathbf{PA}^i \cup \mathbf{PA}^j$. As we have discussed, additional bridging principles, such as the distinguished **PA**, are necessary to obtain the wanted theorem. As the sketch of the proof indicates, only Σ_1-induction is used, such that $\mathbf{I\Sigma}_1$ seems sufficient for the syntactical bridging principles. How these principles are made explicit depends on the understanding of open-ended schematic theories. In the second-order version, the 'bridging principles' are implicit in the impredicative comprehension principles of **CA**.

By comparison with the second-order version of categoricity of the previous subsection, Theorem 2 faces some limitations. In contrast to the second-order version that quantifies over all internal interpretations, the first-order version is restricted to a schematic statement for two arbitrary arithmetical interpretations. An additional limitation of the first-order approach is that it is not possible to prove the existence of the wanted isomorphism explicitly. Due to these difficulties, the first-order version lacks the generality of the second-order version. When focusing on the questions of determinacy and agreement (to be introduced in the next section), we believe that the first-order approach is worth to be further explored, despite the mentioned difficulties.

3 Uniqueness, Determinacy, Agreement

Now that the two relevant versions of internal categoricity have been introduced we take a step back and discuss these results in more detail.

14 [19, p. 32] show that otherwise a counter-model based on Ehrenfeucht-Fraïssé games is constructable.

15 Compare [19, p. 35] for a similar remark about this.

16 An alternative strategy is sketched in [10], employing a particular version of open-ended theories allowing for a more general form of recursion. An alternative proof is to be found in [29, Theorem 14.]

Both versions of the categoricity theorems have been employed to underpin a uniqueness claim for the natural numbers. Following Parsons, it is promising to take uniqueness in the context of communication as reaching a form of agreement between different speakers.[17] Parsons considers two mathematicians, Kurt and Michael, who accept respectively a copy of first-order arithmetic. Parsons intends to argue for agreement about arithmetical statements employing a first-order, internal categoricity theorem. One question deserving more attention in Parsons' discussion is the following: how exactly is internal categoricity able to imply agreement?

Button and Walsh provide an answer to this question by relying on a corollary of their second-order version of categoricity, labelled *intolerance*.[18] Intolerance entails that for any arithmetical sentence φ, φ is evaluated in the same way across all internal structures: deviations are not tolerated. In this sense, intolerance is an internal version of elementary equivalence.

Remember that Button and Walsh are working in a pure second-order logical setting. The language $\mathcal{L}^2$ expands the first-order language by predicate-variables X, Y, Z and function-variables p, q and their respective second-order quantifiers $\forall X$ and $\exists Y$.

Theorem 3 (Button and Walsh) *For any* $\varphi \in \mathcal{L}^2$, **CA** *proves*

$$\mathsf{ISO}(R, N, f, z, M, g, w) \rightarrow [\varphi(N, f, z) \leftrightarrow \varphi(M, g, w)]$$

The proof is by induction on the complexity of φ. Although we allow for additional parameters (first and second-order) in the formula φ we omitted them in the formulation of the theorem for better readability.[19]

From the internalist's perspective, the theorem states that 'no object-language deviation between internal-structures is tolerated' [4, p. 232]. In other words, two internal structures must evaluate all arithmetical sentences uniformly. When considering the question of agreement, one can see that Theorem 3 provides a reasonable version of agreement between agents (accepting respectively an internal structure) given that they share the logical system of **CA**.[20]

Button and Walsh do not stop at Theorem 3, and argue for the following corollary as the wanted version of intolerance:

Corollary 1 *For any second-order formula* φ *whose only free variables are* N, f, z, *and whose quantifiers are restricted to* N, **CA** *proves:*

$$\forall N \forall f \forall z\, (\mathsf{PA}(N, f, z) \rightarrow \varphi(N, f, z)) \vee \forall N \forall f \forall z\, (\mathsf{PA}(N, f, z) \rightarrow \neg\varphi(N, f, z))$$

The argument for this corollary is simple. Informally, if we assume the negation of the left disjunct, then classically, this is nothing but the claim that there is an internal

[17] See for [23, §48] Related questions could be of broader interest. One of the problems according to McGee is the 'doxological' problem [20] of ensuring that our mathematical terms manage to refer to a unique concept or structure.

[18] Button and Walsh also discuss how internal categoricity obviously fails to 'pin down the natural numbers' [4, p. 231] when structures are understood externally. However, this is not a problem for our investigation, since we are focusing on internalism.

[19] For a proof of Theorem 3 see [4, p. 245]

[20] For a discussion see [2, p. 180]

structure, such that $\neg\varphi$. By Theorem 3, this implies that in all internal structures $\neg\varphi$ holds.[21] For the remainder of the article, in our discussion intolerance, we focus on Theorem 3. The reason is that we think that this already contains the interesting fact that deviations are not tolerated, and this points towards agreement.

Button and Walsh 'commend second-order logic [...] to Parsons, as a clean route to achieving his philosophical ends' [4, p. 241]. The main advantage of the second-order logic approach according to [4, p.240] is that it allows one to formulate these relevant results in a 'crisp manner, without any threat of semantic ascent'.

Despite the apparent attractiveness of Button and Walsh's suggestion, there are several reasons for Parsons to be hesitant in adopting such a second-order approach for his purposes. When discussing the use of second-order logic, Parsons claims the following:

> [It] would raise a question whether second-order quantifiers, in particular their use to define new predicates as in second-order logic, can have a definite sense. The implication of our earlier discussions of second-order logic is that any such sense that would license impredicative logic must derive from the concept of set. That such second-order logic is not forced on us at this point is shown by the fact that a mathematics that assumes the concept of natural number but from there on is entirely predicative is perfectly coherent. [23, §47, p. 270]

The quoted passage points towards two issues: first of all, Parsons seems to be worried that the employment of second-order logic to provide agreement might be in some sense circular or question-begging; Parsons thinks that an adequate understanding of impredicative second-order logic, specifically comprehension, requires an understanding of the concept of set, which is – in Parsons' view – not as basic, or primary, as the concept of the natural numbers itself.[22] So, when providing an argument for agreement about arithmetical statements, it would be question-begging to assume an understanding that already significantly exceeds our understanding of the natural numbers.

The second issue of Parsons concerns the claim that impredicative second-order logic is not necessary for an understanding of the natural number concept along the lines of Parsons' explanation in his [23, § 48]. This is in line with his restriction to a first-order framework within his broader motivation.[23]

The sceptical remarks towards impredicative second-order comprehension lead naturally to the question of whether we can minimize the background assumptions. Since the comprehension applied in the second-order arguments is impredicative we could ask whether a proof by predicative means is possible.[24] Whereas Π^1_1-comprehension

[21] Assuming classical reasoning the argument is simple as it stands. However, in the context of questions of determinacy it might be interesting to reconsider the role of classical principles such as the law of excluded middle.

[22] In [23, §48 p. 274] it is claimed that 'it is commonplace in the foundations of mathematics that the idea of natural number is more elementary than that of set'.

[23] In [19, p. 28] Parsons' restriction to a first-order arithmetical framework is understood as in the spirit of 'purity' of methods. See for instance [8].

[24] Although the result in [27] is clearly predicative, it fails to be fully general due to the assumption that the natural number properties are contained in $\mathbb{N}$. We will discuss this issue later.

(with parameters) appears to be sufficient for the second-order categoricity and intolerance results it is not known to us whether the assumptions could be weakened to something predicatively acceptable, such as Δ^1_1-comprehension.[25]

One of the advantages of the pure second-order version is its clarity concerning the tools employed. The boundary between pure second-order logic and mathematical theory is clear enough to provide 'crisp' results, such as Theorems 1 and 3. However, there is an additional worry regarding the applicability of the pure second-order version of categoricity to the case of agreement. Although Theorem 1 is a purely logical statement, Button and Walsh intend to apply it to Parsons' scenario. They refrain from an interpretation of their approach as a form of if-thenism, and claim that arithmetic is endorsed. It is unclear to us how this is achieved in the pure second-order logic setting. A possibility for the second-order proponent is of course to simply transfer the antecedent into an assumed arithmetical theory in addition to the logical axioms. With such a move the if-thenist interpretation is avoided, but also the purported advantages over the first-order version are lost. On the one hand, in order to apply this approach to Parsons' scenario, one would have to reformulate the theorems schematically, that is, for two given theories, and therefore would lose their purported generality. On the other hand, the clear restriction to a purely logical framework is no longer present. We think that these remarks provide additional reasons for further inquiry into suitable first-order approaches to categoricity.

However, sticking with a first-order approach welcomes an additional challenge: since first-order logic alone is insufficient for such purposes, the restriction to first-order frameworks necessitates the employment of additional means, in this case theories. Ideally, one would prefer to rely only on resources that do not exceed what is accessible to Parsons' mathematicians, Kurt and Michael. For Parsons' understanding, it is clear that first-order arithmetical resources, in the form of Peano arithmetic, are unproblematic. However, it is well-known that first-order arithmetic alone is insufficient to provide the wanted results.

For a first-order internalist approach to succeed, one must extend the range of available resources. Obviously not all possible additional resources are suited for Parsons' purposes. For instance, reintroducing strong comprehension principles via an impredicative set theory would violate his internalist picture. It is an additional challenge for an internalist in the spirit of Parsons to provide a reasonable boundary for theories that are close enough to first-order arithmetic to be acceptable. Deciding these questions also depends on an interpretation of 'wanted results'. We have seen that in the second-order version we have a general result of intolerance. General in the sense that we quantify over all internal structures. In contrast to this generality, the first-order version of internal categoricity can only be stated schematically for two given internal arithmetical structures. An arithmetical first-order version of intolerance could be formulated along the following lines:

$$\mathbf{PA} \cup \mathbf{PA}^i \cup \mathbf{PA}^i \vdash \varphi^i \leftrightarrow \varphi^j \tag{+}$$

[25] In [4, p. 247] Theorem 10.5 answers the question for 'weak' predicative comprehension principles allowing only second-order parameters negatively.

