Journal of Philosophical Logic
https://doi.org/10.1007/s10992-022-09654-8

Higher-Order Logic and Disquotational Truth

Lavinia Picollo[1,2] · Thomas Schindler[3]

Received: 28 August 2020 / Accepted: 28 January 2022

Abstract
Truth predicates are widely believed to be capable of serving a certain logical or quasi-logical function. There is little consensus, however, on the exact nature of this function. We offer a series of formal results in support of the thesis that disquotational truth is a device to simulate higher-order resources in a first-order setting. More specifically, we show that any theory formulated in a higher-order language can be naturally and conservatively interpreted in a first-order theory with a disquotational truth or truth-of predicate. In the first part of the paper we focus on the relation between truth and full impredicative sentential quantification. The second part is devoted to the relation between truth-of and full impredicative predicate quantification.

Keywords Truth · Disquotation · Sentential quantification · Higher-order quantification · Theoretical equivalence

1 Introduction

Truth predicates can serve a certain logical or quasi-logical function. Examples of this function are easily found in the literature. Suppose you agree with Einstein's remarks about relativity but can't quite remember what he said. Then you can express your agreement by saying 'What Einstein said about relativity is true'. Or suppose you want to express your disagreement with a theory that isn't finitely axiomatised but you can't put your finger on the particular axioms that are to blame. Then you

✉ Lavinia Picollo
l.picollo@nus.edu.sg

Thomas Schindler
thomas.schindler1980@gmail.com

1 University College London, London, UK

2 National University of Singapore, Singapore, Singapore

3 University of Bristol, Bristol, UK

Published online: 07 May 2022

can say ‘Not every theorem of the theory is true’. However, it is not entirely clear how this function can be best understood or more precisely characterised.

Frequently, the truth predicate is described as a device for expressing infinite conjunctions and disjunctions (e.g. Putnam [1], Gupta [2]); sometimes it is said to provide finite axiomatisability (e.g. Halbach [3]); others have characterised it as a device for sentential quantification (e.g. Azzouni [4], Grover et al. [5], Parsons [6]). In [7] we have argued that the first two accounts are only partially correct and fall short of providing a full account of the logical function of truth. In [8] we have proposed a positive account based on the third proposal: the truth predicate, together with ordinary first-order quantifiers, enables us to emulate quantification into sentence and predicate position within a first-order framework – i.e. we can emulate all inferences. In other words, the truth predicate provides us with the same *inferential* power as higher-order quantifiers but in a grammatically conservative way, that is, without abandoning the convenient framework of first-order logic. In [8] we offered philosophical arguments in support of this thesis; in the present one we provide a series of formal results to this effect.

We focus first on the relation between truth and quantification into sentence position – ‘sentential quantification’, for short. Assume we want to express the factivity of knowledge as a single claim, this is, to universally generalise on φ’s position in the schema

$$\mathrm{K}\varphi \rightarrow \varphi$$

where K is a sentential operator for knowledge. This is usually done with the aid of sentential quantifiers, as follows:

$$\forall \alpha\, (\mathrm{K}\alpha \rightarrow \alpha) \qquad (\text{Factivity}_\alpha)$$

We submit that the following principle formulated in terms of a truth predicate T instead does exactly the same job as far as the inferential behaviour is concerned:

$$\forall v\, (\mathrm{Sent}(v) \rightarrow (\mathrm{K}(\mathrm{T}v) \rightarrow \mathrm{T}v)) \qquad (\text{Factivity}_\mathrm{T})$$

provided that $\mathrm{Sent}(v)$ applies to the appropriate range of sentences, and each sentence in this range is intersubstitutable with its truth ascription also in the context of the knowledge operator.[1] That is, we maintain that the inferential behaviour of ($\text{Factivity}_\mathrm{T}$) is equivalent to that of (Factivity_α).

Similarly, if one wishes to express that some statements made by Einstein hold, one can turn to a one-place operator E, for ‘Einstein said that’, and sentential quantifiers and write

$$\exists \alpha\, (\mathrm{E}\alpha \wedge \alpha) \qquad (\text{Einstein}_\alpha)$$

[1] Alternatively, one could deploy a one-place predicate K^* for knowledge and simply write

$$\forall v\, (\mathrm{Sent}(v) \rightarrow (\mathrm{K}^*(v) \rightarrow \mathrm{T}v)) \qquad (\text{Factivity}^*_\mathrm{T})$$

In that case, the intersubstitutability of sentences and their truth ascriptions in the context of knowledge attributions is not required. However, it can be argued that, if the truth predicate is understood purely disquotationally, the intersubstitutability of truth in the scope of knowledge or belief attributions is philosophically unproblematic (cf. Stern [9, §7.3.2]).

We submit that the same effect can be achieved by means of a truth predicate instead, by

$$\exists v\,(\mathrm{Sent}(v) \wedge (\mathrm{E}(\mathrm{T}v) \wedge \mathrm{T}v)) \tag{Einstein$_\mathrm{T}$}$$

with the same proviso as in the previous example.

In general, sentential quantifiers allow us to quantify into sentence position over a given class of sentences (possibly containing sentential quantifiers themselves), namely those for which we have instances of a comprehension principle for the new quantifiers (cf. Section 3.1). We would like to suggest that disquotational truth predicates work in a similar way: they allow us to generalise into sentence position precisely over the class of sentences for which we have an instance of disquotation.

To support our claims, we show that every admissible inference in a language with (full impredicative) sentential quantifiers can be translated into an admissible inference in a suitable disquotational truth theory without sentential quantifiers *in a natural way*, that is, along the lines of the given examples. Roughly, expressions like (Factivity$_\alpha$) are translated as (Factivity$_\mathrm{T}$), and expressions like (Einstein$_\alpha$) as (Einstein$_\mathrm{T}$). This is a natural translation inasmuch as (Factivity$_\mathrm{T}$), (Einstein$_\mathrm{T}$), and expressions of the like are straightforward formalisations of the natural language readings we usually make of (Factivity$_\alpha$), (Einstein$_\alpha$), and other expressions involving sentential quantification. We typically deploy a truth predicate to pronounce these latter expressions in natural language, since a more direct reading would be ungrammatical. The idea behind the truth theory is to include an instance of disquotation for each translation of a formula in the range of comprehension.

Moreover, we show that the truth theory does not license any more inferences between translations than sentential quantifiers between the original formulae. In other words, sentential quantifiers allow for the same inferences as disquotational truth, modulo the given translation.

We propose an analogous analysis of the relation between satisfaction and (full impredicative) second-order quantification, understood as a theoretical device for quantification into predicate position.[2] Assume we wish to express (the following version of) Leibniz's Law as a single claim by universally generalising on φ's position in

$$\forall u \forall v\,(u = v \rightarrow (\varphi(u) \rightarrow \varphi(v))) \tag{LL-schema}$$

Using second-order quantifiers, we can write

$$\forall V \forall u \forall v\,(u = v \rightarrow (Vu \rightarrow Vv)) \tag{LL}$$

We submit that a disquotational truth-of or satisfaction predicate, Sat, affords us the same inferential power. For one can formulate Leibniz's Law as a single claim as follows:

$$\forall w\,(\mathrm{Form}^1(w) \rightarrow \forall u \forall v\,(u = v \rightarrow (\mathrm{Sat}(w, u) \rightarrow \mathrm{Sat}(w, v)))) \tag{LL$_\mathrm{T}$}$$

where $\mathrm{Form}^1(v)$ holds exactly of all relevant formulae with just one free individual variable.

[2]That is, we understand second-order quantifiers as a primitive device for quantification into predicate position. We do not commit to a particular understanding of what kind of entities these quantifiers range over, or whether they range over entities at all.

To support our thesis, we show that every admissible inference in a language with (full impredicative) second-order quantification can be translated into an admissible inference in a suitable disquotational truth or satisfaction theory without second-order quantification *in a natural way*, that is, along the lines of the given examples. Expressions like (LL) are roughly translated as (LL_T). Again, this is a natural translation, as (LL_T) and expressions of the like are straightforward formalisations of the natural language readings we typically make of (LL) and other expressions involving second-order quantification. We usually employ a truth-of or satisfaction predicate to pronounce the latter into natural language, as more direct readings would be ungrammatical. As before, the idea behind the truth or satisfaction theory is to include an instance of (uniform) disquotation for each naturally translated formula in the range of second-order comprehension.

Furthermore, we show that the disquotational theories do not license any more inferences between translations than second-order quantification does between the original formulae. In other words, second-order quantifiers afford us exactly the same inferences as disquotational truth or satisfaction, modulo the given translation.

In fact, the results mentioned above can be generalised to higher orders: every inference in a language with (full impredicative) n-th-order quantifiers can be translated into an admissible inference in a suitable disquotational truth or satisfaction theory and no more inferences are valid in the latter.

In the more formal literature on the subject one can find many results relating theories of truth to theories formulated in a higher-order language. For example, it is well known that the theory of arithmetical comprehension ACA and the theory of typed compositional truth CT are mutually interpretable (cf. Takeuti [10, chap. 3.18] and Halbach [11, chap. 8]). A similar result holds between the system of ramified truth $\mathsf{RT}_{<\epsilon_0}$ (essentially a formalisation of Tarski's hierarchical truth theory) and ramified analysis $\mathsf{RA}_{<\epsilon_0}$, and between the Kripke-Feferman theory of truth KF and $\mathsf{RA}_{<\epsilon_0}$ (cf. Feferman [12] and Halbach [13, 14]).

However, these results are somehow partial and limited in scope. First, they often relate *predicative* fragments of higher-order logic to theories of truth,[3] leaving the wrong impression that full impredicative higher-order quantification cannot be handled by a truth predicate. Second, they often relate these fragments of higher-order logic to *compositional* theories of truth, suggesting that they cannot be handled by a disquotational truth predicate.[4] On the other hand, in the philosophical literature there is an almost exclusive focus on the relation between *sentential* quantifiers and the truth predicate, neglecting the fact that the truth predicate can actually allow us to mimic quantification into predicate position as well.

By contrast, the results offered in this paper show that full impredicative sentential and predicate quantification (for all finite orders) can be emulated by suitable disquotational theories of truth or satisfaction. We would like to emphasise, however,

[3] Of course, there are exceptions. For example, the impredicative theory of inductive definitions ID_1 is known to be mutually interpretable with the truth theory known as VF (cf. Cantini [15]). See Friedman and Sheard [16] for another example.

[4] Again, there are exceptions. For example, the theory of positive uniform disquotation PUTB has been shown to be proof-theoretically equivalent to $\mathsf{RA}_{<\epsilon_0}$ (cf. Halbach [17]).

that we do not necessarily endorse these as our best truth theories, as the choice of axioms is a bit derivative. It would be desirable if more natural systems, from which these axioms follow, could be devised.

Having a proper grasp of just how powerful disquotational truth can be is philosophically important. For example, in [8] and [18] we have argued that our account of the function of truth helps to dispel the objection that a deflationary theory of truth needs to be conservative:[5] if it is the job of the truth predicate to emulate higher-order quantification, then we shouldn't be surprised that the addition of a truth predicate can lead to non-conservative extensions, because the same holds for higher-order quantification. For a more detailed philosophical analysis of our results we refer the reader to our aforementioned papers.

Before we move to the formal part of the paper, let us briefly address an objection to our account raised by Nicolai [21]. He argues that, if disquotational truth were a tool to simulate higher-order quantification in a first-order setting, a strong form of theoretical equivalence should hold between disquotational truth theories and corresponding higher-order systems.[6] As Nicolai correctly points out, mere mutual interpretability, proof-theoretic equivalence, or mutual truth-definability do not suffice to support the view defended in this paper. The reason is that in general these forms of theoretical equivalence are unreliable. They hold between obviously disparate systems, as they can be based on unnatural translations.[7] As Nicolai [21, p. 20] argues "a natural correspondence between truth and quantification" is required. However, stronger forms of theoretical equivalence that would avoid this issue – e.g. bi-interpretability and definitional equivalence – fail to hold even in the simplest cases (cf. Nicolai [21, Prop. 4]). Nicolai concludes that the absence of an adequate form of theoretical equivalence between higher-order systems and truth theories to support the view that truth is a device to mimic higher-order quantification strongly suggests that the view is mistaken.

In this paper we propose an alternative criterion for theoretical equivalence between truth and higher-order resources that relies on *natural translations*. Thus, despite being weaker than bi-interpretability and definitional equivalence, it avoids the charges held against mutual interpretability and other insufficient forms of theoretical equivalence. We believe that our criterion successfully supports our claim that disquotational truth (or satisfaction) is a device to simulate higher-order resources in a first-order setting.

The paper is structured as follows. In Section 2, we give some technical preliminaries for the remainder of the paper. In Section 3 we focus on the relation between truth and sentential quantification. Section 4 is devoted to truth and second-order quantification in languages that contain a name for every object in the intended domain of the first-order quantifiers, and Section 5 to satisfaction and second-order quantification, without the previous qualification.

[5]See, for instance, Shapiro [19] and Ketland [20].

[6]See Halbach [11, chap. 6] for an overview of the different kinds of theoretical equivalence deployed in the literature and Fujimoto [22] for a more in-depth discussion.

[7]For example, PA is mutually interpretable with $\mathsf{PA} + \neg\mathrm{Con}(\mathsf{PA})$ (i.e. PA together with an axiom stating that PA is inconsistent), but few would consider these theories to have the same theoretical status.

2 Preliminaries

Let $\mathcal{L}$ be any effectivised[8] (standard) first-order language, with logical symbols $=, \neg, \wedge$, and $\forall$ and individual variables $v_1, v_2, \ldots$ Other logical symbols should be understood as the usual abbreviations. $\mathcal{L}$ may also include individual constants $a_1, a_2, \ldots$ and predicates $F_1^n, F_2^n, \ldots$ of each arity n, and other non-logical symbols to be specified. To simplify our presentation we don't allow function symbols in $\mathcal{L}$, but our results can be rather easily extended to languages with them. $\mathcal{L}$ is intended as (a formal representation of) the language we use *prior* to the addition of expressive resources that may allow us to generalise into sentence or formula position.

In what follows, let $u, v, \ldots$ stand for arbitrary individual variables, c for arbitrary constants, t for arbitrary terms, F for arbitrary predicates, $\varphi, \psi, \ldots$ for arbitrary formulae, and $\Gamma, \Delta, \ldots$ for arbitrary sets of formulae, possibly with subindexes. We identify each language with the set of its formulae and terms.

Let $\mathcal{L}_\mathrm{T}$ expand $\mathcal{L}$ with a fresh copy of the signature of the language of Peano arithmetic (PA), $\mathscr{L}_\mathsf{PA}$, and monadic predicates D, for the domain of discourse of $\mathcal{L}$, and T, for truth. $\mathscr{L}_\mathsf{PA}$ will be used to express syntactic notions of $\mathcal{L}$ and the expansions thereof with which we will work in Section 3 and Section 4.[9]

(By contrast, in Section 5 we will use a set-theoretic language to talk about the syntax of $\mathcal{L}$.) We assume a fixed effective coding # of expressions of $\mathcal{L}$ and these expansions by numbers. If n is the code of an expression ϵ, we write $\ulcorner\epsilon\urcorner$ for the numeral of n in $\mathcal{L}_\mathrm{T}$ – i.e. the term that consists of n occurrences of the successor function symbol S followed by the constant 0.

Note that neither '$\mathcal{L}$' nor '$\mathcal{L}_\mathrm{T}$' denote particular languages; they are both variables ranging over languages with the aforementioned characteristics. By contrast, '$\mathscr{L}_\mathsf{PA}$' is a proper name for the language of PA. In subsequent sections we will introduce further assumptions about $\mathcal{L}$ and $\mathcal{L}_\mathrm{T}$ as appropriate.

We assume $\mathscr{L}_\mathsf{PA}$ contains predicate and function symbols for primitive recursive (pr) sets and functions to be specified, including a monadic predicate N for the set of natural numbers.[10] As is well known, finite sequences of numbers can be recursively coded by numbers themselves. Let $\langle v_1, \ldots, v_n\rangle^n$ be an n-place function symbol for the pr function that maps each n-tuple of natural numbers to its code. We often omit the superscript, for readability. Also, let $\mathrm{sub}(u, v)$ define the pr function *substitution*,

[8] Very roughly, an effectivised language is a language whose symbols are gödelised – entailing that the set of its symbols is countable and the set of its codes computable. For further details, see Monk [23, def. 10.2].

[9] To prove general results to the effect that the apparatus of higher-order quantification can be consistently and conservatively substituted with that of truth, the signatures of $\mathcal{L}$ and $\mathscr{L}_\mathsf{PA}$ must be disjoint (and the truth predicate of $\mathcal{L}_\mathrm{T}$ must not occur in $\mathcal{L}$ either). This is why we work with *fresh* copies of the arithmetical and truth-theoretic vocabulary, thereby excluding theories that say something about the (same) syntactic objects – i.e. the natural numbers. Note, however, that these theories are consistently extensible with our truth-theoretic machinery all the same, although the resulting systems might have more (translated) consequences than their higher-order counterparts. For, if any such theory Γ is indeed about the natural numbers (or truth), then it must be sound, so it should not contradict our syntactic or truth-theoretic principles (provided, of course, that the latter are themselves sound). Nonetheless, Γ might entail some (sound) syntactic or truth principles already, which is why conservativeness isn't guaranteed.