In this case, we are not able to quantify over all internal structures, but we can only formulate the result schematically over two internal structures given by $\mathbf{PA}^i$ and $\mathbf{PA}^j$. This formulation of intolerance is *prima facie* less general than its counterpart in pure second-order logic. With respect to this issue, Button and Walsh argue that 'lacking second-order resources, it is not immediately clear what general result we are supposed to be pointed to, by considering the specific interaction between Kurt and Michael' [4, p. 240]. We think that a first-order approach cannot reach a result exactly as general as the one in pure second-order logic. However, following Button and Walsh's challenge, a desideratum for an adequate first-order version could be to regain some generality, possibly 'close enough' to the generality of the second-order version, but within the bounds of a predicatively acceptable framework.

Additionally, the first-order version of intolerance fails to be general from the perspective of agreement for the following reason: the statement of intolerance (+) is given 'locally', in the sense that, *for any fixed* φ, Kurt and Michael are going to agree about φ, modulo the relativisation to i, j. The statement (+) does not quantify over the arithmetical sentences φ. By inspection of Theorem 3, it is clear that even the pure second-order version encounters a similar issue.

A more fundamental worry concerning the overall feasibility of Parsons' agreement strategy has been raised in [12]. The issue addresses the justification for the underlying assumptions employed in the argument for agreement and charges Parsons' strategy of being question-begging. Field argued that employing an open-ended conception of arithmetic is only sufficient for Kurt to show that his numbers can be mapped injectively into Michael's.[26] In contrast, to establish the surjectivity of such a mapping, further assumptions need to be in place. And these further assumptions are – according to Field – question-begging.[27] Parsons acknowledges the problem as based on a version involving some form of 'radical translation' between Kurt and Michael, but argues that the problem is no longer pressing if one assumes that Kurt can expand his language in the appropriate sense. For the sake of this investigation we will be charitable about Parsons' strategy and follow his suggestion.[28] Field's criticism is important, and the assumption of an overarching arithmetical theory in the first-order case deserves further discussion. We think that making these assumptions as explicit as possible is essential to further the discussion on this issue.

In the next section, we will introduce a notion of primitive truth with the aim to circumvent some of the expressive limitations of the first-order approach (discussed in Sections 2.2 and 3) to categoricity and intolerance.

4 Internal Categoricity and Truth

In this section, we introduce a primitive notion of truth by extending our arithmetical theories to an axiomatic theory of truth. The blueprint will be a typed Tarskian compo-

[26] See [12, Chapter 12, Postscript].

[27] Button and Walsh also discuss Field's worry in [4, p. 241].

[28] We think that Field's worry is important in general, but not essential for this investigation. To adjudicate the issue between Field and Parsons would exceed the scope of this article.

sitional theory of truth, known in the literature as **CT**.[29] It is well-known that the use of a truth predicate improves the expressiveness of the underlying base systems. Truth theories are able to mimic some forms of comprehension and therefore it appears not implausible that the truth predicate could be a useful device for our purposes.[30]

The main aim of this section is to provide a first-order version of categoricity that is able to overcome some of the limitations of Theorem 2: we improve on the first-order version of categoricity by expanding the isomorphism to include truth itself. We show that the isomorphism can be lifted from the two given internal arithmetical interpretations to their respective concepts of truth (Proposition 1). After that, we show that employing a truth predicate allows us to prove the isomorphism as an existential claim (Proposition 2). The remainder of this section, Section 4.4, focuses on providing a more general version of categoricity and intolerance. For this, we adopt an asymmetric framework: one distinguished arithmetical structure is taken as given, whereas the alternative structures are represented via an arbitrary sortal predicate P that plays a role similar to a second-order parameter. With Proposition 3 we show that for an arbitrary alternative internal arithmetical structure, we can always find an isomorphism to our distinguished structure. Finally, this allows us to provide a more general version of intolerance (Corollary 2) and to establish a stronger, uniform version of agreement.

4.1 Technical Preliminaries and Conventions

In this section, we expand the first-order arithmetical categoricity theorem to the setting of theories of truth. In order to do this, we will expand our mixed arithmetical theory of arithmetic $\mathbf{PA}^i \cup \mathbf{PA}^j$ (Section 2.2) by two compositional theories of truth. The language of our theory contains, like in the arithmetical case, a distinguished arithmetical vocabulary, as well as two arithmetical vocabularies i, j, which are now expanded by two additional truth predicates T^i, T^j. In the informal intended reading, the expression '$\mathsf{T}^i(t)$' states the truth of the sentence φ from the arithmetical language indexed by i, where t is a term denoting φ. The intended reading of the expression '$\mathsf{T}^j(t)$' is analogous. For concreteness we distinguish between the following signatures: $\mathrm{SIG}_A := \{0, \mathsf{S}, +, \times\}$; $\mathrm{SIG}^i_A := \{\mathsf{N}^i, 0^i, \mathsf{S}^i, +^i, \times^i\}$; $\mathrm{SIG}^j_A := \{\mathsf{N}^j, 0^j, \mathsf{S}^j, +^j, \times^j\}$.

Our mixed language $\mathcal{L}^{ij}_{\mathsf{T}^{ij}}$ is based on the union of the three signatures with the additional truth predicates and a finite set SYN of additional function and predicate symbols representing primitive recursive syntactic sets and functions (to be stated explicitly in the following). We will be working in the following signature:

$$\mathrm{SIG}^{ij}_{\mathsf{T}} := \mathrm{SIG}_A \cup \mathrm{SIG}^i_A \cup \mathrm{SIG}^j_A \cup \{\mathsf{T}^i, \mathsf{T}^j\} \cup \mathrm{SYN}.$$

The intended range of the truth predicate T^i (respectively T^j) are the sentences of the arithmetical sub-language $\mathcal{L}^i_A$ based on SIG^i_A (respectively $\mathcal{L}^j_A$ based on SIG^j_A). The distinguished arithmetical language $\mathcal{L}_A$ is based on SIG_A.

[29] See for example [14, Chapter 8, p. 63] for an introduction to this topic.

[30] For example, **CT** interprets **ACA** a second-order systems of arithmetic based on arithmetical comprehension. See for [14, Chapter 8.6].

In order to have suitable means to make these restrictions precise we assume that our vocabulary contains suitable predicate and function symbols representing primitive recursive syntactic operations.[31] In the usual setting of arithmetic, i.e. for theories containing $\mathbf{I\Sigma}_1$, language expansion by function symbols for primitive recursive functions is unproblematic.[32] With our additional assumption of an overarching syntax theory it is possible to do the same now for restricted parts of the language. We assume a standard Gödelnumbering #, identifying symbols of both signatures $\mathrm{SIG}^i_A \cup \mathrm{SIG}^j_A$ with our distinguished numbers.

So we allow for the following predicates for $k \in \{i, j\}$ in SYN: a predicate symbol $\mathsf{ct}^k(x)$, representing the decidable set of closed terms of the language $\mathcal{L}^k_A$; a predicate symbol $\mathsf{form}^k(x)$, representing the decidable set of formulas of the language $\mathcal{L}^k_A$, and a predicate symbol $\mathsf{sent}^k(x)$, representing the decidable set of sentences of the language $\mathcal{L}^k_A$. We also include function symbols for primitive recursive term and formula building operations in SYN. We use a standard convention employing a dot to indicate the representation: for example $\dot{\neg}(x)$ represents the primitive recursive function that takes the Gödelnumber of a formula φ and returns the Gödelnumber of its negation $\neg\varphi$. Similar for $\dot{=}(x, y)$, $\dot{\mathsf{N}}^k$, $\dot{\wedge}(x, y)$ and $\dot{\forall}(x, y)$. We abuse the notation and use the more intuitive $(x \dot{=} y)$ for $\dot{=}(x, y)$, and analogous for the other function symbols. Similarly, we have for $k \in \{i, j\}$ the representation of the term building functions $\dot{\mathsf{S}}^k$, $\dot{+}^k$, $\dot{\times}^k$.

Additionally we allow for the following special function symbols in SYN: for $k \in \{i, j\}$ the function symbol $\mathsf{num}^k(x)$ represents the primitive recursive function, that takes a distinguished number and returns the Gödelnumber of its k-numeral. The k-numerals are the numerals build up from 0^k and S^k, so for example $\overline{n}^j$ is short for $\underbrace{\mathsf{S}^j ... \mathsf{S}^j}_{n-times} 0^j$. Moreover, we also have the function symbols $\mathsf{sub}(x, y, z)$ representing the primitive recursive substitution function that for the inputs of the Gödelnumbers of a formula $\#\varphi$, a variable $\#x$ and a term $\#t$, outputs the Gödelnumber of the result of uniformly substituting all occurrences of the variable x in φ by t, i.e. $\#\varphi(x/t)$. Finally, we also allow for a valuation function symbol $\mathsf{val}(x)$ that represents the function that for the input $\#t$ for a closed terms t, returns t itself. For example if we have $\overline{n}^i$, then $\mathsf{val}(\overline{\#(\overline{n}^i)}) = \overline{n}^i$.

We frequently make use of relativisations φ^i, where φ^i is to be understood as the result of substituting in the formula φ from $\mathcal{L}_A$ the respective vocabulary by its i-counterparts and relativising all quantifiers to N^i. We make use of the relativisation φ^j analogously.

4.2 Unique Truth

The aim of this section is to establish the internal categoricity in the theory $\mathbf{PA} \cup \mathbf{CT}^i \cup \mathbf{CT}^j$. For $k \in \{i, j\}$, the truth-theoretic axioms of $\mathbf{CT}^k$ are the conjuction of

[31] Such a method is employed in several investigations of axiomatic theories of truth as for example in [5] and [14].

[32] These primitive recursive functions are Σ_1-definable, without increase of quantifier complexity, and their properties are provable in $\mathbf{I\Sigma}_1$. See for example [1, p. 88]

the universal closures of the following:

$$(\mathsf{CT1}^k) \qquad \mathsf{ct}^k(x) \wedge \mathsf{ct}^k(y) \rightarrow (\mathsf{T}^k(x \mathbin{\dot{=}} y) \leftrightarrow \mathsf{val}^k(x) = \mathsf{val}^k(y))$$

$$(\mathsf{CT2}^k) \qquad \mathsf{sent}^k(\dot{\mathsf{N}}^k x) \rightarrow (\mathsf{T}^k(\dot{\mathsf{N}}^k x) \leftrightarrow \mathsf{N}^k(\mathsf{val}^k(x)))$$

$$(\mathsf{CT3}^k) \qquad \mathsf{sent}^k(x) \rightarrow (\mathsf{T}^k(\dot{\neg} x) \leftrightarrow \neg \mathsf{T}^k(x))$$

$$(\mathsf{CT4}^k) \qquad \mathsf{sent}^k(x) \wedge \mathsf{sent}^k(y) \rightarrow (\mathsf{T}^k(x \mathbin{\dot{\wedge}} y) \leftrightarrow \mathsf{T}^k(x) \wedge \mathsf{T}^k(y))$$

$$(\mathsf{CT5}^k) \qquad \mathsf{sent}^k(\dot{\forall} v\, x) \rightarrow (\mathsf{T}^k(\dot{\forall} v\, x) \leftrightarrow \forall y(\mathsf{T}^k(\mathsf{sub}(x, v, \mathsf{num}^k(y)))))$$

$$(\mathsf{CT6}^k) \qquad \mathsf{T}^k(x) \rightarrow \mathsf{sent}^k(x)$$

The theory $\mathbf{CT}^k$ is then the extension of the arithmetical base theory, given by the k-relativised axioms of $\mathbf{Q}$ with the schema of induction IND^k allowing for formulas φ of our mixed language $\mathcal{L}_{\mathsf{T}}^{ij}$:

$$\varphi(0^k) \wedge \forall x(\mathsf{N}^k(x) \rightarrow (\varphi(x) \rightarrow \varphi(\mathsf{S}^k(x))) \rightarrow \forall x(\mathsf{N}^k(x) \rightarrow \varphi(x))$$

So we allow for both truth predicates to appear in both induction principles. The distinguished theory **PA** consists of the axioms of $\mathbf{Q}$ for the language $\mathcal{L}_A$ plus the induction schema on the distinguished numbers, so without any relativisation via sortal predicates, for formulas φ of the mixed arithmetical language $\mathcal{L}_A^{ij}$. Similar to the case of Theorem 2 it seems plausible that the bridging assumptions can be reduced to the theory $\mathbf{I\Sigma}_1$.

We want to expand the isomorphism f of Theorem 2 in a natural way. The isomorphism that we want to establish is now expressed by $\mathsf{ISO}^{\mathsf{T}}_{i \cong j}(\chi)$, which is the conjunction of (iso:1)–(iso:6) from (Section 2.2) and the additional conjunct:

$$\forall x \forall y(\mathsf{sent}^i(x) \wedge \mathsf{sent}^j(y) \wedge \chi(x, y) \rightarrow (\mathsf{T}^i(x) \leftrightarrow \mathsf{T}^j(y))) \qquad \text{(iso:7)}$$

In order to have suitable means to carry out the inductive argument we will make use of the logical complexity of formulas. Since the functions mapping Gödelnumbers of formulas of the respective languages $\mathcal{L}_A^k$ to their logical complexity is primitive recursive, we assume that we have function symbols lc^k in our language. We take $\mathsf{lc}^k(t) = 0$ if t is not a formula, and $\mathsf{lc}^k(t) = 1$ if t represents an atomic formula of $\mathcal{L}_A^k$. The logical complexity of complex formulas is as expected, i.e. the logical complexity of a conjunction is the maximum of the the logical complexities of each conjunct plus one, and the logical complexity of a negated formula or a universally quantified formula is the logical complexity of the formula plus one.

The construction of the witnessing formula for the isomorphism proceeds basically analogously to the arithmetical first-order case. In order to guarantee that the isomorphism also preserves the truth set we add some additional properties on our isomorphism F by identifying the relevant syntactical primitives. We add suitable additional clauses in the definition of our witnessing formula χ, so that for all syntactical constants e, $F(\#e^i) = \#e^j$. So with the assumption of our Gödelnumbering this means for example that $F(\#0^i) = \#0^j$. Moreover, we ensure that F preserves the term-forming operations, i.e. for $f(x) = y$ as short for $\chi(x, y)$ representing our F,

f commutes with the syntactic operations, for example $f(x \dot{+}^i y) = f(x) \dot{+}^j f(y)$. Then we get the following properties provable in $\mathbf{PA} \cup \mathbf{CT}^i \cup \mathbf{CT}^j$:

$$\mathsf{ct}^i(x) \to f(\mathsf{val}^i(x)) = \mathsf{val}^j(f(x)) \tag{6}$$

$$\mathsf{ct}^i(x) \wedge \mathsf{ct}^i(y) \to (\mathsf{val}^i(x) = \mathsf{val}^i(y) \leftrightarrow \mathsf{val}^j(f(x)) = \mathsf{val}^j(f(y))) \tag{7}$$

This suffices to expand our isomorphism to all the syntactical predicates and functions since the first-order version of the isomorphism establishes that for all $\varphi(x) \in \mathcal{L}_A$:

$$\forall x(\varphi^i(x) \leftrightarrow \varphi^j(f(x))) \tag{‡}$$

Moreover, since the formula build-up in both arithmetical languages $\mathcal{L}^i_A$ and $\mathcal{L}^j_A$ is completely parallel, there are primitive recursive functions for the transformation of φ^i into φ^j.[33] This enables a formulation of χ in such a way that it is closed under subformulas, again provable in $\mathbf{PA} \cup \mathbf{CT}^i \cup \mathbf{CT}^j$:

$$\chi(\dot{\neg} x, \dot{\neg} y) \to \chi(x, y) \tag{8}$$

$$\chi(x \dot{\wedge} y, x' \dot{\wedge} y') \to \chi(x, x') \wedge \chi(y, y') \tag{9}$$

$$\chi(\dot{\forall} vx, \dot{\forall} vy) \to \forall z \forall z' \chi(\mathsf{sub}(x, v, \mathsf{num}^i(z)), \mathsf{sub}(y, v, \mathsf{num}^j(z'))) \tag{10}$$

Proposition 1 *There is some* $\chi \in \mathcal{L}^{ij}_{\mathsf{T}}$*, such that*

$$\mathbf{CT}^i \cup \mathbf{CT}^j \cup \mathbf{PA} \vdash \mathsf{ISO}^{\mathsf{T}}_{i \cong j}(\chi)$$

Proof We construct χ as in the proof of Theorem 2 witnessing the isomorphism F with the additional properties on the syntactic part. Then (iso:1)–(iso:6) can be established as in Theorem 2.

In the following we show the additional (iso:7) arguing in $\mathbf{CT}^i \cup \mathbf{CT}^j \cup \mathbf{PA}$ by formal induction on the logical complexity of sentences that if $\mathsf{T}^i(x)$ then $\mathsf{T}^j(f(x))$. The induction argument is standard so we only sketch some cases.

(i) If $\mathsf{lc}(x) = 0$, then $\neg\mathsf{sent}^i(x)$ and by (‡) $\neg\mathsf{sent}^j(f(x))$. So, by ($\mathsf{CT6}^i$) (respectively ($\mathsf{CT6}^j$)) there is no equivalence to show.

(ii) If $\mathsf{sent}^i(x)$ and $\mathsf{lc}(x) = 1$, then if x is $\dot{\mathsf{N}}^i(y)$ for some closed term s. Then

$$\begin{aligned} \mathsf{T}^i(\dot{\mathsf{N}}^i(s)) &\Leftrightarrow \mathsf{N}^i(\mathsf{val}^i(s)) && \text{by } (\mathsf{CT2}^i) \\ &\Leftrightarrow \mathsf{N}^j(\mathsf{val}^j(f(s))) && \text{by (‡) and (7)} \end{aligned}$$

[33] A possibly more elaborate version would employ some disentangled setting as in [18]. In contrast to that setting, in which bridging principles (so-called 'coding axioms') are postulated, in [21] a categoricity argument to establish a form of intolerance in a second-order setting is provided, in which the bridging principles are derivable.

$$\Leftrightarrow \mathsf{T}^j(\underset{\cdot}{\mathsf{N}}^j(f(s))) \qquad \text{by } (\mathsf{CT2}^j).$$

(iii) If $\mathsf{sent}^i(x)$ and x is $\underset{\cdot}{\forall} vy$ for some $\mathsf{var}(v)$ with $\mathsf{lc}(x) = \mathsf{lc}(y) + 1$, then

$$\begin{aligned} \mathsf{T}^i(\underset{\cdot}{\forall} vy) &\Leftrightarrow \forall z(\mathsf{T}^i(\mathsf{sub}(y, v, \mathsf{num}^i(z)))) && \text{by } (\mathsf{CT5}^i) \\ &\Leftrightarrow \forall z(\mathsf{T}^j(f(\mathsf{sub}(y, v, \mathsf{num}^i(z))))) && \text{by IH and (10)} \\ &\Leftrightarrow \mathsf{T}^j(f(\underset{\cdot}{\forall} vy)) && \text{by } (\mathsf{CT5}^j) \text{ and properties of } f \end{aligned}$$

□

4.3 A Unified Theory of Truth

In [4, Chapter 10.5, p. 232] it is suggested that an important consequence of the intolerance theorem is that it allows us to introduce a canonical theory of arithmetic: the idea is that the equivalence provided by the intolerance theorem is a good reason to abandon the (relevant) indices. Our Proposition 1, especially our (iso:7), provides similar reasons for accepting that the relativisation of the two truth predicates is superfluous and that one can introduce a canonical theory of truth. In our setting, this means that we remove the indices and work with a unified theory of truth, i.e. a single truth predicate that works for both languages.