[10] We assume that in the formulation of the axioms of Peano arithmetic all quantifiers are relativised to N.

that takes the code u of a formula φ and the code v of a finite sequence of numbers $k_1, \ldots, k_n$ and returns the code of the formula that results from replacing in φ the free variable with the lowest subindex with the numeral of k_1, the free variable with the second lowest subindex with the numeral of k_2, and so on. Finally, if $v_1, \ldots, v_n$ (with increasing subindexes) occur free in φ, we abbreviate $\mathrm{sub}(\ulcorner\varphi\urcorner, \langle v_1, \ldots, v_n\rangle)$ by $\ulcorner\varphi(\dot{v}_1, \ldots, \dot{v}_n)\urcorner$.

For any expressions ϵ_1, ϵ_2 of the same grammatical category, let $\varphi[\epsilon_1/\epsilon_2]$ be the result of substituting ϵ_1 for every occurrence of ϵ_2 (which should be free in case ϵ_2 is a variable), where ϵ_2 is free for ϵ_1 in φ. Finally, let any sequence of distinct expressions $\epsilon_1, \ldots, \epsilon_n$ be abbreviated by $\boldsymbol{\epsilon}$.

All the calculi we work with throughout the article extend a standard natural deduction calculus for first-order logic with identity, FOL, given by the following axiom and rules of inference:

(IL) $\forall v\, v = v$
(LL) $t_1 = t_2, \varphi[t_1/v] \vdash \varphi[t_2/v]$
($\neg$I) If $\varphi \vdash \psi$ and $\varphi \vdash \neg\psi$, then $\vdash \neg\varphi$
(EFQ) $\varphi, \neg\varphi \vdash \psi$
(DN) $\neg\neg\varphi \vdash \varphi$
($\wedge$I) $\varphi, \psi \vdash \varphi \wedge \psi$
($\wedge$E) $\varphi \wedge \psi \vdash \varphi/\psi$
($\forall v$I) $\varphi \vdash \forall v\, \varphi$, if v doesn't occur in an undischarged assumption
($\forall v$E) $\forall v\, \varphi \vdash \varphi[t/v]$

3 Truth and Sentential Quantification

In this section we show that truth can mimic full impredicative sentential quantification in a natural way, as indicated in the introduction – i.e. we can simulate all inferences. Adding sentential quantifiers to any first-order language can be perfectly simulated by the addition of a truth predicate governed by a restricted form of disquotation instead. Furthermore, we show that adding a truth predicate of those characteristics doesn't add any inferential power to our language, above and beyond those already afforded by sentential quantification. Interestingly, although the disquotational truth theories in question are formulated in a language containing arithmetical vocabulary, no arithmetical axioms are needed.

In more detail, we first expand $\mathcal{L}$ to a language $\mathcal{L}_\mathrm{S}$ with sentential quantifiers and extend FOL to a calculus SQL for $\mathcal{L}_\mathrm{S}$ with inference rules for the new quantifiers. Then, in Section 3.2, we provide a semantics for $\mathcal{L}_\mathrm{S}$ and show that SQL is sound and complete with respect to it. In Section 3.3 we give a *natural* translation η of expressions of $\mathcal{L}_\mathrm{S}$ into $\mathcal{L}_\mathrm{T}$, which will help us formulate a truth theory $\mathsf{UTB}[\eta]$ in $\mathcal{L}_\mathrm{T}$, roughly consisting of all instances of (uniform) disquotation for translations. We then show in Section 3.4 that *any* theory Γ formulated in SQL can be relatively interpreted in the $\mathcal{L}_\mathrm{T}$-theory $\eta(\Gamma)$, formulated over $\mathsf{UTB}[\eta]$. Moreover, in Section 3.5 we show that if an inference between translations is admissible in $\mathsf{UTB}[\eta]$ then the inference between the corresponding original formulae of $\mathcal{L}_\mathrm{S}$ is admissible in SQL, relying on the completeness result proved in Section 3.2. Hence, the inferential power of our truth theory doesn't exceed the inferential power of sentential quantification.

3.1 Sentential Quantification

Throughout this section we allow $\mathcal{L}$ to contain any (computable) subset of sentential constants $s_1, s_2, \ldots$, which play the syntactic role of atomic sentences of the language, and finitely many non-logical operators $\Box_1^n, \Box_2^n, \ldots$ of various arities n. Our definitions and proofs, however, are given for monadic operators only, to facilitate readability. (When it isn't entirely obvious how to generalise to arbitrary arities, we provide precise formulations in footnotes.) Non-logical operators are used to express modalities such as 'Einstein said that ...', 'it's a theorem of arithmetic that ...', or '... grounds ...'. Unlike predicates, operators of this kind can interact directly with sentential quantifiers, allowing us to generalise on schematic principles as we've seen in the examples in the introduction.[11]

Let $\mathcal{L}_S$ expand $\mathcal{L}$ with sentential – also known as 'propositional' – variables $\alpha_1, \alpha_2, \ldots$ and a λ-term $\lambda.\varphi$ for each sentence φ of $\mathcal{L}_S$. Sentential variables are (open) formulae of $\mathcal{L}_S$. By contrast, λ-terms are sentences of this language. They are introduced to the language solely for technical reasons, as we explain after Definition 7.

It should be noted that, in the absence of non-logical operators (and quotation functions or functions of the like), sentential quantification can be simulated in first-order logic alone, without the need of a truth predicate. For, in that case, $\forall\alpha\,\varphi$ can be defined simply as $\varphi[\top/\alpha] \wedge \varphi[\bot/\alpha]$ and $\exists\alpha\,\varphi$ as $\varphi[\top/\alpha] \vee \varphi[\bot/\alpha]$, where $\top$ stands for any logical truth and $\bot$ for any logical falsity. By contrast, non-logical operators bring forth an intensional aspect. For instance, we would like to keep open the possibility that, for some sentence φ that is neither a logical truth nor a logical falsity, $\Box\varphi$ but neither $\Box\top$ nor $\Box\bot$ – e.g. if $\Box$ is to express 'Einstein said that ...'.

In what follows, let $\alpha, \beta, \ldots$ stand for arbitrary sentential variables, s for arbitrary sentential constants, and $\Box$ for arbitrary non-logical sentential operators.

Definition 1 Let SQL formulated in $\mathcal{L}_S$ extend FOL with the following rules of inference:

($\forall\alpha$I)	$\varphi \vdash \forall\alpha\,\varphi$, if α doesn't occur in an undischarged assumption
($\forall\alpha$E)	$\forall\alpha\,\varphi \vdash \varphi[\psi/\alpha]$, if ψ contains no free individual variables
(λI)	$\varphi \vdash \lambda.\varphi$, if φ is a sentence
(λE)	$\lambda.\varphi \vdash \varphi$, if φ is a sentence
(SLE)	If $\vdash \varphi \leftrightarrow \psi$, then $\vdash \Box\varphi \leftrightarrow \Box\psi$[12]

[11] If one used predicates instead of operators, a device structurally equivalent to a quotation function would be needed. However, these functions are not without problems. See, for instance, Tarski [24, pp. 160-2].

[12] The rule can be generalised to modalities with arbitrarily many arguments in the following way: If $\vdash \psi \leftrightarrow \chi$, then $\vdash \Box(\varphi_1, \ldots, \psi, \ldots, \varphi_n) \leftrightarrow \Box(\varphi_1, \ldots, \chi, \ldots, \varphi_n)$.

Note that ($\forall\alpha$E) only allows us to instantiate the sentential universal quantifier with a formula without free individual variables. This is for quantificational hygiene purposes. Otherwise, sentential quantifiers would allow us to quantify not only into sentence but also into formula position. We will deal with quantification into predicate position in subsequent sections.

(SLE) states that logically equivalent expressions are intersubstitutable in all modal contexts. It entails a coarse-grained account of the kind of "objects" to which the modalities apply. This is highly undesirable, for instance, if one wishes to deploy an operator $\Box_G$ to express a hyperintensional notion such as the logical grounding relation, as it is usually done. According to most theories of logical grounding, while for each true sentence φ, $\Box_G(\varphi, \neg\neg\varphi)$, i.e. φ grounds $\neg\neg\varphi$, it's not the case that $\Box_G(\varphi, \varphi)$, i.e. φ doesn't ground itself. Note, however, that the relative interpretability result we offer in Section 3.4 doesn't rely on (SLE). It is only when we prove the converse direction – that our truth theory doesn't outstrip the inferential resources available in SQL – that substitution under logical equivalents is required. Thus, every inference that can be carried out in SQL minus (SLE) can be naturally simulated in a disquotational truth theory without (SLE), whereas it's an open question whether more inferences can be drawn in the truth theory without (SLE) than in SQL in absence of this rule.

Finally, note that SQL entails the following:

(Sentential Comprehension) $\exists\alpha\,(\alpha \leftrightarrow \varphi)$, if neither α nor individual variables occur free in φ

Note that free as well as bound sentential variables may occur in φ, i.e. this is a principle of full impredicative comprehension. To see why it follows, consider the following derivation in SQL:

1.	$\forall\alpha\,\neg(\alpha \leftrightarrow \varphi)$	assumption
2.	$\neg(\varphi \leftrightarrow \varphi)$	($\forall\alpha$E)
3.	$\neg\forall\alpha\,\neg(\alpha \leftrightarrow \varphi)$	($\neg$I), 1-2

3.2 A Semantics for Sentential Quantification

In this section, we provide a semantics for $\mathcal{L}_S$ and show that SQL is sound and complete with respect to it. In Section 3.5, the completeness result will help us show that every inference between translations of formulae of $\mathcal{L}_S$ that can be drawn in our truth theory can be also drawn in SQL between the corresponding original expressions.

Fine [25] and Kaplan [26] have provided a semantics for languages with sentential quantifiers and intensional operators that generalises Kripke's possible-world semantics for modal logic. Unfortunately, both of them focus on propositional rather than first-order languages. Moreover, neither can be extended to a semantics for first-order languages that is adequate for our present purposes. Whilst in Kaplan's semantics the sentential operators are only intended to express alethic modalities, in Fine's the truth conditions of sentences of the form $\Box\varphi$ at a world are given in terms of the accessibility relation, as is usual. As a consequence, in both cases the interpretations of the sentential operators are closed under logical consequence so, if $\Box$ were to express, e.g. 'Einstein said that ...', we would be committed to the idea that Einstein said – the content of – every logical consequence of 'The most incomprehensible thing about the universe is that it is comprehensible'. But this seems undesirable. Note also that, in Kaplan's semantics, the range of the sentential quantifiers in each model is the whole power set of the set of possible worlds. Hence, if we extended this semantics to a first-order language, no calculus would be complete with respect to it.

By contrast, we provide a Henkin-style semantics, in which sentential quantifiers are allowed to range over just a subset of the power set of the set of possible worlds, and the interpretation of the sentential operators is only closed under logical equivalence. It should be emphasised that our semantics is intended merely as a formal tool to establish our results: we wouldn't like to imply that this is the intended semantics for $\mathcal{L}_S$.

Definition 2 Let an $\mathcal{L}_S$-structure $\mathcal{M}$ be a quadruple $\langle D_\mathcal{M}, W_\mathcal{M}, P_\mathcal{M}, \|\cdot\|_{\mathcal{M}^+}\rangle$ such that $D_\mathcal{M} \neq \varnothing$, $W_\mathcal{M} \neq \varnothing$, $P_\mathcal{M} \subseteq \wp(W_\mathcal{M})$ and $\|\cdot\|_{\mathcal{M}^+} = \{\|\cdot\|^w_{\mathcal{M}^+} \mid w \in W_\mathcal{M}\}$, where each $\|\cdot\|^w_{\mathcal{M}^+}$ is an interpretation function for the non-logical vocabulary of $\mathcal{L}_S$ satisfying the following conditions:

- $\forall w' \in W_\mathcal{M},\ \|c\|^{w'}_{\mathcal{M}^+} = \|c\|^w_{\mathcal{M}^+} \in D_\mathcal{M}$
- $\|F^n_i\|^w_{\mathcal{M}^+} \subseteq D^n_\mathcal{M}$
- $\|s\|^w_{\mathcal{M}^+} \in \{\mathrm{t}, \mathrm{f}\}$
- $\|\Box\|^w_{\mathcal{M}^+} \in \wp(P_\mathcal{M})$[13]

Informally, $D_\mathcal{M}$ is the universe of discourse for the first-order variables to range over, $W_\mathcal{M}$ is a set of possible worlds, $P_\mathcal{M}$ is a set of propositions (where a proposition is simply a set of possible worlds) for the sentential variables to range over, and each $\|\cdot\|^w_{\mathcal{M}^+}$ assigns an interpretation to the non-logical vocabulary built from elements of $D_\mathcal{M}$ as usual, except sentential constants take truth values t or f and non-logical operators take sets of propositions in $P_\mathcal{M}$ as their interpretation. Since individual constants are rigid designators (they receive the same interpretation in each world), let us simply write $\|c\|_{\mathcal{M}^+}$.

Definition 3 An assignment g for an $\mathcal{L}_S$-structure $\mathcal{M}$ maps each individual term to $D_\mathcal{M}$ and each sentential variable to $P_\mathcal{M}$ so that $g(c) = \|c\|_{\mathcal{M}^+}$.

For each variable ξ of any type, let $g[o : \xi]$ be an assignment identical to g except it maps ξ to o.

Definition 4 Let $\mathcal{M}$ be an $\mathcal{L}_S$-structure, $w \in W_\mathcal{M}$, and g an assignment for $\mathcal{M}$. Truth in $\mathcal{M}, w, g$ is given as follows:

- $\mathcal{M}, w, g \models t_1 = t_2$ iff $g(t_1) = g(t_2)$
- $\mathcal{M}, w, g \models F(t_1, \dots, t_n)$ iff $\langle g(t_1), \dots, g(t_n)\rangle \in \|F\|^w_{\mathcal{M}^+}$
- $\mathcal{M}, w, g \models \alpha$ iff $w \in g(\alpha)$

[13] In general, if $\Box$ is an n-place operator, we have $\|\Box\|^w_{\mathcal{M}^+} \in \wp(P^n_\mathcal{M})$.

- $\mathcal{M}, w, g \vDash \lambda.\varphi$ iff $\mathcal{M}, w, g \vDash \varphi$
- $\mathcal{M}, w, g \vDash s$ iff $\|s\|^{w}_{\mathcal{M}+} = \mathrm{t}$
- $\mathcal{M}, w, g \vDash \Box\varphi$ iff $\{w' \in W_{\mathcal{M}} \mid \mathcal{M}, w', g \vDash \varphi\} \in \|\Box\|^{w}_{\mathcal{M}+}$ [14]
- $\mathcal{M}, w, g \vDash \neg\varphi$ iff $\mathcal{M}, w, g \nvDash \varphi$
- $\mathcal{M}, w, g \vDash \varphi \wedge \psi$ iff $\mathcal{M}, w, g \vDash \varphi$ and $\mathcal{M}, w, g \vDash \psi$
- $\mathcal{M}, w, g \vDash \forall v\, \varphi$ iff $\forall d \in D_{\mathcal{M}}, \mathcal{M}, w, g[d : v] \vDash \varphi$
- $\mathcal{M}, w, g \vDash \forall \alpha\, \varphi$ iff $\forall p \in P_{\mathcal{M}}, \mathcal{M}, w, g[p : \alpha] \vDash \varphi$

$\mathcal{M}, w \vDash \varphi$ just in case $\mathcal{M}, w, g \vDash \varphi$ for every assignment g for $\mathcal{M}$, and $\mathcal{M} \vDash \varphi$ just in case $\mathcal{M}, w \vDash \varphi$ for every $w \in W_{\mathcal{M}}$.

A formula of the form $\Box\varphi$ is true in $\mathcal{M}, w, g$ just in case the set of possible worlds in which φ is true under g belongs to the interpretation of $\Box$ in w. Simply put, the proposition expressed by φ belongs to the interpretation of $\Box$ in w.

If $\mathcal{M}$ is an $\mathcal{L}_{\mathrm{S}}$-structure, let us write $\mathcal{M}, w, g \vDash \Gamma$ as short for '$\mathcal{M}, w, g \vDash \psi$ for every $\psi \in \Gamma$'. The following definition of a faithful $\mathcal{L}_{\mathrm{S}}$-structure is modelled after the Henkin semantics for second-order languages.

Definition 5 $\mathcal{M}$ is a faithful $\mathcal{L}_{\mathrm{S}}$-structure iff every instance of (Sentential Comprehension) is true in $\mathcal{M}$.

Definition 6 If $\varphi \in \mathcal{L}_{\mathrm{S}}$ and $\Gamma \subseteq \mathcal{L}_{\mathrm{S}}$, φ is an SQL-consequence of Γ ($\Gamma \vDash_{\mathsf{SQL}} \varphi$) iff, for every faithful $\mathcal{L}_{\mathrm{S}}$-structure $\mathcal{M}$, $w \in W_{\mathcal{M}}$, and assignment g for $\mathcal{M}$, if $\mathcal{M}, w, g \vDash \Gamma$, then $\mathcal{M}, w, g \vDash \varphi$.

The rest of this section is devoted to proving the following adequacy theorem:

Theorem 1 *Let $\Gamma \subseteq \mathcal{L}_{\mathrm{S}}$ and $\varphi \in \mathcal{L}_{\mathrm{S}}$:*

$$\Gamma \vDash_{\mathsf{SQL}} \varphi \ \textit{ iff } \ \Gamma \vdash_{\mathsf{SQL}} \varphi.$$

We only prove this result for finite Γ as, first, this is all we need for the purposes of this paper and, second, the proof of the more general result is considerably longer and more convoluted.[15] The right-to-left direction – i.e. the soundness of SQL with respect to the class of faithful $\mathcal{L}_{\mathrm{S}}$-structures – follows easily by induction on the length of proofs. For the proof of the left-to-right direction, that is, the completeness of SQL, we will follow Henkin's standard strategy for the most part. We first extend every consistent set of sentences of $\mathcal{L}_{\mathrm{S}}$ to a maximally consistent set with the witness property. Then, we show that any such set has a model. This latter result will become useful in Section 3.5 as well.

[14] In general, if $\Box$ is an n-place operator, $\mathcal{M}, w, g \vDash \Box(\varphi_1 \ldots, \varphi_n)$ iff $\langle\{w' \in W_{\mathcal{M}} \mid \mathcal{M}, w', g \vDash \varphi_1\}, \ldots, \{w' \in W_{\mathcal{M}} \mid \mathcal{M}, w', g \vDash \varphi_n\}\rangle \in \|\Box\|^{w}_{\mathcal{M}+}$.

[15] The difficulty lies on the left-to-right direction, that is, the completeness of the calculus with respect to SQL-consequence, for which we rely on a version of Henkin's Lemma – our Lemma 1. For reasons that will become clear later on (cf. fn 19), if we let Γ be infinite in Theorem 1, we should generalise Lemma 1 to sets of formulae, not just sentences. But this complicates the proof of the latter a great deal (cf. fn 18).

Lemma 1 *Every consistent set of sentences of $\mathcal{L}_S$ has a faithful model $\mathcal{M}$ in which $P_{\mathcal{M}}$ is countable.*

Proof Let $\mathcal{L}_S^*$ expand $\mathcal{L}_S$ with fresh sentential constants $s_1^*, s_2^*, \ldots$ and fresh individual constants $a_1^*, a_2^*, \ldots$. These fresh constants will play the role of witnesses in our maximally consistent sets. Let $\varphi_1, \varphi_2, \ldots$ be an enumeration of the sentences of $\mathcal{L}_S^*$ s.t., for each n, both a_n^* and s_n^* occur in sentences only after the n-th place.

Let $\Gamma \subseteq \mathcal{L}_S$ be a consistent set of sentences. We first extend Γ to a maximally consistent set of sentences $\Delta \subseteq \mathcal{L}_S^*$ with the witness property, that is:[16]

1. For every sentence φ of $\mathcal{L}_S^*$, if $\varphi \notin \Delta$, then $\Delta, \varphi \vdash_{\mathsf{SQL}} \bot$.
2. If $\varphi_n := \exists v\, \psi \in \Delta$, then $\psi[a_n^*/v] \in \Delta$.
3. If $\varphi_n := \exists \alpha\, \psi \in \Delta$, then $\psi[s_n^*/\alpha] \in \Delta$.

Thus, in Δ every existential claim φ_n has a witness, a_n^* or s_n^*. Let:

- $\Gamma_0 = \Gamma$
- $\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_n\}, & \text{if } \Gamma_n, \varphi_n \nvdash_{\mathsf{SQL}} \bot \text{ and } \varphi_n \neq \exists \xi\, \psi \\ \Gamma_n \cup \{\varphi_n, \psi[a_n^*/v]\}, & \text{if } \Gamma_n, \varphi_n \nvdash_{\mathsf{SQL}} \bot \text{ and } \varphi_n := \exists v\, \psi \\ \Gamma_n \cup \{\varphi_n, \psi[s_n^*/\alpha]\}, & \text{if } \Gamma_n, \varphi_n \nvdash_{\mathsf{SQL}} \bot \text{ and } \varphi_n := \exists \alpha\, \psi \\ \Gamma_n, & \text{otherwise} \end{cases}$
- $\Delta = \bigcup_{i \in \omega} \Gamma_i$

Δ can be easily seen to satisfy conditions 1-3 and to be a subset of Γ. Δ is also guaranteed to be consistent by construction, as adding together with each existential statement a corresponding witnessing claim results in inconsistency only if adding the existential statement alone does so too. If $\varphi_n := \exists v\, \psi$, given our enumeration of sentences of $\mathcal{L}_S^*$, a_n^* cannot occur in Γ_n, φ_n. Thus, if $\Gamma_n \cup \{\varphi_n, \psi[a_n^*/v]\}$ is inconsistent a_n^* does not play an essential role in the derivation of a contradiction: everything that can be derived from $\psi[a_n^*/v]$ in this context follows equally from $\exists v\, \psi$ (i.e. φ_n). Therefore, Γ_n, φ_n must be inconsistent too. Similar remarks hold if $\varphi_n := \exists \alpha\, \psi$.

Let $[\Delta]$ be the class of maximally consistent sets of sentences of $\mathcal{L}_S^*$ with the witness property that share with Δ the same identity statements and the same sentences of the form $(s_n^* \leftrightarrow \varphi)$, where $\exists \alpha\, (\alpha \leftrightarrow \varphi)$ – which belongs to every maximally consistent set – is the n-th sentence in our enumeration. Note that the relation between two individual constants c_1 and c_2 of $\mathcal{L}_S^*$ that holds just in case $c_1 = c_2 \in \Delta$ – or, what is the same, $c_1 = c_2$ belongs to every member of $[\Delta]$ – is easily seen to be an equivalence relation. Let $[c]$ be the equivalence class to which c belongs.

Next, we define a faithful $\mathcal{L}_S^*$-structure with countable domains for both first-order and sentential quantifiers and a world in which all members of Δ – and, thus, all members of Γ – are true. Let $\mathcal{M}_{[\Delta]} = \langle D_{[\Delta]}, W_{[\Delta]}, P_{[\Delta]}, \|\cdot\|_{[\Delta]} \rangle$ be given as follows:

i. $D_{[\Delta]} = \{[c] \mid c \in \mathcal{L}_S^*\}$
ii. $W_{[\Delta]} = [\Delta]$

[16] Recall formulae of the form $\exists \xi\, \varphi$ are short for $\neg \forall \xi\, \neg \varphi$.

iii. $P_{[\Delta]} = \{\{\Sigma \in [\Delta] \mid s \in \Sigma\} \mid s \in \mathcal{L}^*_S\}$
iv. For each $\Sigma \in W_{[\Delta]}$,

(a) $\|c\|^{\Sigma}_{[\Delta]} = [c]$
(b) $\|F\|^{\Sigma}_{[\Delta]} = \{\langle [c_1], \ldots, [c_n]\rangle \mid F(c_1, \ldots, c_n) \in \Sigma\}$
(c) $\|\Box\|^{\Sigma}_{[\Delta]} = \{\{\Pi \in [\Delta] \mid \varphi \in \Pi\} \mid \Box\varphi \in \Sigma\}$[17]
(d) $\|s\|^{\Sigma}_{[\Delta]} = \text{t iff } s \in \Sigma$

Thus, the first-order domain of $\mathcal{M}_{[\Delta]}$, $D_{[\Delta]}$, consists of the equivalence classes of individual constants that co-denote according to all sets in $[\Delta]$, and each member of the latter set serves as a possible world in the model – i.e. belongs to $W_{[\Delta]}$. The domain for the sentential quantifiers, $P_{[\Delta]}$, contains exactly one proposition for each sentential constant s, namely, the set of worlds in which s is true. Thus, $P_{[\Delta]}$ must be countable. Finally, in each world Σ, the interpretation of predicate letters and non-logical operators is grounded on sentences belonging to Σ in which those symbols occur.

Note that, for clause iv.(c) to work, we must have that, for each sentence φ, $\{\Pi \in [\Delta] \mid \varphi \in \Pi\} \in P_{[\Delta]}$. By iii, there must be a constant s s.t. $\{\Pi \in [\Delta] \mid \varphi \in \Pi\} = \{\Pi \in [\Delta] \mid s \in \Pi\}$, that is, $(s \leftrightarrow \varphi)$ must belong to every $\Pi \in [\Delta]$. This is guaranteed by the fact that $\exists\alpha\,(\alpha \leftrightarrow \varphi) \in \Delta$ and so, if this existential claim occurs in the n-th place of our enumeration, by the way in which Δ is constructed, $(s^*_n \leftrightarrow \varphi) \in \Delta$ too. By definition of $[\Delta]$, $(s^*_n \leftrightarrow \varphi)$ must also belong to every member of $[\Delta]$.

Next, we show that, for every sentence φ and every $\Sigma \in [\Delta]$,

$$\mathcal{M}_{[\Delta]}, \Sigma \models \varphi \text{ iff } \varphi \in \Sigma,$$

by induction on the logical complexity of φ.

$\varphi := s$: The result follows directly from iv(d).

$\varphi := F(c_1, \ldots, c_n)$: The result follows directly from iv(a) and iv.(b).

Assume the result holds of every sentence of lower complexity than φ. We reason by cases:

$\varphi := \Box\psi$:

$$\begin{aligned} \Box\psi \in \Sigma \quad &\text{iff } \{\Pi \in [\Delta] \mid \psi \in \Pi\} \in \|\Box\|^{\Sigma}_{[\Delta]} && \text{iv(c)} \\ &\text{iff } \{\Pi \in [\Delta] \mid \mathcal{M}_{[\Delta]}, \Pi \models \psi\} \in \|\Box\|^{\Sigma}_{[\Delta]} && \text{I.H.} \\ &\text{iff } \mathcal{M}_{[\Delta]}, \Sigma \models \Box\psi && \text{Def. 4} \end{aligned}$$

[17] In general, if $\Box$ is an n-place operator, $\|\Box\|^{\Sigma}_{[\Delta]}$ is the following set:

$$\{\langle\{\Pi \in [\Delta] \mid \varphi_1 \in \Pi\}, \ldots, \{\Pi \in [\Delta] \mid \varphi_n \in \Pi\}\rangle \mid \Box(\boldsymbol{\varphi}) \in \Sigma\}$$

$\varphi := \forall\alpha\,\psi$:

$$
\begin{array}{llr}
\forall\alpha\,\psi \in \Sigma & \text{iff } \psi[s/\alpha] \in \Sigma, \text{ for each } s & 3 \\
& \text{iff } \mathcal{M}_{[\Delta]}, \Sigma \models \psi[s/\alpha], \text{ for each } s & \text{I.H.} \\
& \text{iff } \forall p \in P_{[\Delta]},\ \mathcal{M}_{[\Delta]}, \Sigma, g[p:\alpha] \models \psi & \text{iii} \\
& \text{iff } \mathcal{M}_{[\Delta]}, \Sigma, g \models \forall\alpha\,\psi & \text{Def. 4} \\
& \text{iff } \mathcal{M}_{[\Delta]}, \Sigma \models \forall\alpha\,\psi & \text{Def. 4}
\end{array}
$$

The remaining cases can be established in an analogous way. This concludes our induction.[18]

Since $\Delta \in [\Delta]$, we can conclude that all members of Δ are true in $\mathcal{M}_{[\Delta]}, \Delta$. Let $\mathcal{M}$ be the reduct of $\mathcal{M}_{[\Delta]}$ to $\mathcal{L}_\mathrm{S}$. Since $W_\mathcal{M} = W_{[\Delta]}$, we have that $\mathcal{M}, \Delta \models \Gamma$. And, since $P_\mathcal{M} = P_{[\Delta]}$, $P_\mathcal{M}$ must also be countable. Finally, note that $\mathcal{M}$ is a faithful $\mathcal{L}_\mathrm{S}$-structure as, for every formula $\varphi \in \mathcal{L}_\mathrm{S}$ with only $\boldsymbol{\beta}$ free, $\forall\boldsymbol{\beta}\exists\alpha\,(\alpha \leftrightarrow \varphi)$ must belong to every $\Sigma \in [\Delta]$, so it follows from our previous result that, for every $\Sigma \in [\Delta]$, $\mathcal{M}, \Sigma \models \forall\boldsymbol{\beta}\exists\alpha\,(\alpha \leftrightarrow \varphi)$. Hence, $\mathcal{M} \models \exists\alpha\,(\alpha \leftrightarrow \varphi)$ too. This concludes the proof of Lemma 1. □

To finish the proof of Theorem 1, let us assume that $\Gamma \nvdash_\mathsf{SQL} \varphi$ and let $\boldsymbol{\alpha}$ and $\mathbf{v}$ be the only free variables occurring in Γ, φ (recall Γ must be finite). Thus, $\neg\forall\boldsymbol{\alpha}\forall\mathbf{v}\,(\bigwedge_{\psi\in\Gamma}\psi \rightarrow \varphi)$ must be consistent. By Lemma 1, there is a faithful $\mathcal{L}_\mathrm{S}$-structure $\mathcal{M}$ with a $w \in W_\mathcal{M}$ such that $\mathcal{M}, w \not\models \forall\boldsymbol{\alpha}\forall\mathbf{v}\,(\bigwedge_{\psi\in\Gamma}\psi \rightarrow \varphi)$.[19] By Definition 6, $\Gamma \not\models_\mathsf{SQL} \varphi$. This completes our proof.

3.3 Translation and Truth

Let $\mathcal{L}_\mathrm{S}^+$ expand $\mathcal{L}_\mathrm{S}$ with fresh sentential constants $s_1^*, s_2^*, \ldots$. Similarly, let $\mathcal{L}_\mathrm{T}^+$ expand $\mathcal{L}_\mathrm{T}$ with $s_1^*, s_2^*, \ldots$. We now provide a translation η of $\mathcal{L}_\mathrm{S}^+$ into $\mathcal{L}_\mathrm{T}^+$ and formulate a corresponding disquotational truth theory $\mathsf{UTB}[\eta]$ that will allow us to simulate the inferential resources of SQL. The fresh sentential constants are needed to ensure that for every consistent set of sentences Γ of $\mathcal{L}_\mathrm{S}$ we can find a faithful model $\mathcal{M}$ in which $P_\mathcal{M}$ is countable (as in Lemma 1). If there were no such model, we wouldn't

[18] If we wish to generalise Lemma 1 to consistent sets of *formulae* (not of just sentences), we must not only find a model $\mathcal{M}_{[\Delta]}$ and a world $w \in W_{[\Delta]}$ but also an assignment g under which all members of our new Δ are true. To establish the induction case for $\varphi := \forall\alpha\,\psi$, it's important that every set in $P_{[\Delta]}$ corresponds to a sentential constant, as given by iii. So we have to make sure that g assigns to each sentential variable α a set $\{\Sigma \in W_{[\Delta]} \mid s \in \Sigma\}$. This is not trivial, as we could in principle have, for instance, that $\neg(\alpha \leftrightarrow s) \in \Delta$ (or $\neg(\Box\alpha \leftrightarrow \Box s) \in \Delta$) for each sentential constant s of $\mathcal{L}_\mathrm{S}^*$. Making sure this doesn't happen adds a thick layer of complication we would rather avoid.

[19] Note that the standard completeness proof-strategy is blocked here. We can infer that $\Gamma \cup \{\neg\varphi\}$ is consistent from $\Gamma \nvdash_\mathsf{SQL} \varphi$, but Lemma 1 doesn't allow us to conclude that the set has a model (and, so, that $\Gamma \not\models_\mathsf{SQL} \varphi$). Lemma 1 is restricted to sets of sentences, whilst Theorem 1 isn't. This is important for the proof of Theorem 3, which (very roughly) shows that a disquotational truth predicate does not license any more inferences between *formulae* (not just sentences) than do sentential quantifiers. This is why we focus on the sentence $\neg\forall\boldsymbol{\alpha}\forall\mathbf{v}\,(\bigwedge_{\psi\in\Gamma}\psi \rightarrow \varphi)$ instead of $\Gamma \cup \{\neg\varphi\}$. However, this is only possible if Γ is finite, as otherwise the antecedent would be infinitely long. For infinite Γ, a generalisation of Lemma 1 to formulae is required, but the proof is considerably more complicated (cf. fn 18), as noted in fn 15.

be able to translate Γ into the language of truth while preserving the meaning of the arithmetical vocabulary, because the translation maps statements involving sentential variables into statements involving truth ascriptions, and of course there are only countably many sentences to which we can ascribe truth. We could avoid the use of these fresh sentential constants if we used (structured) propositions instead of sentences as truth bearers. However, at least as far as the present paper is concerned, we prefer to use sentences as truth bearers, as there are well-established theories of sentences available in the literature (e.g. Peano arithmetic) but no widely accepted, formal theory of structured propositions.[20]

In order to provide the translation, let us (recursively) split the variables $v_1, v_2, \ldots$ of $\mathcal{L}_{\mathrm{T}}^{+}$ into two infinite sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ Variables belonging to the first sequence are called 's-variables', as they are the variables to which we will map the sentential variables of $\mathcal{L}_{\mathrm{S}}^{+}$. Variables in the second sequence are reserved for translating the individual variables of $\mathcal{L}_{\mathrm{S}}^{+}$.