In order to enable a better readability of the following we introduce some additional conventions. First of all we will adapt our set of primitive syntactic vocabulary for the language $\mathcal{L}_A^{ij}$ so that we have sent^{ij}, ct^{ij}. Instead of employing the relativised quantifiers $\forall x(\mathsf{sent}^{ij}(x) \rightarrow ...)$, we are now going to use the more intuitive shorthand $\forall \ulcorner \varphi \urcorner$ to enhance readability, similarly for the existential quantifier.

Moreover, we introduce a convention for the use of variables of different sorts. We use x, y, z as neutral variables ranging over the full domain, l, m, n for variables ranging over N^i and u, v, w for variables ranging over N^j. We also introduce an abbreviation for the substitution function. If $\varphi(v_1, ..., v_h)$ is a formula with at most h free variables, then $\ulcorner \varphi(\dot{m}) \urcorner$ is short for $\mathsf{sub}(\ulcorner \varphi \urcorner, \ulcorner v_1 \urcorner, \mathsf{num}^i(m))$. This generalizes to simultaneous substitution in the expected way. The num^j and num cases are analogous. Moreover, we also allow for mixed substitutions: for instance, $\ulcorner \varphi(\dot{m}, \dot{u}) \urcorner$ is short for the result of substituting the mth i-numeral for m and the uth j-numeral for u in φ. Finally, we will also employ the standard pairing function $(\cdot, \cdot)$ in our distinguished **PA**. Given the overarching theory **PA**, this primitive recursive pairing function is definable and allows us to pair syntactic objects of both sorts.

The truth theory $\mathbf{CT}[\mathbf{PA}^{ij}] \cup \mathbf{PA}$ is formulated in the language $\mathcal{L}_\mathsf{T}^{ij}$. Similarly to the previous section, the language $\mathcal{L}_\mathsf{T}^{ij}$ is build up on the unified signature SIG_T, which is the union of the following signatures: $\mathrm{SIG}_A \cup \mathrm{SIG}_A^i \cup \mathrm{SIG}_A^j \cup \mathrm{SYN} \cup \{\mathsf{T}\}$. In contrast to the previous section, here the arithmetical language is expanded with a single truth predicate T, whose intended range is now the set of sentences from the mixed arithmetical language $\mathcal{L}_A^{ij}$ based on the union of the signatures SIG_A^i,SIG_A^j, SIG_A and SYN.

In the following, we will show that this unified theory of truth proves the internal isomorphism as an existential claim and thereby overcomes one of the *prima facie* disadvantages of the first-order strategy.

The theory $\mathbf{CT}[\mathbf{PA}^{ij}] \cup \mathbf{PA}$ is the extension of the theory $\mathbf{PA}^i \cup \mathbf{PA}^j \cup \mathbf{PA}$ by the universal closures of the following axioms:

$$\text{(CT1)} \quad \mathsf{ct}^k(x) \wedge \mathsf{ct}^k(y) \rightarrow (\mathsf{T}(x \underset{\cdot}{=} y) \leftrightarrow \mathsf{val}^k(x) = \mathsf{val}^k(y))$$

$$\text{(CT2)} \quad \mathsf{sent}^k(\underset{\cdot}{\mathsf{N}}^k(x)) \rightarrow (\mathsf{T}\,\underset{\cdot}{\mathsf{N}}^k(x) \leftrightarrow \mathsf{N}^k(\mathsf{val}^k(x)))$$

$$\text{(CT3)} \quad \mathsf{T}\ulcorner \underset{\cdot}{\neg}\varphi \urcorner \leftrightarrow \neg \mathsf{T}\ulcorner \varphi \urcorner$$

$$\text{(CT4)} \quad \mathsf{T}\ulcorner \varphi \underset{\cdot}{\wedge} \psi \urcorner \leftrightarrow \mathsf{T}\ulcorner \varphi \urcorner \wedge \mathsf{T}\ulcorner \psi \urcorner$$

$$\text{(CT5}^i\text{)} \quad \mathsf{T}\ulcorner \underset{\cdot}{\forall} v(\underset{\cdot}{\mathsf{N}}^i(v) \underset{\cdot}{\rightarrow} \varphi)\urcorner \leftrightarrow \forall y(\mathsf{N}^i(y) \rightarrow \mathsf{T}\ulcorner \mathsf{sub}(\varphi, v, \mathsf{num}^i(y))\urcorner)$$

$$\text{(CT5}^j\text{)} \quad \mathsf{T}\ulcorner \underset{\cdot}{\forall} v(\underset{\cdot}{\mathsf{N}}^j(v) \underset{\cdot}{\rightarrow} \varphi)\urcorner \leftrightarrow \forall y(\mathsf{N}^j(y) \rightarrow \mathsf{T}\ulcorner \mathsf{sub}(\varphi, v, \mathsf{num}^j(y))\urcorner)$$

In the following, we show that the closure of Δ_1-formulas under primitive recursion can be adapted to the present setting. Recursive functions are definable by Δ_1-formulas. The primitive recursive functions are provably recursive. With our internal semantic reflection we can make this explicit as follows: First of all we have Δ_1 characterizations of the classes of formulas such as Σ_1-formulas, i.e. $\Sigma_1(\ulcorner\varphi\urcorner)$. Second we can characterise for a formula φ to represent a total function $F : \mathsf{N}^i \rightarrow \mathsf{N}^j$ by $\mathsf{func}_{i \rhd j}(\ulcorner\varphi\urcorner)$ iff $\forall n \exists! u \mathsf{T}(\ulcorner\varphi(\dot{n}, \dot{u})\urcorner)$. For total functions represented by a Σ_1-formula we can also represent its complement by a Σ_1-formula. The class of Δ_1-total functions from N^i to N^j is then defined as $\Delta_1^{\mathsf{N}}(\ulcorner\varphi\urcorner)$ iff $\Sigma_1(\ulcorner\varphi\urcorner) \wedge \mathsf{func}_{i \rhd j}(\ulcorner\varphi\urcorner)$. It is then possible to lift the closure of Δ_1 under primitive recursion in the following form:

Lemma 1

$$\begin{aligned} \mathbf{CT}[\mathbf{PA}^{ij}] \cup \mathbf{PA} \vdash \Delta_1^{\mathsf{N}}(\ulcorner\psi_{\mathsf{S}}\urcorner) \rightarrow \exists \ulcorner\psi\urcorner(&\Delta_1^{\mathsf{N}}(\ulcorner\psi\urcorner) \wedge \\ &\forall x(\mathsf{T}\ulcorner\psi(\overline{\#0}, \dot{x})\urcorner \leftrightarrow x = (\overline{0}^i, \overline{0}^j)) \wedge \\ &\forall z(\mathsf{T}\ulcorner\psi(\underset{\cdot}{\mathsf{S}}(\dot{z}), \dot{x})\urcorner \leftrightarrow \exists r(\mathsf{T}\ulcorner\psi(\dot{z}, \dot{r})\urcorner \wedge \mathsf{T}\ulcorner\psi_{\mathsf{S}}(\dot{r}, \dot{x})\urcorner))) \end{aligned}$$

Proof (Sketch) The idea is basically similar to the construction of our formula ψ in Theorem 2. Given a primitive recursive function F, that maps an ordered pair (n, u) to $(S^i(n), S^j(u))$, for the successor case represented by ψ_{S}, there is a primitive recursive function G represented by ψ, such that $G(0)$ is the ordered pair $(0^i, 0^j)$ and $G(x+1) = F(G(x))$. The semantic ascent and the truth predicate allow us to state this in the more general form.[34] □

Then $\mathsf{ISO}_{i \cong j}(\ulcorner\chi\urcorner)$ is the conjunction of the following

$$\forall y \forall z(\mathsf{T}\ulcorner\chi(\dot{y}, \dot{z})\urcorner \rightarrow (\mathsf{N}^i(y) \wedge \mathsf{N}^j(z))) \quad \text{(iso:1)}$$

$$\forall n \exists! u\, \mathsf{T}\ulcorner\chi(\dot{n}, \dot{u})\urcorner \quad \text{(iso:2)}$$

$$\forall v \exists! m\, \mathsf{T}\ulcorner\chi(\dot{v}, \dot{m})\urcorner \quad \text{(iso:3)}$$

[34] Compare [13], Lemma 1.79. They show that this lemma is also provable in $\mathbf{I\Sigma_1}$ with a truth predicate for Σ_1 statements definable in $\mathbf{I\Sigma_1}$.