Via our fixed coding of expressions by numbers, Kleene's Recursion Theorem (cf. Hinman [27]) entails the existence of a pr function $\eta : \mathcal{L}_{\mathrm{S}}^{+} \to \mathcal{L}_{\mathrm{T}}^{+}$ such that:[21,22]

$$\eta(t) = \begin{cases} y_i, & \text{if } t := v_i \\ c, & \text{if } t := c \end{cases}$$

$$\eta(\varphi) = \begin{cases} s, & \text{if } \varphi := s \\ F(\eta(t_1), \ldots, \eta(t_n)), & \text{if } \varphi := F(t_1, \ldots, t_n) \\ \mathrm{T}x_i, & \text{if } \varphi := \alpha_i \\ \mathrm{T}\ulcorner\eta(\psi)\urcorner, & \text{if } \varphi := \lambda.\,\psi \\ \Box\eta(\psi), & \text{if } \varphi := \Box\psi \\ \neg\eta(\psi), & \text{if } \varphi := \neg\psi \\ \eta(\psi) \wedge \eta(\chi), & \text{if } \varphi := \psi \wedge \chi \\ \forall y_i\,(\mathrm{D}(y_i) \to \eta(\psi)), & \text{if } \varphi := \forall v_i\,\psi \\ \forall x_i\,(\mathrm{Trsl}_\eta(x_i) \to \eta(\psi)), & \text{if } \varphi := \forall \alpha_i\,\psi \end{cases}$$

where D is a monadic predicate of $\mathcal{L}_{\mathrm{T}}$ intended to express the domain of discourse of $\mathcal{L}$ (cf. Section 2) and $\mathrm{Trsl}_\eta(v)$ abbreviates $\exists u\,(\mathrm{Sent}_{\mathcal{L}_{\mathrm{S}}^{+}}(u) \wedge \varphi_\eta(u, v))$, where $\mathrm{Sent}_{\mathcal{L}_{\mathrm{S}}^{+}}(v)$ is a predicate of $\mathcal{L}_{\mathsf{PA}}$ for the set of (codes of) sentences of $\mathcal{L}_{\mathrm{S}}^{+}$, $\varphi_\eta(u, v)$ a formula of $\mathscr{L}_{\mathsf{PA}}$ expressing the function η itself, and u is a fresh variable. In other words, $\mathrm{Trsl}_\eta(v)$ defines the set of expressions of $\mathcal{L}_{\mathrm{T}}^{+}$ that are translations of $\mathcal{L}_{\mathrm{S}}^{+}$-sentences.

Our translation behaves as we anticipated in the introduction: sentential variables α_i are translated into singular truth ascriptions to the corresponding s-variable x_i and logical structure is otherwise preserved, except quantifiers are relativised

[20] Note that the fresh sentential constants play no role in the proofs of Theorems 2 and 3; they are only used in Corollaries 4 and 5.

[21] If $\epsilon \in \mathcal{L}_{\mathrm{S}}$, we write $\eta(\epsilon)$ instead of $\eta(\#\epsilon)$, for perspicuity.

[22] If $\Box$ is an n-place operator and $\varphi := \Box(\varphi_1, \ldots, \varphi_n)$, $\eta(\varphi)$ is $\Box(\eta(\varphi_1), \ldots, \eta(\varphi_n))$.

appropriately. Moreover, the non-logical vocabulary of $\mathcal{L}_\mathrm{S}^+$ remains unmodified. Thus, η is a natural translation.

Next, we introduce a theory formulated in $\mathcal{L}_\mathrm{T}^+$ with a disquotational truth predicate for translations. To simulate inferences of SQL, our truth theory must contain some principles that allow it to tell between syntactic objects and the objects that $\mathcal{L}_\mathrm{S}^+$-sentences are about. In addition, a syntactic principle governing translations and a version of (SLE) governing the non-logical operators is be required. Note that although $\mathcal{L}_\mathrm{T}^+$ contains the signature of $\mathscr{L}_{\mathsf{PA}}$, no axioms of PA governing these symbols will be included in our truth theory.

The last point deserves some emphasis: in order to simulate higher-order quantification, none of the usual arithmetical axioms are required in the truth theory. (Nor does our truth theory include any truth-theoretic axioms for the language $\mathscr{L}_{\mathsf{PA}}$.) On the one hand, this supports the idea that the function of the truth predicate is indeed of a quasi-logical nature: our results do not rely on a strong base theory. On the other hand, since the formulation of our truth theory involves function symbols and other expressions from $\mathscr{L}_{\mathsf{PA}}$ (as is standard practice in the literature), one may expect the truth theory to prove the standard principles governing these functions. Otherwise, those symbols can completely deviate from their intended meanings, making it unclear whether the language can be truly understood as talking about expressions.[23] Fortunately, all our results – in particular, the conservativeness results in Section 3.5 – still hold if the usual arithmetical axioms, together with truth-theoretic axioms for $\mathscr{L}_{\mathsf{PA}}$, are added to our truth theory, as we'll see later on (cf. Corollaries 4 and 5).

For the following definition, note that if $\varphi \in \mathcal{L}_\mathrm{S}^+$, then α_i occurs free in φ just in case x_i is free in $\eta(\varphi)$ and, similarly, v_i occurs free in φ just in case y_i is free in $\eta(\varphi)$.

Definition 7 $\mathsf{UTB}[\eta] \subseteq \mathcal{L}_\mathrm{T}^+$ is the theory formulated in FOL given by the following axioms and rule:

(D) $\quad \mathrm{D}c$, for each individual constant c of $\mathcal{L}$

(Trsl$_\eta$) $\quad \forall \mathbf{x}_k\, (\bigwedge_{i\leqslant n} \mathrm{Trsl}_\eta(x_{k_i}) \to \mathrm{Trsl}_\eta(\ulcorner\varphi(\dot{\mathbf{x}}_k)\urcorner))$, if φ is in the range of η with only s-variables free, $\mathbf{x}_k$

(SLE$_\eta$) $\quad$ If φ and ψ are in the range of η with only $\mathbf{x}_j$ and $\mathbf{y}_k$ free and

$$\vdash_{\mathsf{UTB}[\eta]} \forall \mathbf{x}_j \forall \mathbf{y}_k\, (\bigwedge_{i\leqslant m} \mathrm{Trsl}_\eta(x_{j_i}) \wedge \bigwedge_{i\leqslant n} \mathrm{D}(y_{k_i}) \to (\varphi \leftrightarrow \psi)), \text{ then}$$

$$\vdash_{\mathsf{UTB}[\eta]} \forall \mathbf{x}_j \forall \mathbf{y}_k\, (\bigwedge_{i\leqslant m} \mathrm{Trsl}_\eta(x_{j_i}) \wedge \bigwedge_{i\leqslant n} \mathrm{D}(y_{k_i}) \to (\Box\varphi \leftrightarrow \Box\psi))$$

(UTB$_\eta$) $\quad \forall \mathbf{x}_k\, (\bigwedge_{i\leqslant n} \mathrm{Trsl}_\eta(x_{k_i}) \to (\mathrm{T}\ulcorner\varphi(\dot{\mathbf{x}}_k)\urcorner \leftrightarrow \varphi(\mathbf{x}_k)))$, if φ is in the range of η with only s-variables free, $\mathbf{x}_k$

Following the notation introduced in the preliminaries, $\mathrm{Trsl}_\eta(\ulcorner\varphi(\dot{\mathbf{x}}_k)\urcorner)$ is short for $\mathrm{Trsl}_\eta(\mathrm{sub}(\ulcorner\varphi(\mathbf{x}_k)\urcorner, \mathbf{x}_k\rangle))$, and $\mathrm{T}\ulcorner\varphi(\dot{\mathbf{x}}_k)\urcorner$ is short for $\mathrm{Tsub}(\ulcorner\varphi(\mathbf{x}_k)\urcorner, \langle\mathbf{x}_k\rangle)$.

Axiom (D) guarantees that every individual constant of $\mathcal{L}$ denotes an object in D. (Trsl$_\eta$) states that, if we substitute translated sentences for the free s-variables in a

[23] We are grateful to an anonymous referee for raising this issue.

translation, the result is also a translated sentence. Note that this axiom, needed for proofs later, wouldn't be true if we didn't have λ-terms in $\mathcal{L}_S^+$: while Tx_i and $\eta(\varphi)$ are translations, $T\ulcorner\eta(\varphi)\urcorner$ isn't unless $\lambda.\varphi$ is in the domain of η. (SLE_η) is just a translated version of (SLE), where the corresponding ranges of the variables occurring in φ and ψ are indicated by the assumptions in the premises and in the conclusion of the rule. Finally, (UTB_η) is a uniform version of disquotation for sentences in the range of η, with s-parameters only, ranging accordingly over translated sentences.

As already mentioned in the introduction, we do not necessarily recommend $\mathsf{UTB}[\eta]$ as our ultimate truth theory, as the choice of axioms is parasitic on $\mathcal{L}_S^+$. However it would be a mistake to view them as cherry-picked; they are precisely what is required to make our philosophical point: While sentential quantifiers allow us to quantify into sentence position over the class of sentences for which we have comprehension, the truth predicate allows us to quantify into sentence position over the class of sentences for which we have disquotation. If we restrict our attention to the simplified case of a class C of sentences which don't themselves contain sentential variables, then the whole machinery of $\mathsf{UTB}[\eta]$ would not be needed in order to prove inferential equivalence results;[24] all that is needed on the side of the truth theory are the instances of disquotation for the members of C – a natural, non-*ad hoc*, set of truth-theoretic axioms. The reason why $\mathsf{UTB}[\eta]$ involves a more complex (and perhaps less perspicuous) set of truth-theoretic axioms is to handle the more general case where C may contain sentences in which sentential quantifiers occur. Here we need to add instances of disquotation for the *translations* of the sentences in C instead, as members of C may contain vocabulary alien to the language of truth. But as we have argued, our translation is a natural one; if that is so, then there is nothing *ad hoc* about adding the disquotation instances for the translations of $\mathcal{L}_S^+$-sentences. Of course, if we opt for a truth predicate rather than sentential quantifiers to generalise on sentence positions to begin with, the class of translations of sentences of the language with sentential quantifiers would not be the most natural class we might wish to generalise over, but other classes will be.

3.4 Preservation of Inferences

The main result of this section establishes that any inference one can draw in SQL from some premises to some conclusion can also be drawn in $\mathsf{UTB}[\eta]$ from the translations of the premises to the translation of the conclusion, together with assumptions fixing the respective range of the variables that occur free in these translations. If only sentences are involved in premises and conclusion, no extra assumptions are needed. Simply put, we show that η preserves all valid inferences from SQL to $\mathsf{UTB}[\eta]$.[25]

If $\Gamma \subseteq \mathcal{L}_S$, let $\eta(\Gamma) = \{\eta(\psi) \mid \psi \in \Gamma\}$, $\mathrm{Trsl}_\eta(\Gamma) = \{\mathrm{Trsl}_\eta(x_i) \mid x_i$ occurs free in $\eta(\Gamma)\}$ and $\mathrm{D}_\eta(\Gamma) = \{\mathrm{D}(y_i) \mid y_i$ occurs free in $\eta(\Gamma)\}$.

[24] In this case, quantification into sentence position over members of C requires parameter-free, predicative comprehension, and this can easily be mimicked by typed disquotation. A proof of this result, for the case of quantification into predicate position, can be found in Halbach [11, Corollary 8.37].

[25] As we already mentioned in the previous subsection, it is noteworthy that no arithmetical axioms are required in the truth theory.

For perspicuity, we will often drop the subindexes of the variables v_i and α_i and assume that $\eta(v) = y$ and $\eta(\alpha) = \mathrm{T}x$.

Theorem 2 *Let $\Gamma \subseteq \mathcal{L}_{\mathrm{S}}$ and $\varphi \in \mathcal{L}_{\mathrm{S}}$:*

$$\textit{If } \Gamma \vdash_{\mathsf{SQL}} \varphi, \textit{ then } \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi), \eta(\Gamma) \vdash_{\mathsf{UTB}[\eta]} \eta(\varphi).$$

To prove this result we first need the following lemmata.

Lemma 2 *If t_2 doesn't occur in the scope of a λ-term in $\varphi \in \mathcal{L}_{\mathrm{S}}^{+}$,*

$$\eta(\varphi)[\eta(t_1)/\eta(t_2)] = \eta(\varphi[t_1/t_2]).$$

This result follows easily from the definition of η by induction on the logical complexity of φ. The proof of the following is slightly more convoluted.

Lemma 3 *If the only variables occurring free in $\eta(\psi)$ are $\mathbf{x}_k$,*

$$\mathsf{UTB}[\eta] \vdash \forall \mathbf{x}_k \, (\bigwedge_{i \leqslant n} \mathrm{Trsl}_\eta(x_{k_i}) \to (\eta(\varphi)[\ulcorner \eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] \leftrightarrow \eta(\varphi[\psi/\alpha])))$$

Roughly, this lemma establishes (a parametrised version of) the following: applying η to the result of substituting α for ψ in φ is equivalent in $\mathsf{UTB}[\eta]$ to applying η to α, φ, and ψ first and then performing the substitution on the resulting expressions.

Proof The proof is by induction on the logical complexity of φ. If α doesn't occur free in φ, the result follows trivially, as there's nothing to substitute. Thus, we assume α occurs free in the formula and reason in $\mathsf{UTB}[\eta]$ from the assumption that $\bigwedge_{i \leqslant n} \mathrm{Trsl}_\eta(x_{k_i})$:

$\varphi := \alpha$:

$$\begin{aligned} \eta(\alpha)[\ulcorner \eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] &\leftrightarrow \mathrm{T}x[\ulcorner \eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] && \text{def. } \eta \\ &\leftrightarrow \mathrm{T}\ulcorner \eta(\psi)(\dot{\mathbf{x}}_k)\urcorner \\ &\leftrightarrow \eta(\psi) && (\mathrm{UTB}_\eta) \\ &\leftrightarrow \eta(\alpha[\psi/\alpha]) \end{aligned}$$

Assume that the result holds for every formula of lower complexity than φ. We reason by cases:[26]

$\varphi := \Box\chi$:

$$\begin{aligned} \eta(\Box\chi)[\ulcorner \eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] &\leftrightarrow \Box\eta(\chi)[\ulcorner \eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] && \text{def. } \eta \\ &\leftrightarrow \Box\eta(\chi[\psi/\alpha]) && \text{I.H., } (\mathrm{SLE}_\eta) \\ &\leftrightarrow \eta(\Box\chi[\psi/\alpha]) && \text{def. } \eta \end{aligned}$$

[26]Note that the case for $\varphi := \lambda.\psi$ can be skipped, as no free variables may occur in ψ.

$\varphi := \forall\alpha_i\ \chi$ (where $\alpha_i \neq \alpha$):

$$\begin{aligned}
\eta(\forall\alpha_i\ \chi)[\ulcorner\eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] &\leftrightarrow \forall x_i\ (\mathrm{Trsl}_\eta(x_i) \to \eta(\chi))[\ulcorner\eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x] \qquad \text{def. } \eta\\
&\leftrightarrow \forall x_i\ (\mathrm{Trsl}_\eta(x_i) \to \eta(\chi)[\ulcorner\eta(\psi)(\dot{\mathbf{x}}_k)\urcorner/x])\\
&\leftrightarrow \forall x_i\ (\mathrm{Trsl}_\eta(x_i) \to \eta(\chi[\psi/\alpha])) \qquad \text{I.H.}\\
&\leftrightarrow \eta(\forall\alpha_i\ \chi[\psi/\alpha]) \qquad \text{def. } \eta
\end{aligned}$$

The other cases can be proved in a similar fashion. □

We can now go back to Theorem 2. We prove it by induction on the length of the derivation of φ from Γ in SQL. Assume for now that no λ-terms occur in this derivation. We will later show how to modify our proof to account for the occurrence of λ-terms. Throughout the proof we only reason in UTB$[\eta]$, so let '$\vdash$' stand for derivability in this theory.

If the proof is one-step long, then either $\varphi \in \Gamma$, in which case the result follows trivially, or φ is a logical axiom. In that case, we have the following:

(IL) $\varphi := \forall v\ v = v$. Then, $\eta(\varphi)$ is $\forall y\ (\mathrm{D}(y) \to y = y)$, which is a logical truth and, thus, can be derived in UTB$[\eta]$ without assumptions.