$$\mathsf{T}\ulcorner\chi(\overline{0^i}, \overline{0^j})\urcorner \land \forall n, u(\mathsf{T}\ulcorner\chi(\dot{n}, \dot{u})\urcorner \to \mathsf{T}\ulcorner\chi(\underset{\cdot}{\mathsf{S}}^{i}(\dot{n}), \underset{\cdot}{\mathsf{S}}^{j}(\dot{u}))\urcorner) \tag{iso:4}$$

$$\forall n, m\forall u, v, w(\mathsf{T}\ulcorner\chi(\dot{n}, \dot{u})\urcorner \land \mathsf{T}\ulcorner\chi(\dot{m}, \dot{v})\urcorner \land \mathsf{T}\ulcorner\chi(\dot{n} \,\underset{\cdot}{+}^{i}\, \dot{m}, \dot{w})\urcorner \to w = u +^{j} v) \tag{iso:5}$$

$$\forall n, m\forall u, v, w(\mathsf{T}\ulcorner\chi(\dot{n}, \dot{u})\urcorner \land \mathsf{T}\ulcorner\chi(\dot{m}, \dot{v})\urcorner \land \mathsf{T}\ulcorner\chi(\dot{n} \,\underset{\cdot}{\times}^{i}\, \dot{m}, \dot{w})\urcorner \to w = u \times^{j} v) \tag{iso:6}$$

The internal semantic reflection provided by the truth predicate allows us to lift the usual course-of-values recursion for the definition of χ to the construction of the term $\ulcorner\chi\urcorner$ witnessing the existential claim for the isomorphism. Lemma 1 allows us to prove the following version of categoricity:[35]

Proposition 2

$$\mathbf{CT}[\mathbf{PA}^{ij}] \cup \mathbf{PA} \vdash \exists\ulcorner\chi\urcorner \mathsf{ISO}_{i\cong j}(\ulcorner\chi\urcorner).$$

4.4 General Categoricity with Truth

In the previous sections, we considered two arithmetical theories with additional truth predicates, and employed syntactical 'bridging' principles. Although this approach naturally applies to the setting of communication as intended by Parsons, the assumptions (that we made) of an overarching theory of syntax could be challenged.[36]

There are at least two points of contention for our approach: Firstly, our bridging principles are formulated in an overarching theory of syntax: this amounts to the assumption that the two 'domains' are subsets of the overarching 'domain'. Secondly, the symmetric treatment of $\mathbf{PA}^i$ and $\mathbf{PA}^j$ presupposes that the two mathematicians already share a more general framework.

The following section aims to overcome these issues by adapting our framework to an asymmetric treatment. In contrast to the symmetric approach, which aims to provide internal categoricity between two given internal structures using a distinguished, overarching theory, the asymmetric approach aims to prove internal categoricity between our distinguished internal structures and 'alternative' ones.

Such an asymmetric treatment has already been used in the reverse mathematics investigation of categoricity in [27]. The reverse mathematics result establishes that $\mathbf{WKL}_0$ is sufficient to establish an internal categoricity result.[37] However, the setting employed by the reverse mathematics case has an implicit assumption: the alternative domain is assumed to be contained in the distinguished domain of natural numbers. For its application to Parsons' strategy, this assumption faces similar challenges to our first contention, as the assumption of an overarching theory.[38]

35 As one of the referees pointed out, it seems possible that for a proof of Proposition 2 and Lemma 1 in isolation, a (suitably modified) partial Σ_1-truth predicate is sufficient, following [13]. However, in the context of the previous result, Proposition 1, it appears natural to use a primitive truth predicate.

36 [32, Chapter 10.3] for example expresses scepticism towards the philosophical significance of what he calls the intra-language versions of the categoricity argument. His scepticism is motivated by the worry that the assumption of an overarching theory might be too strong.

37 For more details about $\mathbf{WKL}_0$ see [26, p. 35].

38 [4, p. 246] also indicate the importance of this implicit assumption. We should point out that we do not criticise the mathematical significance of the result. Here we are only interested in the philosophical application of this result for Parsons' strategy.

Similarly to the reverse mathematics case, we will use a primitive set of notions only for one distinguished internal structure, breaking the symmetry with the previous cases, where we worked with two given arithmetical structures and two sets of primitive vocabularies.

Under these assumptions, we will provide an additional version of internal categoricity (Proposition 3) and of intolerance (Corollary 2). Our internal categoricity result is going to be based on the categoricity result due to [11]. For the bridging principles, they use a theory of finite sets and classes called **EFSC**.[39] In our approach, we replace the finite set theory, that is responsible for the bridging principles, with our expanded theory of arithmetic and truth.[40] In the following, we make the assumptions on the expanded theory explicit.

We work within a language $\mathcal{L}_{\mathsf{N}}$ of first-order **PA**, with a predicate N for 'our' natural numbers. The vocabulary $\mathcal{L}_{\mathsf{N}}^{\mathsf{T}}$ expands $\mathcal{L}_{\mathsf{N}}$ and is based on the signature $\mathrm{SIG}_{\mathsf{N}}^{\mathsf{T}} := \mathrm{SIG}_A \cup \{\mathsf{N}\} \cup \{\mathsf{T}\}$. The arithmetical axioms are formulated relativised to this N and the theory of syntax is standard for our language.

Whereas we use our mathematical vocabulary for our usual mathematical discourse, to talk about *our* natural numbers, we interpret the 'alternative' internal arithmetical structures by language expansions, mainly due to an arbitrary new predicate P, intended to represent the alternative domain. We assume that P and N are disjoint, so specifically we do not assume that the alternative domain is contained in our distinguished domain.

In order to have suitable bridging principles we expand our linguistic repertoire by allowing for additional syntactic vocabulary, for example a primitive notion of ordered pair. The language $\mathcal{L}_{\mathsf{N}}^{\mathsf{T}}(P)$ is then based on the signature $\mathrm{SIG}_{\mathsf{N}}^{\mathsf{T}}(P) := \mathrm{SIG}_{\mathsf{N}}^{\mathsf{T}} \cup \{P\} \cup \mathrm{SYN}^{+}$. We again allow for an expansion of predicates, constants and function symbols. SYN^{+} contains additional resources (relative to P). We use $n, m, ...$ for variables ranging over N, and $v, w, ...$ for variables ranging over P, and $x, y, ...$ for unrestricted quantification.

In the following we explain the additional primitives in SYN^{+}: we have primitive unary predicates $\mathsf{pred}^{P}(x)$, $\mathsf{const}^{P}(x)$, $\mathsf{func}^{P}(x)$, with the intended interpretation 'x is a predicate symbol, whose range is in P', 'x is a constant with value in P', and 'x is a function symbol representing a function from P to P'. The intended range of these additional syntactic objects is the alternative arithmetical structure. Finally, we add a primitive ternary function symbol $\mathsf{appl}(n, x, y)$ to SYN^{+} with the intended reading 'the result of applying n-times the syntactical operation x to the syntactical object y'. In the case that $n = 1$ we also use the abbreviation $\mathsf{appl}(x, y)$. The idea behind the application function is to allow an enumeration with our numbers of the applications of syntactical operations in the expanded language.[41] Our syntactic repertoire includes also predicates representing the syntactic categories of the expanded language. So

[39] For the presentation of **EFSC**, see [11, p. 3-4] Although it is interpretable in $\mathbf{ACA_0}$, the interpreted version is not as interesting as it could be, since by interpreting all objects as elements of $\mathbb{N}$ also the internal models will be subsets of $\mathbb{N}$. This is a quite similar assumption as in the reverse version that we gave up.

[40] It is well known that there is a close connection between adjunctive set theory and concatenation theory, see for example [7].

[41] The introduction of this application function symbol can be challenged. This will be discussed in Section 5.

there is a formula $\mathsf{ct}^P(x)$ that represents the closure of the set of individual constants under function application, and similarly $\mathsf{term}^P(x)$, $\mathsf{form}^P(x)$ and $\mathsf{sent}^P(x)$.

4.4.1 Peano Systems and CT[P]

We intend to characterise arbitrary arithmetical interpretations, which we call Peano systems.[42] More specifically, we will relativise these system to our parameter P and therefore say that a P-Peano system (a Peano system relative to P) is a triple (p, a, h), such that $\mathsf{pred}^P(p)$, $\mathsf{const}^P(a)$ and $\mathsf{func}^P(h)$. The first component is intended to be a sortal predicate for the range of the quantifiers of the Peano system, the second component is a constant denoting the zero-element of the Peano system, and the third component is a one-place function symbol, representing the successor function of the Peano system. In order to have a suitable form of quantification over terms we assume that the constant and the function symbol form a systematic naming device for an alternative internal structure. In the following we list the linguistic part of our assumptions on a P-Peano system:

$$\mathsf{const}^P(x) \to \mathsf{term}^P(x) \tag{L_1}$$

$$\mathsf{var}(x) \to \mathsf{term}^P(x) \tag{L_2}$$

$$\mathsf{term}^P(x) \wedge \mathsf{func}^P(y) \to \forall n(\mathsf{term}^P(\mathsf{appl}(n, x, y))). \tag{L_3}$$

Additionally, we expand the notion of a P-formula, form^P, for formulas only containing notions of the expanded language and closed under application of conjunction, negation and N-restricted quantification and P-restricted quantification.

$$\mathsf{term}^P(x) \wedge \mathsf{term}^P(y) \to \mathsf{form}^P(\mathsf{appl}(\dot{=}, (x, y))) \tag{L_4}$$

$$\mathsf{term}^P(x) \wedge \mathsf{pred}^P(p) \to \mathsf{form}^P(\mathsf{appl}(p, x)) \tag{L_5}$$

The usual closure conditions for $\neg, \wedge, \forall$ are labelled as (L_6) - (L_8). In order to be able to formulate our theory of truth, we intend to expand the notion of a numeral and the value function to P. For this, we assume that all the systems that we consider have a distinguished constant and a distinguished function symbol.