Assume that the result holds for every Δ and every ψ if ψ can be derived from Δ in fewer than n steps. We consider the possible ways in which φ could have been inferred from Γ:

(LL) $t_1 = t_2, \psi[t_1/v] \vdash \psi[t_2/v]$, where $\varphi := \psi[t_2/v]$. We have:

$$\begin{aligned}
\eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, t_1 = t_2), \mathrm{D}_\eta(\Gamma, t_1 = t_2) &\vdash \eta(t_1 = t_2) \qquad \text{I.H.}\\
\eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \psi[t_1/v]), \mathrm{D}_\eta(\Gamma, \psi[t_1/v]) &\vdash \eta(\psi[t_1/v]) \qquad \text{I.H.}
\end{aligned}$$

that is,

$$\begin{aligned}
\eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi, t_1 = t_1) &\vdash \eta(t_1) = \eta(t_2) \qquad \text{def. } \eta\\
\eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi, t_1 = t_1) &\vdash \eta(\psi)[\eta(t_1)/\eta(v)]) \qquad \text{Lemma 2}
\end{aligned}$$

as we can replace both $t_1 = t_2$ and $\psi[t_1/v]$ on the left-hand side of the turn-style with $t_1 = t_1$, since the only variables that occur in $t_1 = t_2$ and $\psi[t_1/v]$ but not in φ occur in t_1. Thus, we have the following:

$$\begin{aligned}
\eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi, t_1 = t_1) &\vdash \eta(\psi)[\eta(t_2)/\eta(v)]) \qquad \text{(LL)}\\
&\vdash \eta(\psi[t_2/v]) \qquad \text{Lemma 2}\\
&\vdash \eta(\varphi)
\end{aligned}$$

If all free variables in t_1 occur in Γ, φ, the proof is complete. Otherwise, let $\mathbf{v}_k$ be the variables occurring free in t_1 but not in Γ, φ:

$$\begin{aligned}
\eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi) &\vdash \bigwedge_{i \leqslant n} \mathrm{D}(y_{k_i}) \to \eta(\varphi) \qquad \text{Deduction}\\
&\vdash \forall\mathbf{y}_k\ (\bigwedge_{i \leqslant n} \mathrm{D}(y_{k_i}) \to \eta(\varphi)) \qquad (\forall v\mathrm{I})\\
&\vdash \exists\mathbf{y}_k\ \bigwedge_{i \leqslant n} \mathrm{D}(y_{k_i}) \to \eta(\varphi)\\
&\vdash \eta(\varphi)\ \ \text{(D)}
\end{aligned}$$

Since $\mathbf{y}_k$ don't occur free in $\eta(\varphi)$ and axiom (D) entails that there is at least one object satisfying D, one can easily get rid of the unwanted assumption that the free variables in t_1 range over D by an application of the Deduction Theorem and a logical principle governing quantifiers. We use this trick to prove other subsequent cases as well.

($\neg$I) If $\psi \vdash \chi$ and $\psi \vdash \neg\chi$, then $\vdash \neg\psi$, where $\varphi := \neg\psi$. We have the following:

$$\begin{array}{rcll} \eta(\Gamma), \eta(\psi), \mathrm{Trsl}_\eta(\Gamma, \psi, \chi), \mathrm{D}_\eta(\Gamma, \psi, \chi) & \vdash & \eta(\chi) & \text{I.H.} \\ \eta(\Gamma), \eta(\psi), \mathrm{Trsl}_\eta(\Gamma, \psi, \neg\chi), \mathrm{D}_\eta(\Gamma, \psi, \neg\chi) & \vdash & \eta(\neg\chi) & \text{I.H.} \\ & \vdash & \neg\eta(\chi) & \text{def. } \eta \end{array}$$

Since the variables that occur free in $\neg\chi$, ψ are the same as in χ, φ,

$$\begin{array}{rcl} \eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi, \chi), \mathrm{D}_\eta(\Gamma, \varphi, \chi) & \vdash & \neg\eta(\psi)(\neg\mathrm{I}) \\ & \vdash & \eta(\neg\psi)\text{def. } \eta \end{array}$$

If all free variables in χ occur in Γ, φ, the proof is complete. Otherwise, let $\boldsymbol{\alpha}_j$ and $\mathbf{v}_k$ be the only variables occurring free in χ but not in Γ, φ. Applying the same trick as before, since, by (Trsl_η), $\mathsf{UTB}[\eta]$ entails $\exists v\, \mathrm{Trsl}_\eta(v)$, we obtain the desired result.

($\forall v$E) $\forall v\, \psi \vdash \psi[t/v]$, where $\varphi := \psi[t/v]$. We have:

$$\begin{array}{rcll} \eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \forall v\, \psi), \mathrm{D}_\eta(\Gamma, \forall v\, \psi) & \vdash & \eta(\forall v\, \psi) & \text{I.H.} \\ & \vdash & \forall y\, (\mathrm{D}(y) \to \eta(\psi)) & \text{def. } \eta \\ & \vdash & \mathrm{D}(\eta(t)) \to \eta(\psi)[\eta(t)/y] & (\forall v\mathrm{E}) \end{array}$$

Since all variables occurring free in $\forall v\, \psi$ are also free in φ, by weakening:

$$\begin{array}{rcll} \eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi) & \vdash & \mathrm{D}(\eta(t)) \to \eta(\psi)[\eta(t)/y] & \\ & \vdash & \eta(\psi)[\eta(t)/y] & (\to\mathrm{E}), (\mathrm{D}) \\ & \vdash & \eta(\psi[t/v]) & \text{Lemma 2} \end{array}$$

($\forall\alpha$E) $\forall\alpha\, \psi \vdash \psi[\chi/\alpha]$, where no individual variables occur free in χ and $\varphi := \psi[\chi/\alpha]$. Let $\boldsymbol{\alpha}_k$ be the variables occurring free in χ. By inductive hypothesis, the definition of η, and ($\forall v$E), we obtain the following:

$$\begin{array}{rcl} \eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \forall\alpha\, \psi), \mathrm{D}_\eta(\Gamma, \forall\alpha\, \psi) & \vdash & \eta(\forall\alpha\, \psi) \\ & \vdash & \forall x\, (\mathrm{Trsl}_\eta(x) \to \eta(\psi)) \\ & \vdash & \mathrm{Trsl}_\eta(\ulcorner\eta(\chi)(\dot{\mathbf{x}}_k)\urcorner) \to \eta(\psi)[\ulcorner\eta(\chi)(\dot{\mathbf{x}}_k)\urcorner/x] \end{array}$$

Since all variables occurring free in $\forall\alpha\, \psi$ are also free in φ, by weakening, ($\to$E), (Trsl_η), and Lemma 3:

$$\begin{array}{rcl} \eta(\Gamma), \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi) & \vdash & \mathrm{Trsl}_\eta(\ulcorner\eta(\chi)(\dot{\mathbf{x}}_k)\urcorner) \to \eta(\psi)[\ulcorner\eta(\chi)(\dot{\mathbf{x}}_k)\urcorner/x] \\ & \vdash & \eta(\psi)[\ulcorner\eta(\chi)(\dot{\mathbf{x}}_k)\urcorner/x] \\ & \vdash & \eta(\psi[\chi/\alpha]) \end{array}$$

The case for (SLE) follows easily from the inductive hypothesis and (SLE_η). The other cases can be established in a similar fashion.

This completes the proof of Theorem 2 for Γ, φ without λ-terms. If λ-terms $\lambda.\psi_1, \ldots, \lambda.\psi_n$ occurred in Γ, φ, then one could replace each of them with $\psi_1, \ldots, \psi_n$, translate the resulting inference into an admissible inference in

$\mathsf{UTB}[\eta]$, and then replace back each relevant occurrence of $\eta(\psi_1), \ldots, \eta(\psi_n)$ with $\mathrm{T}\ulcorner\psi_1\urcorner, \ldots, \mathrm{T}\ulcorner\psi_n\urcorner$. Since $\mathsf{UTB}[\eta]$ entails the equivalence between each sentence ψ and $\mathrm{T}\ulcorner\psi\urcorner$ *even in the scope of the non-logical operators*, the resulting inference is admissible in the theory.

Corollary 1 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_\mathrm{S}$:*

$$\textit{If } \Gamma \vdash_{\mathsf{SQL}} \varphi, \textit{ then } \eta(\Gamma) \vdash_{\mathsf{UTB}[\eta]} \eta(\varphi).$$

Slightly abusing terminology, one could take this result to show that η is a *relative interpretation* of each theory Γ formulated in SQL in the first-order theory $\eta(\Gamma) + \mathsf{UTB}[\eta]$.[27] Note, however, that η isn't just *any* relative interpretation but one that preserves the non-logical vocabulary of $\mathcal{L}_\mathrm{S}^+$ and, in addition, maps every sentential variable α to a truth ascription $\mathrm{T}x$, i.e. η is a *natural* relative interpretation, as we argued in the introduction.

3.5 Associated Models

The purpose of this section is to show that disquotational truth for a class of (translations of) expressions doesn't afford us more inferential power than sentential quantification over this class. In other words, we show that if the inference from some premises to a conclusion is not admissible in SQL, then the inference from the translations of the premises to the translation of the conclusion is not admissible in $\mathsf{UTB}[\eta]$ either. Together with the results proved in Section 3.4, this allows us to conclude that disquotational truth and sentential quantification are *equivalent means* of achieving one and the same goal, that is, to enable quantification into sentence position over a given class of sentences.

To establish the main result of this section, we first provide a semantics for $\mathcal{L}_\mathrm{T}$ and extensions thereof with respect to which FOL is sound. This is needed because $\mathcal{L}_\mathrm{T}$ is not a standard first-order language but contains non-logical sentential operators. (Again, we do not necessarily endorse this semantics as the most natural one, but it is sufficient for our purposes – i.e. the proof of Theorem 3 below.) We then show how to transform any faithful $\mathcal{L}_\mathrm{S}^+$-structure $\mathcal{M}$ with a countable domain for the sentential quantifiers into a model $\mathcal{M}^\eta$ for $\mathsf{UTB}[\eta]$, preserving truth under η (Lemma 5). Our main result – that $\mathsf{UTB}[\eta]$ is conservative over SQL (Theorem 3) – then follows easily. Indeed, since $\mathcal{M}^\eta$ interprets the arithmetical vocabulary in the standard way, we have as a corollary that $\mathsf{UTB}[\eta] + \mathsf{PA}$ – actually, $\mathsf{UTB}[\eta]$ together with any sound "syntax theory" formulated in $\mathscr{L}_{\mathsf{PA}}$ – is conservative over SQL as well. Moreover, a simple modification of our argument shows that our results still hold if we add disquotation axioms for $\mathcal{L}_{\mathsf{PA}}$ to $\mathsf{UTB}[\eta]$. That is, $\mathsf{UTB}[\eta] + \mathsf{UTB}[\mathsf{PA}]$ is conservative over SQL as well, where $\mathsf{UTB}[\mathsf{PA}]$ denotes the result of extending PA with axioms of uniform disquotation for the language of $\mathcal{L}_{\mathsf{PA}}$.

[27] Strictly speaking, no theory formulated in $\mathcal{L}_\mathrm{S}$ can be relatively interpreted in a first-order theory, as logical structure cannot be preserved since first-order languages lack sentential quantifiers.

Definition 8 An $\mathcal{L}_T$-structure $\mathcal{M}$ is a quadruple $\langle D_\mathcal{M}, W_\mathcal{M}, R_\mathcal{M}, \|\cdot\|_{\mathcal{M}^+}\rangle$ such that $D_\mathcal{M} \neq \varnothing$, $W_\mathcal{M} \neq \varnothing$, $R_\mathcal{M} \subseteq W_\mathcal{M} \times D_\mathcal{M}$ and $\|\cdot\|_{\mathcal{M}^+} = \{\|\cdot\|^w_{\mathcal{M}^+} \mid w \in W_\mathcal{M}\}$, where each $\|\cdot\|^w_{\mathcal{M}^+}$ is an interpretation function for the non-logical vocabulary satisfying the following conditions:

- $\forall w' \in W_\mathcal{M},\ \|c\|^{w'}_{\mathcal{M}^+} = \|c\|^w_{\mathcal{M}^+} \in D_\mathcal{M}$
- $\|F^n_i\|^w_{\mathcal{M}^+} \subseteq D^n_\mathcal{M}$
- $\|s\|^w_{\mathcal{M}^+} \in \{\mathrm{t}, \mathrm{f}\}$
- $\|\Box\|^w_{\mathcal{M}^+} \subseteq D_\mathcal{M}$[28]

Predicates, sentential constants, and non-logical operators may receive a different interpretation in each possible world of an $\mathcal{L}_T$-structure, just like in $\mathcal{L}_S$-structures. Non-logical operators, however, are no longer assigned sets of sets of possible worlds as their interpretation but rather subsets of members of the first-order domain, $D_\mathcal{M}$. Thus, to evaluate the truth of formulae of the form $\Box\varphi$ in each world w and assignment g, instead of looking for a set in $\|\Box\|^w_{\mathcal{M}^+}$ containing exactly those worlds in which φ is true under g, we will look for a member of $D_\mathcal{M}$ in $\|\Box\|^w_{\mathcal{M}^+}$ that is *related* via $R_\mathcal{M}$ to exactly those worlds in which φ is true under g. This is the purpose of $R_\mathcal{M}$.

Definition 9 Let $\mathcal{M}$ be an $\mathcal{L}_T$-structure, $w \in W_\mathcal{M}$, and g an assignment for $\mathcal{M}$.

- $\mathcal{M}, w, g \models t_1 = t_2$ iff $g(t_1) = g(t_2)$
- $\mathcal{M}, w, g \models F(t_1, \ldots, t_n)$ iff $\langle g(t_1), \ldots, g(t_n)\rangle \in \|F\|^w_{\mathcal{M}^+}$
- $\mathcal{M}, w, g \models s$ iff $\|s\|^w_{\mathcal{M}^+} = \mathrm{t}$
- $\mathcal{M}, w, g \models \Box\varphi$ iff $\exists p \in \|\Box\|^w_{\mathcal{M}^+}\ \forall w' \in W_\mathcal{M}, \langle w', p\rangle \in R_\mathcal{M}$ iff $\mathcal{M}, w', g \models \varphi$[29]
- $\mathcal{M}, w, g \models \neg\varphi$ iff $\mathcal{M}, w, g \not\models \varphi$
- $\mathcal{M}, w, g \models \varphi \wedge \psi$ iff $\mathcal{M}, w, g \models \varphi$ and $\mathcal{M}, w, g \models \psi$
- $\mathcal{M}, w, g \models \forall v\, \varphi$ iff $\forall d \in D_\mathcal{M}, \mathcal{M}, w, g[d : v] \models \varphi$

$\mathcal{M}, w \models \varphi$ just in case $\mathcal{M}, w, g \models \varphi$ for every assignment g for $\mathcal{M}$, and $\mathcal{M} \models \varphi$ just in case $\mathcal{M}, w \models \varphi$ for every $w \in W_\mathcal{M}$.

Lemma 4 (Soundness) *Let $\mathcal{M}$ be an $\mathcal{L}_T$-structure, $w \in W_\mathcal{M}$, and g an assignment for $\mathcal{M}$. If $\Gamma \vdash_{\mathsf{FOL}} \varphi$ and $\mathcal{M}, w, g \models \Gamma$, then $\mathcal{M}, w, g \models \varphi$.*

This lemma follows easily from Definition 9 by induction on the length of the derivation of φ from Γ in FOL. Note that there are no rules or axioms for non-logical operators in the calculus, so FOL imposes no special constraints on the truth values formulae of the form $\Box\varphi$ may take.

[28] If $\Box$ is an n-place operator, we have that $\|\Box\|^w_{\mathcal{M}^+} \subseteq D^n_\mathcal{M}$.

[29] In general, if $\Box$ is an n-place operator, $\mathcal{M}, w, g \models \Box(\varphi_1 \ldots, \varphi_n)$ iff, for some $p_1, \ldots, p_n$ s.t. $\langle \mathbf{p}\rangle \in \|\Box\|^w_{\mathcal{M}^+}$, it is the case that $\forall i \leqslant n \forall w' \in W_\mathcal{M}, \langle w', p_i\rangle \in R_\mathcal{M}$ iff $\mathcal{M}, w', g \models \varphi_i$.

The following important lemma shows that we can transform any faithful $\mathcal{L}_\mathrm{S}$-structure $\mathcal{M}$ with a countable domain for the sentential quantifiers into a $\mathcal{L}_\mathrm{T}^+$-structure, preserving truth under η.

Lemma 5 *For every faithful $\mathcal{L}_\mathrm{S}$-structure $\mathcal{M}$ in which $P_\mathcal{M}$ is countable, there is an $\mathcal{L}_\mathrm{T}^+$-structure $\mathcal{M}^\eta = \langle D^\eta_\mathcal{M}, W_\mathcal{M}, R^\eta_\mathcal{M}, \|\cdot\|_{\mathcal{M}^\eta}\rangle$ satisfying the following conditions:*

1. $\mathcal{M}^\eta$ *extends and expands both* $\mathcal{M}$ *and* $\mathbb{N}$ *to* $\mathcal{L}_\mathrm{T}^+$.
2. $\|\mathrm{D}\|_{\mathcal{M}^\eta} = D_\mathcal{M}$.
3. *For each assignment g for $\mathcal{M}$ there is an assignment g^η for $\mathcal{M}^\eta$ s.t., for every $w \in W_\mathcal{M}$ and $\varphi \in \mathcal{L}_\mathrm{S}$:*

$$\mathcal{M}^\eta, w, g^\eta \vDash \eta(\varphi) \textit{ iff } \mathcal{M}, w, g \vDash \varphi.$$

4. $\mathcal{M}^\eta \vDash \mathsf{UTB}[\eta]$.

Proof Let us first expand $\mathcal{M}$ to an $\mathcal{L}_\mathrm{S}^+$-structure $\mathcal{M}^+$. Recall that $\mathcal{L}_\mathrm{S}^+$ extends $\mathcal{L}_\mathrm{S}$ by fresh sentential constants $s_1^*, s_2^*, \ldots$ Let $p_1, p_2, \ldots$ be an enumeration of $P_\mathcal{M}$ (the propositions existing in $\mathcal{M}$). For every $w \in W_\mathcal{M}$ and $n > 0$, let

$$\|s_n^*\|^w_{\mathcal{M}^+} = \mathrm{t} \text{ iff } w \in p_n.$$

Thus, for each $p \in P_\mathcal{M}$, there is a sentential constant s such that p is the proposition expressed by s in $\mathcal{M}^+$ – i.e. $p = \{w \in W_{\mathcal{M}^+} \mid \mathcal{M}^+, w \vDash s\}$. It is easy to check that $\mathcal{M}^+$ is a faithful $\mathcal{L}_\mathrm{S}^+$-structure.

In what follows, if φ is a sentence of $\mathcal{L}_\mathrm{S}^+$, let us write $\|\varphi\|_{\mathcal{M}^+}$ for $\{w \in W_\mathcal{M} \mid \mathcal{M}^+, w \vDash \varphi\}$ – i.e. $\|\varphi\|_{\mathcal{M}^+}$ is the proposition expressed by φ in $\mathcal{M}^+$.
Let $\mathcal{M}^\eta$ be given as follows:

i. $D^\eta_\mathcal{M} = \omega \cup D_\mathcal{M}$[30]
ii. $R^\eta_\mathcal{M} = \{\langle w, \#\eta(\varphi)\rangle \mid w \in \|\varphi\|_{\mathcal{M}^+}\}$
iii. For each $w \in W_\mathcal{M}$:
   (a) $\|\epsilon\|^w_{\mathcal{M}^\eta} = \|\epsilon\|^w_{\mathcal{M}^+}$, if ϵ is a non-logical symbol of $\mathcal{L}_\mathrm{S}^+$ other than a sentential operator
   (b) $\|\epsilon\|^w_{\mathcal{M}^\eta} = \|\epsilon\|_\mathbb{N}$, if ϵ is a non-logical symbol of $\mathscr{L}_\mathsf{PA}$[31]
   (c) $\|\mathrm{D}\|^w_{\mathcal{M}^\eta} = D_\mathcal{M}$
   (d) $\|\Box\|^w_{\mathcal{M}^\eta} = \{\#\eta(\varphi) \mid \|\varphi\|_{\mathcal{M}^+} \in \|\Box\|^w_{\mathcal{M}^+}\}$[32]
   (e) $\|\mathrm{T}\|^w_{\mathcal{M}^\eta} = \{\#\eta(\varphi) \mid \langle w, \#\eta(\varphi)\rangle \in R^\eta_\mathcal{M}\}$

[30] We assume w.l.o.g. that ω and $D_\mathcal{M}$ are disjoint, as we can always work with a set different from but isomorphic to ω instead.

[31] Note that our domain contains in addition to the elements of ω also some non-numerical objects. We assume that, whenever ϵ is a function symbol of $\mathscr{L}_\mathsf{PA}$, $\|\epsilon\|^w_{\mathcal{M}^\eta}$ yields some dummy value if applied to a member of $D_\mathcal{M}$.

[32] More generally, if $\Box$ is an n-place operator,

$$\|\Box\|^w_{\mathcal{M}^\eta} = \{\langle \#\eta(\varphi_1), \ldots, \#\eta(\varphi_n)\rangle \mid \langle \|\varphi_1\|_{\mathcal{M}^+}, \ldots, \|\varphi_n\|_{\mathcal{M}^+}\rangle \in \|\Box\|^w_{\mathcal{M}^+}\}.$$

The existence of $\mathcal{M}^\eta$ is guaranteed by our assumption that $\mathcal{L}$ and $\mathscr{L}_{\mathsf{PA}}$ don't share any non-logical vocabulary, as stated in the preliminaries.

Informally, $R^\eta_{\mathcal{M}}$ relates each world w with the (codes of) translations of sentences that express true propositions in w, and $\|\Box\|^w_{\mathcal{M}^\eta}$ is the set of (codes of) translations of sentences that express a proposition in $\|\Box\|^w_{\mathcal{M}^+}$. Moreover, in each world w, T is interpreted as the set of (codes of) translations of sentences that express a true proposition in w. Note that the same proposition might be expressed in $\mathcal{M}$ by more than one sentence, so more than one code might correspond to the same proposition.

Conditions 1 and 2 are clearly satisfied. Regarding condition 3, for each assignment g for $\mathcal{M}$, let g^η be defined as follows:

$$g^\eta(v) = \begin{cases} g(v), & \text{if } v := y_i \\ \min\{\#\eta(\varphi) \mid g(\alpha_i) = \|\varphi\|_{\mathcal{M}^+}\}, & \text{if } v := x_i \end{cases}$$

Since every proposition in $P_{\mathcal{M}}$ is expressed by a sentential constant of $\mathcal{L}^+_{\mathrm{S}}$, for each α_i there must always be a sentence $\varphi \in \mathcal{L}^+_{\mathrm{S}}$ s.t. $g(\alpha_i) = \|\varphi\|_{\mathcal{M}^+}$. This guarantees that, for every s-variable x_i, $g^\eta(x_i) = \#\eta(\varphi)$, for some sentence $\varphi \in \mathcal{L}^+_{\mathrm{S}}$.

Next, we will prove a slightly stronger result that directly entails condition 3. This will become handy later, in the proof of condition 4. We show by induction on the build-up of φ that, for each assignment g for $\mathcal{M}$ – or, what's the same, $\mathcal{M}^+$ – there is an assignment g^η for $\mathcal{M}^\eta$ s.t., for every $w \in W_{\mathcal{M}}$ and $\varphi \in \mathcal{L}^+_{\mathrm{S}}$:

$$\mathcal{M}^\eta, w, g^\eta \vDash \eta(\varphi) \ \text{ iff } \ \mathcal{M}^+, w, g \vDash \varphi. \tag{1}$$

$\varphi := \alpha$: Let $g^\eta(x) = \#\eta(\psi)$, so $g(\alpha) = \{w \in W_{\mathcal{M}} \mid \mathcal{M}^+, w \vDash \psi\}$. Then:

$$\begin{aligned} \mathcal{M}^\eta, w, g^\eta \vDash \mathrm{T}x \quad & \text{iff } \ g^\eta(x) \in \|\mathrm{T}\|^w_{\mathcal{M}^\eta} && \text{Def. 9} \\ & \text{iff } \ \#\eta(\psi) \in \|\mathrm{T}\|^w_{\mathcal{M}^\eta} && \text{def. } g^\eta \\ & \text{iff } \ \langle w, \#\eta(\psi)\rangle \in R^\eta_{\mathcal{M}} && \text{iii.(e)} \\ & \text{iff } \ w \in \|\psi\|_{\mathcal{M}^+} && \text{ii} \\ & \text{iff } \ w \in g(\alpha) && \text{def. } g^\eta \\ & \text{iff } \ \mathcal{M}^+, w, g \vDash \alpha && \text{Def. 4} \end{aligned}$$

The results for $\varphi := F(t_1, \ldots, t_n)$ and $\varphi := s$ follow trivially from iii.(a). Assume the result holds for every formula of lower complexity than φ. We reason by cases:

$\varphi := \Box\psi$: Assume, first, that $\mathcal{M}^\eta, w, g^\eta \vDash \Box\eta(\psi)$. By Definition 9, there must be a $d \in \|\Box\|^w_{\mathcal{M}^\eta}$ s.t.

$$\forall w' \in W_{\mathcal{M}},\ \langle w', d\rangle \in R^\eta_{\mathcal{M}} \ \text{ iff } \ \mathcal{M}^\eta, w', g^\eta \vDash \eta(\psi).$$

By iii.(d), there must be a sentence χ of $\mathcal{L}^+_{\mathrm{S}}$ s.t. $\#\eta(\chi) \in \|\Box\|^w_{\mathcal{M}^\eta}$ and

$$\forall w' \in W_{\mathcal{M}},\ \langle w', \#\eta(\chi)\rangle \in R^\eta_{\mathcal{M}} \ \text{ iff } \ \mathcal{M}^\eta, w', g^\eta \vDash \eta(\psi).$$

By iii.(d), ii, and the inductive hypothesis, we have that there's a sentence χ of $\mathcal{L}^+_{\mathrm{S}}$ s.t. $\|\chi\|_{\mathcal{M}^+} \in \|\Box\|^w_{\mathcal{M}^+}$ and

$$\forall w' \in W_{\mathcal{M}},\ w' \in \|\chi\|_{\mathcal{M}^+} \ \text{ iff } \ \mathcal{M}^+, w', g \vDash \psi.$$

In other words, $\|\chi\|_{\mathcal{M}^+} = \{w' \in W_{\mathcal{M}} \mid \mathcal{M}^+, w', g \models \psi\}$, so $\{w' \in W_{\mathcal{M}} \mid \mathcal{M}, w', g \models \psi\} \in \|\Box\|^w_{\mathcal{M}^+}$. By Definition 4, $\mathcal{M}^+, w, g \models \Box\psi$.

We proved that, if $\mathcal{M}^\eta, w, g^\eta \models \Box\eta(\psi)$, then $\mathcal{M}^+, w, g \models \Box\psi$. The proof of the converse works precisely in reverse, where χ can be a sentential constant expressing the proposition $\{w' \in W_{\mathcal{M}} \mid \mathcal{M}^+, w', g \models \psi\} \in \|\Box\|^w_{\mathcal{M}^+}$ in $\mathcal{M}^+$.

The other cases follow easily from the inductive hypothesis. Note that, since $\mathcal{M}^+$ is just an expansion of $\mathcal{M}$, condition 3 follows immediately from the instances of (1) for formulae of $\mathcal{L}_{\mathrm{S}}$.

It only remains to be shown that condition 4 holds, that is, that $\mathcal{M}^\eta$ is a model of $\mathsf{UTB}[\eta]$. It follows from iii.(a) and iii.(c) that (D) is true in $\mathcal{M}^\eta$. Moreover, since every instance of (Trsl_η) is a sentence of $\mathcal{L}_{\mathsf{PA}}$, the truth of this axiom in $\mathcal{M}^\eta$ follows from iii.(b).

If $\boldsymbol{\alpha}_k$ are the only variables occurring free in $\varphi \in \mathcal{L}^+_{\mathrm{S}}$ and $\boldsymbol{\psi}_k$ are sentences of $\mathcal{L}^+_{\mathrm{S}}$, then, since every free occurrence of an s-variable x in $\eta(\varphi)$ must occur in the context of $\mathrm{T}x$, we have

$$\eta(\varphi)[\ulcorner\eta(\psi_{k_1})\urcorner/x_{k_1}]\dots[\ulcorner\eta(\psi_{k_n})\urcorner/x_{k_n}] = \eta(\varphi[\lambda.\psi_{k_1}/\alpha_{k_1}]\dots[\lambda.\psi_{k_n}/\alpha_{k_n}]). \quad (2)$$

Moreover, if φ and ψ are sentences of $\mathcal{L}^+_{\mathrm{S}}$, it follows by Theorem 1 that

$$\|\varphi\|_{\mathcal{M}^+} = \|\varphi[\lambda.\psi/\psi]\|_{\mathcal{M}^+}. \quad (3)$$

To prove that $\mathcal{M}^\eta \models$(UTB_η), let φ be a formula of $\mathcal{L}^+_{\mathrm{S}}$ with only α free. Our proof can be easily generalised to cases in which φ contains any number of sentential variables free. We reason from the assumption that $\mathcal{M}^\eta, w, g' \models \mathrm{Trsl}_\eta(x)$, so $g'(x) = \#\eta(\psi)$ for some sentence ψ of $\mathcal{L}^+_{\mathrm{S}}$. Since no other variables occur free in φ, we can safely assume that $g' = g^\eta$ for some assignment g for $\mathcal{M}$. Therefore, $g(\alpha) = \|\psi\|_{\mathcal{M}^+}$, so

$$\begin{array}{lll}
\mathcal{M}^\eta, w, g^\eta \models \mathrm{T}\ulcorner\eta(\varphi)(\dot{x})\urcorner & \text{iff } \mathcal{M}^\eta, w \models \mathrm{T}\ulcorner\eta(\varphi)[\ulcorner\eta(\psi)\urcorner/x]\urcorner & \text{iii.(b)}\\
& \text{iff } \mathcal{M}^\eta, w \models \mathrm{T}\ulcorner\eta(\varphi[\lambda.\psi/\alpha])\urcorner & (2)\\
& \text{iff } \#\eta(\varphi[\lambda.\psi/\alpha]) \in \|\mathrm{T}\|^w_{\mathcal{M}^\eta} & \text{Def. 9}\\
& \text{iff } \langle w, \#\eta(\varphi[\lambda.\psi/\alpha])\rangle \in R^\eta_{\mathcal{M}} & \text{iii.(e)}\\
& \text{iff } w \in \|\varphi[\lambda.\psi/\alpha]\|_{\mathcal{M}^+} & \text{ii}\\
& \text{iff } w \in \|\varphi[\psi/\alpha]\|_{\mathcal{M}^+} & (3)\\
& \text{iff } \mathcal{M}^+, w \models \varphi[\psi/\alpha] & \\
& \text{iff } \mathcal{M}^+, w, g \models \varphi & \text{def. } g^\eta\\
& \text{iff } \mathcal{M}^\eta, w, g^\eta \models \eta(\varphi) & (1)
\end{array}$$

Finally, to prove that (SLE_η) holds in $\mathcal{M}^\eta$ we show that, for every natural number n, if only n applications of (SLE_η) are allowed, then $\mathcal{M}^\eta$ is a model of the theory. We've proved the base case already, as we have shown that all axioms of $\mathsf{UTB}[\eta]$ are true in $\mathcal{M}^\eta$. Assume that, if only n applications of (SLE_η) are allowed, then $\mathcal{M}^\eta$ is also a model of the theory. Assume also that

$$\forall x \forall y\, (\mathrm{Trsl}_\eta(x) \wedge \mathrm{D}(y) \rightarrow (\Box\eta(\varphi) \leftrightarrow \Box\eta(\psi))) \quad (4)$$

can be derived in $\mathsf{UTB}[\eta]$ by only $n+1$ applications of (SLE_η), where x, y are the only variables occurring free in $\eta(\varphi), \eta(\psi)$. As before, our proof can be easily generalised to any number of free variables. This entails that

$$\forall x \forall y\, (\mathrm{Trsl}_\eta(x) \wedge \mathrm{D}(y) \rightarrow (\eta(\varphi) \leftrightarrow \eta(\psi))) \tag{5}$$

must be derivable in $\mathsf{UTB}[\eta]$ by fewer than $n+1$ applications of (SLE_η), which means that (5) must be true in $\mathcal{M}^\eta$, by inductive hypothesis.

Assume, for contradiction, that (4) is not true in $\mathcal{M}^\eta$. Thus, there must be a $w \in W_\mathcal{M}$ and an assignment g^η for $\mathcal{M}^\eta$ s.t. $\mathcal{M}^\eta, w, g^\eta \not\models \Box\eta(\varphi) \leftrightarrow \Box\eta(\psi)$. By (1), $\mathcal{M}^+, w, g \not\models \Box\varphi \leftrightarrow \Box\psi$ and, by Definition 4,

$$\{w' \in W_\mathcal{M} \mid \mathcal{M}^+, w', g \models \varphi\} \neq \{w' \in W_\mathcal{M} \mid \mathcal{M}^+, w', g \models \psi\}.$$

Again, by (1), we have

$$\{w' \in W_\mathcal{M} \mid \mathcal{M}^\eta, w', g^\eta \models \eta(\varphi)\} \neq \{w' \in W_\mathcal{M} \mid \mathcal{M}^\eta, w', g^\eta \models \eta(\psi)\},$$

that is, there is a $w' \in W_\mathcal{M}$ s.t. $\mathcal{M}^\eta, w', g^\eta \models \eta(\varphi)$ and $\mathcal{M}^\eta, w', g^\eta \not\models \eta(\psi)$ or vice versa. In other words, $\mathcal{M}^\eta, w', g^\eta \not\models \eta(\varphi) \leftrightarrow \eta(\psi)$, so (5) isn't true in $\mathcal{M}^\eta$ after all, contradicting our initial assumption. □

This completes the proof of Lemma 5. Now, together with Lemmata 1 and 4, Lemma 5 entails a conservativeness result. As we will show next, $\mathsf{UTB}[\eta]$ does not allow us to draw more inferences between translations than SQL already allows between the original formulae.