Then we expand our num function to $\mathsf{num}^P(x)$. The idea is that the distinguished constant plays the role of the P-constant symbol (for zero) in the P-Peano system. The 'zeroth' P-numeral, $\overline{0}^P$, is then this distinguished constant. Moreover, the distinguished function symbol plays the role of the P-successor function symbol in the P-Peano system. The P-numeral $\overline{n}^P$ is then the nth application of the distinguished P-successor function symbol to the distinguished P-constant. Since for the enumeration of applications of the P-successor function we use our numbers, the function is primitive recursive and is represented by $\mathsf{num}^P(x)$.

42 This terminology is used in [27].

For the generalization of our term valuation function $\mathsf{val}(x)$ we assume that every P-constant has a unique value in P and that this is preserved for P-function applications on P-closed terms:

$$\forall x(\mathsf{const}^P(x) \rightarrow \exists! u(\mathsf{val}(x) = u)) \qquad (L_9)$$

$$\forall x, y(\mathsf{func}^P(x) \wedge \mathsf{ct}^P(y) \rightarrow \exists! u(\mathsf{val}(\mathsf{appl}(x, y)) = u)) \qquad (L_{10})$$

We can then use the P-numerals to define the value function by recursion. The value of $\overline{0}^P$ is the unique value of our distinguished constant. The value of $\overline{n+1}^P$ is the unique value of the distinguished function symbol applied to $\overline{n}^P$. Additionally, we assume that for all the objects from the alternative domain there is a P-numeral that has the number as its value[43]

$$\forall u \exists n(u = \mathsf{val}(\mathsf{appl}(n, h, a))) \qquad (L_{11})$$

The uniqueness is guaranteed by the injectivity of the successor function. Then we can expand the dot notation $\dot{x}$ in a natural way. If $P(u)$, then $\dot{u}$ is the unique P-numeral t, such that $\mathsf{val}(t) = u$.

Peano systems are usually characterised as triples, consisting of a set, an element of this set and a function on this set, satisfying three properties: The element is not within the range of the function, the function is injective and induction holds for all subsets.[44] In our setting, the first two properties for Peano systems are stated as follows as (L_{12}):

$$\exists x \exists y\, (\mathsf{const}^P(x) \wedge \mathsf{func}^P(y) \wedge \forall n \geq 1(\mathsf{val}(\mathsf{appl}(n, y, x)) \neq \mathsf{val}(x)) \wedge$$
$$\forall z, z'(\mathsf{ct}^P(z) \wedge \mathsf{ct}^P(z') \rightarrow (\mathsf{val}(\mathsf{appl}(y, z)) = \mathsf{val}(\mathsf{appl}(y, z')) \rightarrow \mathsf{val}(z) = \mathsf{val}(z'))))$$

With this we formulate the system $\mathbf{CT}[P]$, in the language $\mathcal{L}_{\mathsf{N}}^{\mathsf{T}}(P)$. It extends first-order Peano arithmetic with the universal closures of the following axioms, relativised to P:

$$\mathsf{ct}(x) \wedge \mathsf{ct}(y) \rightarrow (\mathsf{T}(x \mathbin{\underset{\cdot}{=}} y) \leftrightarrow \mathsf{val}(x) = \mathsf{val}(y)) \qquad (\mathsf{CT1}^P)$$

$$\mathsf{ct}^P(x) \rightarrow (\mathsf{T}(\mathsf{appl}(p, x)) \leftrightarrow P(\mathsf{val}(x))) \qquad (\mathsf{CT2}^P)$$

$$\mathsf{T}\ulcorner \underset{\cdot}{\neg} \varphi \urcorner \leftrightarrow \neg \mathsf{T}\ulcorner \varphi \urcorner \qquad (\mathsf{CT3}^P)$$

$$\mathsf{T}\ulcorner \varphi \mathbin{\underset{\cdot}{\wedge}} \psi \urcorner \leftrightarrow \mathsf{T}\ulcorner \varphi \urcorner \wedge \mathsf{T}\ulcorner \psi \urcorner \qquad (\mathsf{CT4}^P)$$

$$\forall n \mathsf{T}\ulcorner \varphi(\dot{n}) \urcorner \leftrightarrow \mathsf{T}\ulcorner \underset{\cdot}{\forall} x(\underset{\cdot}{\mathsf{N}}(x) \mathbin{\underset{\cdot}{\rightarrow}} \varphi(x)) \urcorner \qquad (\mathsf{CT5}^P)$$

$$\forall u \mathsf{T}\ulcorner \varphi(\dot{u}) \urcorner \leftrightarrow \mathsf{T}\ulcorner \underset{\cdot}{\forall} x(\mathsf{appl}(p, x) \mathbin{\underset{\cdot}{\rightarrow}} \varphi(x)) \urcorner \qquad (\mathsf{CT6}^P)$$

[43] Similarly to the case in [27] we require an additional assumption to internally state that we can 'reach' all elements of P. An alternative strategy would be to assume a new primitive num: function for P.

[44] For a presentation of Peano systems in the reverse mathematics case, see [27, Definition 2.1., p. 285]

A pre-Peano system (p, a, h) is a triple, such that p is a predicate symbol, a is a distinguished constant and h is a distinguished function symbol, such that $\bigwedge_{1\leq i\leq 12} L_i$ holds of these symbols. With the truth predicate and $\mathbf{CT}[P]$ in the background we can now formulate the third property for a P-Peano system:

$$\forall\ulcorner\varphi\urcorner(\mathsf{T}\ulcorner\varphi(a)\urcorner \wedge \forall u(\mathsf{T}\ulcorner\varphi(\dot{u})\urcorner \rightarrow \mathsf{T}\ulcorner\varphi(\mathsf{appl}(h, \dot{u}))\urcorner) \rightarrow \mathsf{T}\ulcorner\underset{\cdot}{\forall} x(\mathsf{appl}(p, x) \underset{\cdot}{\rightarrow} \varphi(x))\urcorner) \tag{L_{13}}$$

A triple (p, a, h) is a P-Peano system, denoted as $\mathsf{PS}_P(p, a, h)$, just in case (p, a, h) is a pre-Peano system satisfying the additional property L_{13}.

4.4.2 The Isomorphism

The aim of this section is to provide the proof of the wanted internal isomorphism between our internal arithmetical structure and the Peano system relativised to P.

With our naming machinery we can build singleton sets of elements of either N or P using our num-function, for example with L_{13} we get

$$\forall u\exists\ulcorner\varphi\urcorner\forall y(\mathsf{T}\ulcorner\varphi(\dot{y})\urcorner \leftrightarrow y = u).$$

Important for our case is that we can also simulate ordered pairs of elements of N and P simultaneously.

$$\forall n\forall u\exists\ulcorner\varphi\urcorner\forall x, y(\mathsf{T}\ulcorner\varphi(\dot{x}, \dot{y})\urcorner \leftrightarrow x = n \wedge y = u)$$

This is due to our naming machinery and the fact that our arithmetical syntax theory and truth theory work for the wider range. Additionally, we can use disjunctions to mimic the talk about 'finite' subsets by adjunction. With this in place we can carry out a proof in $\mathbf{CT}[P]$ by following Feferman and Hellman's strategy.

The basic task is to define a function from our natural numbers into the domain of P, s.t. the following holds

$$\begin{aligned} f(0) &= \mathsf{val}(a) \\ f(\mathsf{S}(n)) &= \mathsf{val}(\mathsf{appl}(h, f(n))) \end{aligned} \tag{$**$}$$

The proof strategy is to approximate this function by using suitable formulas for 'finite' sets of elements of both 'domains'. We let $\mathsf{rec}(\ulcorner\varphi\urcorner, h, a, n)$ be the conjunction of (R_1) - (R_3) with

$$\forall m\forall w(\mathsf{T}\ulcorner\varphi(\dot{m}, \dot{w})\urcorner \rightarrow m \leq n) \tag{R_1}$$

$$\forall w(\mathsf{T}\ulcorner\varphi(\overline{\#0}, \dot{w})\urcorner \leftrightarrow w = \mathsf{val}(a)) \tag{R_2}$$

$$\forall m < n\forall w(\mathsf{T}\ulcorner\varphi(\underset{\cdot}{\mathsf{S}}\dot{m}, \dot{w})\urcorner \leftrightarrow \exists u(\mathsf{T}\ulcorner\varphi(\dot{m}, \dot{u})\urcorner \wedge \mathsf{T}\ulcorner\mathsf{appl}(h, \dot{u}) \underset{\cdot}{=} \dot{w}\urcorner)) \tag{R_3}$$

Since our compositional theory works for the expanded language we can as usual use the commutation of truth with the connectives to show the following:

Lemma 2 $\exists\ulcorner\psi\urcorner\forall n(\mathsf{T}\ulcorner\psi(\ulcorner\varphi\urcorner, \dot{h}, \dot{a}, \dot{n})\urcorner \leftrightarrow \mathsf{rec}(\ulcorner\varphi\urcorner, h, a, n))$

It is possible to establish standard properties (compare Theorem 3. in [11, p. 8]) of the predicate rec, such that the following holds:

$$\forall n\exists!u\exists\ulcorner\varphi\urcorner(\mathsf{rec}(\ulcorner\varphi\urcorner, h, a, n) \wedge \mathsf{T}\ulcorner\varphi(\dot{n}, \dot{u})\urcorner) \tag{!}$$

Then we define our function $f : \mathsf{N} \to P$ satisfying (∗) by the following formula

$$f(n, u) :\leftrightarrow \exists\ulcorner\varphi\urcorner(\mathsf{rec}(\ulcorner\varphi\urcorner, h, a, n) \wedge \mathsf{T}\ulcorner\varphi(\dot{n}, \dot{u})\urcorner)$$

Now we can establish the internal categoricity for the relevant notion of isomorphism. We let $\mathsf{ISO}_{\mathsf{N}\cong P}(f, (p, h, a))$ be the conjuction of the following:

$$\forall n\forall v(f(n, v) \to (\mathsf{N}(n) \wedge P(v))) \tag{iso:1}$$

$$\forall n\exists!u\ f(n, u) \tag{iso:2}$$

$$\forall v\exists!m\ f(m, v) \tag{iso:3}$$

$$f(0, \mathsf{val}(a)) \wedge \forall n, u(f(n, u) \to f(\mathsf{S}(n), \mathsf{val}(\mathsf{appl}(h, \dot{u})))) \tag{iso:4}$$

The isomorphism is established in a similar fashion in [11, Theorem 5.].

Proposition 3

$$\mathbf{CT}[P] \vdash \mathsf{PS}^P(p, h, a) \to \mathsf{ISO}_{\mathsf{N}\cong P}(f, (p, h, a)))$$

Proof We have (iso:1) by definition of rec. By (!) we have (iso:2) and therefore in order to simplify the presentation we work in the following with f as a function symbol. By (R_3) we also have (iso:4). What remains to be shown is that f is one-to-one and onto, (iso:3). For the former we use induction on n in N to show $f(n) = f(m) \to n = m$.

To show surjectivity we employ induction on P as stated in (L_{13}). The idea is to use it on the subset X of P, such that $X = \{u \in P \mid \exists n f(n) = u\}$. In order to apply (L_{13}) we have to make sure that there is a formula θ, such that $\mathsf{T}\ulcorner\theta(\dot{v})\urcorner$ defines X. We have $v \in X$ iff $P(v) \wedge \exists n(f(n) = v)$, which is by definition of f equivalent to $P(v) \wedge \exists n\exists\ulcorner\varphi\urcorner(\mathsf{rec}(\ulcorner\varphi\urcorner, h, a, n) \wedge \mathsf{T}\ulcorner\varphi(\dot{n}, \dot{v})\urcorner)$. Then we can use the $\ulcorner\psi\urcorner$ from Lemma 2 to reformulate it as $P(v) \wedge \exists n\exists\ulcorner\varphi\urcorner(\mathsf{T}\ulcorner\psi(\ulcorner\varphi\urcorner, \dot{h}, \dot{a}, \dot{n})\urcorner \wedge \mathsf{T}\ulcorner\varphi(\dot{n}, \dot{v})\urcorner))$. Then we can use the expanded T-biconditionals for P and the commutation axioms to see that there is a formula θ, such that $\mathsf{T}\ulcorner\theta(\dot{v})\urcorner$ is extensionally equivalent to the previous (with the parameters h, a hidden).

To apply (L_{13}) we show: $\mathsf{T}\ulcorner\theta(a)\urcorner \wedge \forall w(\mathsf{T}\ulcorner\theta(\dot{w})\urcorner \to \mathsf{T}\ulcorner\theta(\mathsf{appl}(h, \dot{w}))\urcorner)$. The a case is obvious. In the successor case we assume that for some w we have $\mathsf{T}\ulcorner\theta(\dot{w})\urcorner$. By definition of f, $f(\mathsf{S}(n)) = \mathsf{val}(\mathsf{appl}(h, f(n)))$ and so $\mathsf{T}\ulcorner\theta(\mathsf{appl}(h, \dot{w}))\urcorner$. Then by (L_{13}) we have $\forall v\mathsf{T}\ulcorner\theta(\dot{v})\urcorner$ establishing the surjectivity of f. □

5 Determinacy Reconsidered

The previous section presented our truth-theoretic versions of internal categoricity based on a primitive notion of truth. It is time to reconsider our original goal and discuss how and to what degree our approach improves on the first-order, Parsons-style approach. As we discussed in Sections 2 and 3, the drawbacks of the first-order, Parsons-style, approach were mainly the following three: as Button and Walsh noted, first-order approaches seem not to be able to provide a crisp version of the intolerance theorem, meaning that there is no immediate and direct formulation of a first-order intolerance theorem. Additionally, the first-order version of internal categoricity seems to be less general than its second-order counterpart. Finally, in the first-order version, the existence of the wanted isomorphism was not proved explicitly.

Let us start with the complaint that the first-order version lacks a direct and explicit formulation of an analogue of intolerance. In our truth theoretic case of Section 4.2 it was possible to establish a truth-theoretic equivalence between arithmetical sentences:

$$\forall x \forall y(\mathsf{sent}^i(x) \wedge \mathsf{sent}^j(y) \wedge \chi(x, y) \rightarrow (\mathsf{T}^i(x) \leftrightarrow \mathsf{T}^j(y))) \qquad \text{(iso:7)}$$

(iso:7) can be already seen as a first-order formulation of intolerance. Moreover, in the asymmetric case of Section 4.4, one can also show a truth-theoretic equivalence between arithmetical sentences $\ulcorner\varphi\urcorner$ and their relativisations $\ulcorner\varphi^P\urcorner$ to the parameter P:

Corollary 2 *Intolerance*

$$\mathbf{CT}[P] \vdash \mathsf{PS}^P(p, h, a) \rightarrow \forall\ulcorner\varphi\urcorner(\mathsf{sent}^{\mathsf{N}}(\ulcorner\varphi\urcorner) \rightarrow (\mathsf{T}\ulcorner\varphi\urcorner \leftrightarrow \mathsf{T}\ulcorner\varphi^P\urcorner)).$$

These formulations also witness that in the first-order case, intolerance can be formulated quite naturally. Due to the expressive resources provided by the truth predicate, intolerance can be stated explicitly, as a sentence of our language, directly expressing that the truth of arithmetical statements is preserved along different interpretations. Moreover, there is a sense in which these formulations improve even on the second-order formulation. As we pointed out in Section 3, Theorem 3 establishes a local, schematic form of intolerance, for any given arithmetical sentence φ. In contrast to this, our version of intolerance is global: the truth predicate allows us to quantify over all arithmetical sentences $\ulcorner\varphi\urcorner$. For this reason, our versions of intolerance are natural and acceptable from an internalist, Parsons-style perspective.

Turning to the question of generality, we saw that the first-order approach is not as general as its counterpart in pure second-order logic: using the resources of second-order logic one can quantify over all internal structures, and this is not possible in a first-order version. In our truth-theoretic approach, we attempted to overcome some of the expressive limitations. In Proposition 3 and Corollary 2 we allowed for an additional parameter P to recover generality. Although P is taken to be the domain of an arbitrary arithmetical structure, without the means of impredicative second-order quantification we cannot reach the level of generality of the second-order version. Even if full quantification seems to be out of reach within a Parsons-style approach, our truth-theoretic approach – involving language expansions with an arbitrary parameter

P – might be as general as one can expect in a predicative arithmetical setting. This additional first-order constraint also sets us apart from the approach in [11]: although their results are provided by predicative means, they allow for class variables and quantifiers. It is not entirely clear whether their use of class quantifiers is acceptable from a Parsons-style perspective.

Finally, our truth-theoretic approach also improves on the Parsons-style first-order strategy with respect to the existence claim for the isomorphism: in Proposition 2 we are able to show the existence of the isomorphism explicitly. This is possible due to the expressive resources provided by the truth predicate.

We conclude this article by discussing some additional questions and possible objections to our truth-theoretic approach.

On the face of it there seems to be tension between our approach and a result in [15]. They claim that "the definiteness of the theory of truth for a structure does not follow as a consequence of the definiteness of the structure in which that truth resides." [15, p. 26]. So, even if the natural number structure is definite in the sense (exemplified in their theorem) that two **ZFC**-models agree on the interpretation of the arithmetical vocabulary, arithmetical truth remains indeterminate; there is an 'arithmetical' statement σ, such that the two models do not agree and assign two different truth values to σ. This philosophical conclusion is based on their Theorem 1 in [15, p. 5].

From a model-theoretic perspective, the theorem shows that we can expand a **PA**-model $\mathfrak{A}$ with two satisfaction classes S and S' in an incompatible way, i.e., such that $(\mathfrak{A}, S)$ and $(\mathfrak{A}, S')$ are models of **CT** and disagree about the truth of an arbitrary arithmetical statement σ, such that $\sigma \in S$ and $\sigma \notin S'$.

They draw the following philosophical conclusion:

> [T]he definiteness of the theory of truth for a structure does not follow as a consequence of the definiteness of the structure in which that truth resides. Even in the case of arithmetic truth and the standard model of arithmetic $\mathbb{N}$, we claim, it is a *philosophical error* to deduce that arithmetic truth is definite just on the basis that the natural numbers themselves and the natural number structure $\langle \mathbb{N}, +, \times, 0, 1, < \rangle$ is definite. At bottom, our claim is that one does not get definiteness-of-truth *for free* from definiteness-of objects and definiteness-of-structure. [our emphasis] [15, p. 26]

Let us discuss the two emphasised points in the previous passage. The first concerns the claim that it is a philosophically erroneous inference to conclude the determinacy of truth from the definiteness of the structure. The technical result of their Theorem 1 is beyond doubt and supports Hamkins and Yang's philosophical conclusion, at least for an external understanding of determinacy and of structures. However, from an internalist understanding of determinacy and arithmetical interpretations, their technical result does not support an equivalent philosophical conclusion. As our result shows, from an internalist perspective, the acceptance of (the relevant) axiomatic truth-theoretic principles implies determinacy of truth – in the form of internalised intolerance exemplified by Proposition 1. In our case, the scenario sketched by Hamkins and Yang is excluded by our expanded induction to include statements of the mixed vocabulary and the bridging principles. By doing so, we ensure that the two **CT**-models agree on their respective arithmetical truths.[45]

Hamkins' and Yang's picture presupposes expressive resources and distinctions that are not directly available in an internalist conception. For example the relevant counterexample σ is a nonstandard sentence and also the models are nonstandard models of arithmetic and set theory. For the internalist these fine-grained distinctions are not within the range of interpretations that she is able to discriminate. Parsons and others have convincingly argued that such distinctions are not within the reach of internalism.[46] For the internalist the models are not given as something external, existing independently. The internalist can only make sense of these models via descriptions of the model, as given in the object language. However, once the nonstandard model is given by a description, via a suitable language expansion for example, the internalist cannot only recognise the interpretation, but also conceive its inadequacy as a relevant alternative.