Theorem 3 *Let $\Gamma \subseteq \mathcal{L}_\mathrm{S}$ and $\varphi \in \mathcal{L}_\mathrm{S}$:*

$$\textit{If } \mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi), \eta(\Gamma) \vdash_{\mathsf{UTB}[\eta]} \eta(\varphi), \textit{ then } \Gamma \vdash_{\mathsf{SQL}} \varphi.$$

Proof Assume $\Gamma \nvdash_{\mathsf{SQL}} \varphi$. Therefore, for every finite $\Delta \subseteq \Gamma$, $\Delta \cup \{\neg\varphi\}$ must be consistent. By Lemma 1, there is a faithful model $\mathcal{M}$ in which $P_\mathcal{M}$ is countable with a $w \in W_\mathcal{M}$ and an assignment g s.t. $\mathcal{M}, w, g \models \Delta \cup \{\neg\varphi\}$. By Lemma 5 (conditions 3 and 4), there is an $\mathcal{L}_\mathrm{T}^+$-structure $\mathcal{M}^\eta \models \mathsf{UTB}[\eta]$ and an assignment g^η s.t. $\mathcal{M}^\eta, w, g^\eta \models \mathrm{Trsl}_\eta(\Delta, \varphi), \mathrm{D}_\eta(\Delta, \varphi), \eta(\Delta), \eta(\neg\varphi)$. By Lemma 4, we can conclude that, for each finite $\Delta \subseteq \Gamma$, $\mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi), \eta(\Delta) \nvdash_{\mathsf{UTB}[\eta]} \eta(\varphi)$. Therefore, we must also have that $\mathrm{Trsl}_\eta(\Gamma, \varphi), \mathrm{D}_\eta(\Gamma, \varphi), \eta(\Gamma) \nvdash_{\mathsf{UTB}[\eta]} \eta(\varphi)$. □

Corollary 2 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_\mathrm{S}$:*

$$\textit{If } \eta(\Gamma) \vdash_{\mathsf{UTB}[\eta]} \eta(\varphi), \textit{ then } \Gamma \vdash_{\mathsf{SQL}} \varphi.$$

A theory T_2 extending another theory T_1 is said to be proof-theoretically conservative over T_1 just in case, if T_2 proves a sentence of the language of T_1, T_1 proves it as well. This notion can be naturally extended to theories T_1 and T_2 such that T_2 does not extend T_1, provided that T_2 is an extension of *a translation* of T_1. In that case, let us say that T_2 is proof-theoretically conservative over T_1 *modulo the given translation* just in case, if T_2 proves the translation of a sentence φ of the language of T_1, then T_1 proves φ. Thus, Corollary 2 establishes that, for each set of sentences Γ of $\mathcal{L}_\mathrm{S}$, $\eta(\Gamma) + \mathsf{UTB}[\eta]$ is proof-theoretically conservative over $\Gamma + \mathsf{SQL}$ *modulo*

η, as η is a translation of $\mathcal{L}_{\mathrm{S}}$ into $\mathcal{L}_{\mathrm{T}}^{+}$ and $\eta(\Gamma) + \mathsf{UTB}[\eta]$ extends $\eta(\Gamma + \mathsf{SQL})$ (cf. Corollary 1).

Put together, Corollaries 1 and 2 entail the following:

Corollary 3 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_{\mathrm{S}}$:*

$$\Gamma \vdash_{\mathsf{SQL}} \varphi \text{ iff } \eta(\Gamma) \vdash_{\mathsf{UTB}[\eta]} \eta(\varphi).$$

As a consequence, since η is a natural translation, Corollary 3 allows us to conclude that the addition of a uniformly disquotational truth predicate for a given class of sentences affords us exactly the same inferential power as sentential quantifiers. The generalising function of a truth predicate governed by this kind of disquotation is no more and no less than the simulation of quantification into sentence position in a first-order setting.

Earlier we raised the question whether our results still hold if we add the usual axioms of PA to our truth theory (cf. Section 3.3). It follows from Lemma 5(1) that this is the case indeed.[33] Let $\mathrm{Tr}(\mathbb{N})$ be the set of sentences of $\mathscr{L}_{\mathsf{PA}}$ that are true in the standard model.

Corollary 4 $\mathsf{UTB}[\eta] + \mathrm{Tr}(\mathbb{N})$ *is conservative over* SQL *modulo η.*

A fortiori, $\mathsf{UTB}[\eta]$ combined with *any* sound theory formulated in $\mathscr{L}_{\mathsf{PA}}$ to serve as a syntax theory (e.g. PA) is also conservative over SQL modulo η. As the following corollary shows, Lemma 5 also entails that, if $\Gamma \subseteq \mathcal{L}_{\mathrm{S}}$ is consistent, then $\eta(\Gamma) + \mathsf{UTB}[\eta]$ together with *any* sound "syntax theory" formulated in $\mathscr{L}_{\mathsf{PA}}$ has a model in which the arithmetical vocabulary is interpreted standardly.

Corollary 5 *If $\Gamma \subseteq \mathcal{L}_{\mathrm{S}}$ is consistent, then $\eta(\Gamma) + \mathsf{UTB}[\eta]$ has a model in which the arithmetical vocabulary is interpreted by the standard model of arithmetic.*

Note that a simple modification of our argument shows that we can replace $\mathsf{UTB}[\eta]$ with $\mathsf{UTB}[\eta] + \mathsf{UTB}[\mathsf{PA}]$ in Corollaries 4 and 5, where $\mathsf{UTB}[\mathsf{PA}]$ is the result of extending PA with uniform disquotation for the language of $\mathcal{L}_{\mathsf{PA}}$.[34]

4 Truth and Predicate Quantification

In this section our goal is to establish that quantification into predicate position in a language with a name for each object in its intended first-order domain can be simulated by means of a truth predicate, as anticipated in the introduction. We give a natural translation of an arbitrary second-order language of those characteristics into

[33] Recall that the axioms of PA are relativised to the predicate N (cf. fn 10).

[34] One simply needs to add all $\mathcal{L}_{\mathsf{PA}}$-sentences that are true in $\mathbb{N}$ to the extension of the truth predicate $\|\mathrm{T}\|^{\eta}_{\mathcal{M}}$ in the proof of Lemma 5.

a corresponding first-order language with a truth predicate and show that the inferences licensed by full impredicative second-order quantification translate exactly into the admissible inferences in a truth theory whose truth predicate is governed by a restricted version of uniform disquotation. (As before, it is interesting that although the truth theories in question are formulated in a language containing arithmetical vocabulary, no arithmetical axioms are needed.) To facilitate readability, we focus on second-order quantification, but the results in this section are easily generalised to all higher orders.

More fully, we first expand $\mathcal{L}$ to $\mathcal{L}_2$, a language with quantification into predicate position, and extend FOL to a second-order calculus SOL with inference rules and axioms for the new quantifiers. We then provide a *natural* translation τ of $\mathcal{L}_2$ into $\mathcal{L}_\mathrm{T}$ and formulate a corresponding truth theory $\mathsf{UTB}[\tau]$, containing, roughly, an instance of uniform disquotation for each formula in the range of τ. We then show that *any* admissible inference from Γ to φ in SOL can be translated into an inference from $\tau(\Gamma)$ to $\tau(\varphi)$ in $\mathsf{UTB}[\tau]$. In addition, we show that $\mathsf{UTB}[\tau]$ does not license any inference between translations that SOL doesn't already allow for between the corresponding original expressions. We do so by showing that every model of SOL can be transformed into a model of $\mathsf{UTB}[\tau]$ that preserves truth under τ. In other words, we show that SOL and $\mathsf{UTB}[\tau]$ have the same inferential power.[35]

4.1 Quantification into Predicate Position

Let $\mathcal{L}$ be as in the preliminaries. We further assume it contains a canonical name – i.e. a designated closed term – for each object in its intended domain. Let $\mathcal{L}_2$ expand $\mathcal{L}$ with predicate variables $V_1^n, V_2^n, \ldots$ of each arity n, and a λ-term $\{\lambda \mathbf{v}.\, \varphi\}$ (playing the syntactic role of an n-place predicate) for every formula φ with exactly $\mathbf{v}$ free.

Let U, V stand for arbitrary predicate variables.

Definition 10 Let SOL formulated in $\mathcal{L}_2$ extend FOL with the following rules and axioms:

($\forall V$I) $\quad \varphi \vdash \forall V\, \varphi$, if V doesn't occur in an undischarged assumption
($\forall V$E) $\quad \forall V_i^n\, \varphi \vdash \varphi[V_j^n/V_i^n]$
(Comprehension) $\quad \exists V_i^n \forall \mathbf{v}\, (V_i^n \mathbf{v} \leftrightarrow \varphi)$, if V_i^n doesn't occur in φ
(λI) $\quad \varphi[\mathbf{t}/\mathbf{v}] \vdash \{\lambda \mathbf{v}.\, \varphi\}(\mathbf{t})$
(λE) $\quad \{\lambda \mathbf{v}.\, \varphi\}(\mathbf{t}) \vdash \varphi[\mathbf{t}/\mathbf{v}]$

Note that both free and bound predicate variables may occur in φ in (Comprehension), i.e. this is a fully impredicative principle.

[35] Schindler [28, Proposition 25] essentially shows that all theorems of SOL can be translated into *theorems* of a disquotational truth theory. The results in the present section do not only establish that *inferences* can be preserved as well, but also that $\mathsf{UTB}[\tau]$ doesn't license any more inferences than SOL. Moreover, in [28] it hasn't been noticed that essentially no arithmetical or syntax axioms are needed for that result.

4.2 Translation and Truth

Analogously to what we did in Section 3.3, let $\mathcal{L}_2^+$ expand $\mathcal{L}_2$ with a denumerable stock of fresh predicate constants $Q_1^n, Q_2^n, \ldots$, for each arity n. Similarly, let $\mathcal{L}_\mathrm{T}^+$ extend $\mathcal{L}_\mathrm{T}$ by $Q_1^n, Q_2^n, \ldots$[36] We now provide a translation τ of $\mathcal{L}_2^+$ into $\mathcal{L}_\mathrm{T}^+$ and formulate a corresponding disquotational truth theory $\mathsf{UTB}[\tau]$ that will allow us to simulate the inferential resources of SOL.

In order to do this, let us (recursively) split the variables $v_1, \ldots, v_n$ of $\mathcal{L}_\mathrm{T}^+$ into infinite sequences $x_1^1, x_2^1, \ldots, x_1^2, x_2^2, \ldots$ and $y_1, y_2, \ldots$. Variables belonging to all sequences but the last one are called 'p-variables', as they are the variables to which we map the predicate variables of $\mathcal{L}_2^+$.

Given our fixed coding, by Kleene's Recursion Theorem we obtain a pr function $\tau : \mathcal{L}_2^+ \to \mathcal{L}_\mathrm{T}^+$ that translates the expressions of $\mathcal{L}_2^+$ into $\mathcal{L}_\mathrm{T}^+$ as follows:

$$\tau(t) = \begin{cases} y_i, & \text{if } t := v_i \\ c, & \text{if } t := c \end{cases}$$

and

$$\tau(\varphi) = \begin{cases} F(\tau(t_1), \ldots, \tau(t_n)), & \text{if } \varphi := F(t_1, \ldots, t_n) \\ \mathrm{Tsub}(x_i^n, \langle \tau(t_1), \ldots, \tau(t_n) \rangle), & \text{if } \varphi := V_i^n(t_1, \ldots, t_n) \\ \mathrm{Tsub}(\ulcorner \tau(\psi) \urcorner, \langle \tau(t_1), \ldots, \tau(t_n) \rangle), & \text{if } \varphi := \{\lambda \mathbf{v}.\, \psi\}(t_1, \ldots, t_n) \\ \neg \tau(\psi), & \text{if } \varphi := \neg \psi \\ \tau(\psi) \wedge \tau(\chi), & \text{if } \varphi := \psi \wedge \chi \\ \forall y_i\, (\mathrm{D}(y_i) \to \tau(\psi)), & \text{if } \varphi := \forall v_i\, \psi \\ \forall x_i^n\, (\mathrm{Trsl}_\tau^n(x_i^n) \to \tau(\psi)), & \text{if } \varphi := \forall V_i^n\, \psi \end{cases}$$

Again, D is a monadic predicate of $\mathcal{L}_\mathrm{T}$ intended to express the domain of discourse of $\mathcal{L}$ (cf. Section 2) and $\mathrm{Trsl}_\tau^n(v)$ abbreviates the formula

$$\exists u\, (\mathrm{For}_{\mathcal{L}_2^+}(u) \wedge \varphi_\tau(u, v) \wedge \mathrm{nvff}(v) = \overline{n})$$

of $\mathcal{L}_{\mathsf{PA}}$, where $\varphi_\tau(u, v)$ expresses τ in $\mathcal{L}_{\mathsf{PA}}$, $\mathrm{nvff}(v)$ is a function symbol for the function that maps the code of each formula of $\mathcal{L}_\mathrm{T}^+$ to the number of non-p-variables that occur free in it, $\mathrm{For}_{\mathcal{L}_2^+}(v)$ defines the set of formulae of $\mathcal{L}_2^+$ without free occurrences of predicate variables, and u is a fresh variable. In other words, $\mathrm{Trsl}_\tau^n(v)$ is true exactly of all formulae of $\mathcal{L}_\mathrm{T}^+$ in the range of τ with just n free variables, none of which are p-variables. In the remainder of this section, let us refer to these formulae as 'n-place translations'.

Our translation behaves as we anticipated in the introduction: predications of the form $V_i^n t_1, \ldots, t_n$ are translated as 'the result of substituting $\tau(t_1), \ldots, \tau(t_n)$ for the free variables in x_i^n is true' (i.e. x_i^n is true of $\tau(t_1), \ldots, \tau(t_n)$) and logical structure is otherwise preserved, except quantifiers are relativised appropriately. Moreover, the non-logical vocabulary of $\mathcal{L}_2^+$ remains unmodified. Thus, τ is a natural translation.

[36]See the remarks at the beginning of Section 3.3 for the rationale behind this.

Next, we will formulate a uniformly disquotational truth theory in $\mathcal{L}_{\mathrm{T}}^{+}$ for the formulae of this language that fall in the range of τ. The theory will contain, in addition, principles that allow it to distinguish between syntactic objects and the objects that belong to the first-order domain of $\mathcal{L}_2^{+}$. Again, note that although $\mathcal{L}_{\mathrm{T}}^{+}$ contains the signature of $\mathcal{L}_{\mathsf{PA}}$, no axioms governing these symbols will be included in our truth theory. Since no strong base theory is required for our results, this may be taken to support the idea that the function of the truth predicate is of a quasi-logical nature. On the other hand, we emphasise that all results proved here still hold if our truth theory is expanded with the usual arithmetical axioms and truth-theoretic axioms for the sentences of $\mathcal{L}_{\mathsf{PA}}$ (cf. Corollaries 9 and 10).

To increase readability, we often write x_i instead of x_i^n and $\mathrm{Trsl}_\tau(x_i)$ instead of $\mathrm{Trsl}_\tau^n(x_i^n)$, if the arity is clear from the context. In what follows, let $\mathbf{t}_{\mathbf{k}}^{\mathbf{l}}$ abbreviate $x_{j_1}^{l_1}, \ldots, x_{j_m}^{l_m}$.

Definition 11 $\mathsf{UTB}[\tau]$ is the theory formulated in $\mathcal{L}_{\mathrm{T}}^{+}$ extending FOL with the following axioms:

(D) $\quad \mathrm{D}c$, for each individual constant c of $\mathcal{L}$

(Trsl_τ) $\quad \forall \mathbf{x}_j^l \forall \mathbf{y}_k \left(\bigwedge_{i \leqslant m} \mathrm{Trsl}_\tau(x_{j_i}^{l_i}) \wedge \bigwedge_{i \leqslant n} \mathrm{D}(y_{k_i}) \rightarrow \mathrm{Trsl}_\tau^r(\ulcorner \varphi(\dot{\mathbf{x}}_j^l, \dot{\mathbf{y}}_k)\urcorner\right)$, if φ is in the range of τ and its only free variables are $\mathbf{x}_{\mathbf{j}}^{\mathbf{l}}$, $\mathbf{y}_k$, and $y_{p_1}, \ldots, y_{p_r}$

(UTB_τ) $\quad \forall \mathbf{x}_j \forall \mathbf{y}_k \left(\bigwedge_{i \leqslant m} \mathrm{Trsl}_\tau(x_{j_i}) \wedge \bigwedge_{i \leqslant n} \mathrm{D}(y_{k_i}) \rightarrow (\mathrm{T}\ulcorner \varphi(\dot{\mathbf{x}}_j, \dot{\mathbf{y}}_k)\urcorner \leftrightarrow \varphi(\mathbf{x}_j, \mathbf{y}_k))\right)$, if φ is in the range of τ and its only free variables are $\mathbf{x}_k$ and $\mathbf{y}_k$[37]

(Trsl_τ) establishes that replacing in a translation φ each free l_i-place p-variable with the canonical name of an l_i-place translation and n-many of the $n+r$ free non-p-variables with canonical names for members of the domain of $\mathcal{L}$ results in an r-place translation. In turn, (UTB_τ) provides an instance of disquotation for each formula in the range of τ with parameters from the set of translations and the domain of $\mathcal{L}$.