The second point concerns the question of whether we obtain the determinateness of truth 'for free'. We admit that the inference is not obtainable without any further presuppositions, and also that the determinacy of truth is not 'for free'. We do assume that a suitable expansion of the range of a truth predicate to an arbitrary P is possible. However, we should add that it seems plausible (and unsurprising) that categoricity and determinacy cannot be obtained without any additional resources or bridging principles, as all versions of categoricity that we discussed employ (implicitly or explicitly) some additional principles to construct the wanted isomorphism.

Given that there is no determinacy 'for free', we believe that the question of the acceptability of the bridging assumptions becomes essential to evaluate Parsons' project. In contrast to the second-order approach, our bridging assumptions are less committing than the full impredicative comprehension principles. Additionally, our bridging principles are not 'hidden' in the logic, but are made explicit. In Sections 4.2 and 4.3, we employ bridging principles in form of an overarching theory of arithmetic together with a syntactical symmetry within the two languages. This amounts to the assumption that the alternative internal structures 'live within' the overarching structure. A similar assumption is implicit in the reverse mathematics case. Of course, these assumptions could (and should) be open to further philosophical scrutiny. In Section 4.4, we tried to weaken these assumptions by allowing the alternative P to be disjoint from N. In order to obtain the wanted categoricity result, we relied on the assumption that we can enumerate the alternative P-numerals with our natural numbers and that the P-numerals suffice for explicating the truth of quantified P-statements. Although these assumptions are not completely innocent, we think that they are more acceptable and less committing than the assumptions in place in

45 We should point out that Hamkins and Yang don't directly argue against internalism. Their argument is employed in the context of Feferman's philosophy of mathematics and within his claim that the definiteness of the natural numbers implies the determinacy of arithmetical truth. For an analysis of Feferman's argument the consequences are dependent on an interpretation of Feferman's conceptual structuralism. It would be an interesting question to investigate whether Feferman's conceptual structuralism can be understood as a form of internalism.

46 Compare Parsons position spelled out in [23, p. 288] but also the discussion in [4, p. 279 f.].

the second-order approach, especially for the communicative setting considered by Parsons.

Although our truth-theoretic approach displays some attractive features, one might worry that the introduction of a truth predicate goes beyond Parsons' first-order constraints. A primitive truth predicate exceeds pure first-order arithmetic, but we think that in an open-ended conception of arithmetic the notion of truth that we employ is close enough to be acceptable as part of an arithmetical setting. This is also witnessed by the close connection between theories of truth and predicative subsystems of second-order arithmetic. Against the worry that **CT** might be unmotivated, we note that in the open-ended setting there is an alternative strategy. It is possible to start with weak disquotational truth-theoretic principles and recover the compositional principles of typed truth via reflection.[47]

A variation of this worry might question the compatibility of internalism and truth. One might object that a truth predicate 'brings back' the external semantic notions rejected by internalism. We believe that these worries are misplaced. Introducing a primitive truth predicate via an axiomatic theory of truth is not only in line with a Davidsonian conception, but also with a deflationary conception of truth. On a deflationary conception the main purpose of the truth predicate is expressive and we follow this understanding. With this we set our approach apart from several other conceptions, such as the model-theoretic 'external', semantic notion of truth-in-a-structure. It is such a conception that internalism rejects. Internalism does not allow for a strong form of semantic reflection from a meta-theoretical perspective along the lines of model-theory. However, it is compatible with an object linguistic (deflationary) truth.

Although a full evaluation of the truth-theoretic approach is not possible at this early stage, this work provides some important first steps towards a more attractive picture of internalism, and additional motivation for further work on a Parsons-style approach to the uniqueness and agreement.

Acknowledgements Versions of this article have been presented at research seminars in Bristol, Konstanz, and Warsaw and a workshop in Helsinki. Thanks to the audience for their helpful comments and discussion. In particular, we are thankful to Philip Welch, Sean Walsh, Cezary Cieśliński, and the two anonymous referees of this journal.
Martin Fischer thankfully acknowledges the financial support by the DFG, "Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) - Projektnummer 407312485". Matteo Zicchetti is grateful to the South, West and Wales Doctoral Training Partnership and the Arts and Humanities Research Council for their financial support (AH/L503939/1)

Author Contributions All authors whose names appear on the submission (i) made substantial contributions to the conception of the work; (ii) drafted the work and revised it critically for important intellectual content; (iii) approved the version to be published; (iv) agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding Martin Fischer: Deutsche Forschungsgemeinschaft (407312485). Matteo Zicchetti Arts and Humanities Research Council (AH/L503939/1)

Data Availability Not applicable.

[47] An overview of several reflection strategies, including principles of type-free truth, is given in [16].

Declarations

Ethical Approval Not applicable.

Conflict of interest No competing interests.

References

1. Buss, S. R. (Ed.) (1998). *Handbook of Proof Theory*. Elsevier Science Publisher.
2. Button, T. (2022). Mathematical internal realism. In J. Conant, & S. Chakraborty (Eds.), *Engaging Putnam* (pp. 157–182). De Gruyter.
3. Button, T., & Walsh, S. (2016). Structure and Categoricity: Determinacy of reference and truth value in the philosophy of mathematics. *Philosophia Mathematica, 24*, 283–307.
4. Button, T., & Walsh, S. (2018). *Philosophy and Model Theory*. Oxford University Press.
5. Cantini, A. (1989). Notes on formal theories of truth. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 35*, 97–130.
6. Carnap, R. (1950). Empiricism, semantics, and ontology. *Revue Intern de Phil, 4*, 20–40.
7. Damnjanovic, Z. (2017). Mutual interpretability of Robinson arithmetic and adjunctive set theory with extensionality. *The Bulletin of Symbolic Logic, 23*, 381–404.
8. Detlefsen, M., & Arana, A. (2011). Purity of Methods. Philosophers'. *Imprint, 11*, 1–20.
9. Dummett, M. (1963). The philosophical significance of Gödel's theorem. In M. Dummett (Ed.), *Ratio* (pp. 186–214). Duckworth.
10. Feferman, S. (2013). Categoricity and open-ended axiom systems. slides (online).
11. Feferman, S., & Hellman, G. (1995). Predicative foundations of arithmetic. *Journal of philosophical logic, 24*, 1–17.
12. Field, H. (2007). *Truth and the Absence of Fact*. Oxford University Press.
13. Hájek, P., & Pudlák, P. (1998). *Metamathematics of First-Order Arithmetic*. Springer. 2nd ed.
14. Halbach, V. (2014). *Axiomatic Theories of Truth*. (revised ed.). Cambridge University Press.
15. Hamkins, J. D., & Yang, R. (2013). Satisfaction is not absolute. p. arXiv:1312.0670.
16. Horsten, L., & Leigh, G. E. (2017). Truth is simple. *Mind, 126*, 195–232.
17. Lavine, S. (1994). *Understanding the Infinite*. Harvard University Press.
18. Leigh, G. E., & Nicolai, C. (2013). Axiomatic truth, syntax and metatheoretic reasoning. *Review of Symbolic Logic, 6*, 613–636.
19. Maddy, P., & Väänänen, J. (2022). Philosophical uses of categoricity arguments, arXiv:2204.13754.
20. McGee, V. (1997). How we learn mathematical language. *The Philosophical Review, 106*, 35–68.
21. Mount, B. M., & Waxman, D. (2021). Stable and unstable theories of truth and syntax. *Mind, 130*, 439–473.
22. Parsons, C. (1990). The uniqueness of the natural numbers. *Iyyun: The Jerusalem Philosophical Quarterly, 39*, 13–44.
23. Parsons, C. (2008). *Mathematical Thought and its Objects*. Cambridge University Press.
24. Putnam, H. (1980). Models and reality. *The Journal of Symbolic Logic*, *45*, 464–482. Reprinted in Putnam 1983, 1-25.
25. Shapiro, S. (1991). *Foundations without Foundationalism*. Clarendon Press.
26. Simpson, S. G. (1999). *Subsystems of Second Order Arithmetic*. Springer Verlag.

27. Simpson, S. G., & Yokoyama, K. (2013). Reverse mathematics and Peano categoricity. *Annals of Pure and Applied Logic, 164*, 284–293.
28. Väänänen, J. (2012). Second-order logic or set theory. *Bulletin of Symbolic Logic, 18*, 91–121.
29. Väänänen, J. (2020). Tracing internal categoricity. *Theoria*, *0*, 1–10.
30. Väänänen, J., & Wang, T. (2015). Internal categoricity in arithmetic and set theory. *Notre Dame Journal of Formal Logic, 56*, 121–134.
31. Walmsley, J. (2002). Categoricity and indefinite extensibility. *Proceedings of the Aristotelian Society, 102*, 239–257.
32. Warren, J. (2020). *Shadows of Syntax: Revitalizing Logical and Mathematical Conventionalism*. USA: Oxford University Press.