4.3 Preservation of Inferences

The main result of this section establishes that any admissible inference in SOL from some premises to some conclusion can also be drawn in $\mathsf{UTB}[\tau]$ from the translations of the premises to the translation of the conclusion, together with assumptions fixing the respective range of the variables that occur free in these translations. If only sentences are involved in premises and conclusion, no extra assumptions are needed. Put simply, we show that τ preserves all valid inferences from SOL to $\mathsf{UTB}[\tau]$.

If $\Gamma \subseteq \mathcal{L}_2$, let $\tau(\Gamma) = \{\tau(\psi) \mid \psi \in \Gamma\}$, $\mathrm{Trsl}_\tau(\Gamma) = \{\mathrm{Trsl}_\tau^n(x_i^n) \mid x_i^n$ occurs free in $\tau(\Gamma)\}$ and $\mathrm{D}_\tau(\Gamma) = \{\mathrm{D}(y_i) \mid y_i$ occurs free in $\tau(\Gamma)\}$.

[37] To facilitate proofs, we require that at least one non-p-variable occurs free in φ.

Theorem 4 *Let $\Gamma \subseteq \mathcal{L}_2$ and $\varphi \in \mathcal{L}_2$:*

$$\text{If } \Gamma \vdash_{\mathsf{SOL}} \varphi, \text{ then } \tau(\Gamma), \mathrm{Trsl}_\tau(\Gamma, \varphi), \mathrm{D}_\tau(\Gamma, \varphi) \vdash_{\mathsf{UTB}[\tau]} \tau(\varphi).$$

To prove this result we first need the following two lemmata, which follow easily from the definition of τ by induction on the logical complexity of φ.

Lemma 6 *If t doesn't occur in the scope of a λ-term in $\varphi \in \mathcal{L}_2$,*

$$\tau(\varphi)[\tau(t_1)/\tau(t_2)] = \tau(\varphi[t_1/t_2]).$$

Lemma 7 *If x^n_j doesn't occur in the scope of a λ-term in $\varphi \in \mathcal{L}_2$, then*

$$\tau(\varphi)[x^n_i/x^n_j] = \tau(\varphi[V^n_i/V^n_j]).$$

We prove Theorem 4 by induction on the length of the derivation of φ from Γ in SOL. We assume for now that no λ-terms occur in the derivation. Throughout the proof we only reason in $\mathsf{UTB}[\tau]$, so let '$\vdash$' express derivability in this theory.

If the proof is one-step long, then either $\varphi \in \Gamma$, in which case the result follows trivially, or φ is a logical axiom. In that case, there are two possibilities: (IL), for which the result follows trivially as well, and

(Comprehension) $\varphi := \exists V^n_i \forall \mathbf{v}\, (V^n_i \mathbf{v} \leftrightarrow \psi)$, where V^n_i doesn't occur free in ψ. Let $\mathbf{v}_k$ be the individual variables distinct from $\mathbf{v}$ that occur free in ψ, and let $\mathbf{V}_j$ be the predicate variables distinct from V^n_i that occur free in this formula. The following is the instance of (UTB$_\tau$) for $\tau(\psi)$:

$$\forall \mathbf{x}_j \forall \mathbf{y}_k \forall \mathbf{y}\, (\bigwedge_{i \leqslant m} \mathrm{Trsl}_\tau(x_{j_i}) \wedge \bigwedge_{i \leqslant r} \mathrm{D}(y_{k_i}) \wedge \bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow$$
$$(\mathrm{T}\ulcorner \tau(\psi)(\dot{\mathbf{x}}_j, \dot{\mathbf{y}}_k, \dot{\mathbf{y}})\urcorner \leftrightarrow \tau(\psi)))$$

Since $\mathbf{x}_j$ and $\mathbf{y}_k$ are exactly the variables occurring free in $\tau(\varphi)$, by ($\forall v$E) and ($\rightarrow$E), we have that $\mathrm{Trsl}_\tau(\varphi), \mathrm{D}_\tau(\varphi)$ imply the following:

$$\forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Tsub}(\ulcorner \tau(\psi)(\dot{\mathbf{x}}_j, \dot{\mathbf{y}}_k)\urcorner, \langle \mathbf{y} \rangle) \leftrightarrow \tau(\psi)))$$

By (Trsl$_\tau$),

$$\mathrm{Trsl}^n_\tau(\ulcorner \tau(\psi)(\dot{\mathbf{x}}_j, \dot{\mathbf{y}}_k)\urcorner) \wedge \forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Tsub}(\ulcorner \tau(\psi)(\dot{\mathbf{x}}_j, \dot{\mathbf{y}}_k)\urcorner, \langle \mathbf{y} \rangle) \leftrightarrow \tau(\psi))))$$

and, by ($\exists v$I),

$$\exists x^n_i\, (\mathrm{Trsl}^n_\tau(x^n_i) \wedge \forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Tsub}(x^n_i, \langle \mathbf{y} \rangle) \leftrightarrow \tau(\psi))))$$

This is exactly the translation of the instance of (Comprehension) for ψ, i.e. the translation of φ. Thus, we have that $\mathrm{Trsl}_\tau(\varphi), \mathrm{D}_\tau(\varphi) \vdash \tau(\varphi)$.

Assume that the result holds for every Δ and every ψ if ψ can be derived from Δ in fewer than n steps. We consider the possible ways in which φ could have been inferred from Γ:

($\forall V$E) $\forall V_i^n \, \psi \vdash \psi[V_j^n/V_i^n]$, where $\varphi := \psi[V_j^n/V_i^n]$. We have the following:

$$\begin{array}{rcll} \tau(\Gamma), \mathrm{Trsl}_\tau(\Gamma, \forall V_i^n \, \psi), \mathrm{D}_\tau(\Gamma, \forall V_i^n \, \psi) & \vdash & \tau(\forall V_i^n \, \psi) & \text{I.H.} \\ & \vdash & \forall x_i^n \, (\mathrm{Trsl}_\tau(x_i^n) \to \tau(\psi)) & \text{def. } \tau \\ & \vdash & \mathrm{Trsl}_\tau(x_j^n) \to \tau(\psi)[x_j^n/x_i^n] & (\forall v\text{E}) \end{array}$$

Since all variables occurring free in $\forall V_i^n \, \psi$ are also free in φ, by weakening:

$$\begin{array}{rcll} \tau(\Gamma), \mathrm{Trsl}_\tau(\Gamma, \varphi), \mathrm{D}_\tau(\Gamma, \varphi) & \vdash & \mathrm{Trsl}_\tau(x_j^n) \to \tau(\psi)[x_j^n/x_i^n] & \\ & \vdash & \tau(\psi)[x_j^n/x_i^n] & (\to\text{E}) \\ & \vdash & \tau(\psi[V_j^n/V_i^n]) & \text{Lemma 7} \end{array}$$

The other rules are proved in a similar fashion as in Section 3.4, appealing to Lemma 6 and the axioms of $\mathsf{UTB}[\tau]$. This completes the proof of Theorem 4 for Γ, φ without λ-terms. The latter can be easily dealt with deploying a similar trick as in the proof of Theorem 2 at the end of Section 3.4.

Corollary 6 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_2$:*

$$\textit{If } \Gamma \vdash_{\mathsf{SOL}} \varphi, \textit{ then } \tau(\Gamma) \vdash_{\mathsf{UTB}[\tau]} \tau(\varphi).$$

As before, abusing our terminology slightly, this result could be taken to show that τ is a *relative interpretation* of each second-order theory Γ in the first-order theory $\tau(\Gamma) + \mathsf{UTB}[\tau]$.[38] But note that τ isn't just *any* relative interpretation but one that preserves not only the non-logical vocabulary of $\mathcal{L}_2$ but also maps every second-order formula of the form $V(t_1, \ldots, t_n)$ to a truth-of ascription $\mathrm{Tsub}(x, \langle \tau(t_1), \ldots, \tau(t_n) \rangle)$. In other words, τ is a *natural* relative interpretation, as we argued in the introduction.

4.4 Associated Models

The purpose of this section is to show that disquotational truth for a class of (translations of) expressions does not carry more inferential power than second-order quantification over this class. More specifically, we show that if the inference from some premises to a conclusion is not admissible in SOL, then the inference from the translations of the premises to the translation of the conclusion is not admissible in $\mathsf{UTB}[\tau]$ either. Together with the results we obtained in Section 4.3, the result of this section suggests that disquotational truth and second-order quantification provide *equivalent ways* of achieving one and the same goal, that is, to enable quantification into predicate position over a given class of formulae. Again, it should be noted that this conservativeness result still holds if we expand $\mathsf{UTB}[\tau]$ with the usual arithmetical axioms and disquotation axioms for the formulae of $\mathcal{L}_{\mathsf{PA}}$.

To establish the main result of this section, we first introduce a Henkin (or general) semantics for $\mathcal{L}_2$, define faithful structures, and show how to obtain an $\mathcal{L}_\mathrm{T}^+$-structure

[38] Strictly speaking, no theory formulated in $\mathcal{L}_2$ can be relatively interpreted in a first-order theory, as logical structure cannot be preserved because first-order languages lack second-order quantifiers.

from a countable faithful $\mathcal{L}_2$-structure that preserves truth under τ. Again, we do not claim that our Henkin models are the intended models of $\mathcal{L}_2$; we simply use them to establish our results.

Definition 12 A Henkin $\mathcal{L}_2$-structure $\mathcal{M}$ is a triple $\langle D_{\mathcal{M}}, S_{\mathcal{M}}, \|\cdot\|_{\mathcal{M}^+}\rangle$ where $D_{\mathcal{M}} \neq \varnothing$, $S_{\mathcal{M}} = \{S^n_{\mathcal{M}} \subseteq \wp(D^n_{\mathcal{M}}) \mid n > 0 \text{ and } S^n_{\mathcal{M}} \neq \varnothing\}$, and $\|\cdot\|_{\mathcal{M}^+}$ is a standard interpretation function for the non-logical vocabulary of $\mathcal{L}_2$.

$D_{\mathcal{M}}$ is the universe of discourse for the first-order variables to range over and each $S^n_{\mathcal{M}}$ is the set over which the n-place predicate variables range over. If $\mathcal{M}$ is a Henkin $\mathcal{L}_2$-structure, an assignment g for $\mathcal{M}$ maps each individual variable of $\mathcal{L}_2$ to a member of $D_{\mathcal{M}}$ and each n-place predicate variable to a member of $S^n_{\mathcal{M}}$. Truth in $\mathcal{M}, g$ and truth in $\mathcal{M}$ *simpliciter* are defined as usual.

Definition 13 A Henkin $\mathcal{L}_2$-structure $\mathcal{M}$ is faithful iff every instance of (Comprehension) is true in $\mathcal{M}$.

Definition 14 If $\varphi \in \mathcal{L}_2$ and $\Gamma \subseteq \mathcal{L}_2$, then φ is an SOL-consequence of Γ ($\Gamma \vDash_{\mathsf{SOL}} \varphi$) iff, for every faithful Henkin $\mathcal{L}_2$-structure $\mathcal{M}$ and assignment g for $\mathcal{M}$, if $\mathcal{M}, g \vDash \Gamma$, then $\mathcal{M}, g \vDash \varphi$.

It is well known that SOL is sound and complete with respect to the class of faithful Henkin structures (cf. Shapiro [29]).

In what follows we show how to transform each *countable* faithful Henkin $\mathcal{L}_2$-structure $\mathcal{M}$ into an $\mathcal{L}^+_{\mathrm{T}}$-structure $\mathcal{M}^\tau$ that interprets the arithmetical vocabulary in the standard way, preserves truth under τ, and makes all axioms of $\mathsf{UTB}[\tau]$ true. As in the previous section, a simple modification of our argument shows that we can actually pick $\mathcal{M}^\tau$ so that $\mathcal{M}^\tau \vDash \mathsf{UTB}[\tau] + \mathsf{UTB}[\mathsf{PA}]$.

Here, a Henkin $\mathcal{L}_2$-structure $\mathcal{M}$ is countable iff $D_{\mathcal{M}}$ and $S^n_{\mathcal{M}}$, for each $n > 0$, are all countable.

Lemma 8 *For every countable faithful Henkin $\mathcal{L}_2$-structure $\mathcal{M}$ there's an $\mathcal{L}^+_{\mathrm{T}}$-structure $\mathcal{M}^\tau = \langle D^\tau_{\mathcal{M}}, \|\cdot\|_{\mathcal{M}^\tau}\rangle$ satisfying the following conditions:*

- *$\mathcal{M}^\tau$ extends and expands both $\mathcal{M}$ and $\mathbb{N}$ to $\mathcal{L}^+_{\mathrm{T}}$.*
- *For each assignment g for $\mathcal{M}$ there is an assignment g^τ for $\mathcal{M}^\tau$ s.t. for every $\varphi \in \mathcal{L}_2$:*

$$\mathcal{M}, g \vDash \varphi \text{ iff } \mathcal{M}^\tau, g^\tau \vDash \tau(\varphi).$$

- $\mathcal{M}^\tau \vDash \mathsf{UTB}[\tau]$.

Proof Let $\mathcal{M}$ be given. W.l.o.g. assume that $\mathcal{M}$ is pythagorean, i.e. that $D_{\mathcal{M}} \subseteq \omega$. (If $\mathcal{M}$ is not pythagorean, we simply map it isomorphically to a suitable subset of ω.) We also assume that no element of $D_{\mathcal{M}}$ is the Gödel code of an expression.

We first expand $\mathcal{M}$ to a faithful Henkin $\mathcal{L}_2^+$-structure $\mathcal{M}^+$. Let $q_1^n, q_2^n, \ldots$ be an enumeration of $S^n_{\mathcal{M}}$. For every $n, i > 0$, let

$$\|Q_i^n\|_{\mathcal{M}^+} = q_i^n.$$

Thus, for each $q_i^n \in S^n_{\mathcal{M}}$ there is a predicate constant Q_i^n in $\mathcal{L}_2^+$ whose extension is q_i^n. It is easy to check that $\mathcal{M}^+$ is a faithful Henkin $\mathcal{L}_2^+$-structure.

Let φ be a formula of $\mathcal{L}_2^+$ with exactly $\mathbf{v}$ free, g be an assignment for $\mathcal{M}^+$ and

$$\mathcal{T}_{\mathcal{M}} = \{sub(\#\tau(\varphi), \#\langle d_1, \ldots, d_n\rangle) \mid \mathcal{M}^+, g[d_1 : v_1]\ldots[d_n : v_n] \models \varphi\} \subseteq \omega,$$

where sub is the substitution function for $\mathcal{L}_2^+$ and $\#\langle d_1, \ldots, d_n\rangle$ is the code of the sequence of natural numbers $\langle d_1, \ldots, d_n\rangle$.

$\mathcal{M}^\tau$ is given as follows:

- $D^\tau_{\mathcal{M}} = \omega$
- $\|\epsilon\|^\tau_{\mathcal{M}} = \|\epsilon\|_{\mathbb{N}}$, if ϵ is a non-logical symbol of $\mathscr{L}_{\mathsf{PA}}$
- $\|\epsilon\|_{\mathcal{M}^+} = \|\epsilon\|_{\mathcal{M}^+}$, if ϵ is a non-logical symbol of $\mathcal{L}_2^+$
- $\|\mathrm{D}\|^\tau_{\mathcal{M}} = D_{\mathcal{M}}$
- $\|\mathrm{T}\|^\tau_{\mathcal{M}} = \mathcal{T}_{\mathcal{M}}$

Condition 1 is clearly satisfied. Regarding condition 2, note that every assignment for $\mathcal{M}$ is also an assignment for $\mathcal{M}^+$. So, for each assignment g for $\mathcal{M}$, let the associated assignment g^τ for $\mathcal{M}^\tau$ be defined as follows:

$$g^\tau(v) = \begin{cases} g(v_i), & \text{if } v := y_i \\ \min\{\#\tau(\varphi) \mid g(V_i^n) = \|\varphi\|_{\mathcal{M}^+}\}, & \text{if } v := x_i^n \end{cases}$$

where $\|\varphi\|_{\mathcal{M}^+}$ denotes the "extension" of the formula φ in $\mathcal{M}^+$ – i.e. the set of n-tuples that satisfy φ in $\mathcal{M}^+$. (Note that φ must have n free individual variables.)

The fact that, for each V_i^n, there is a predicate constant Q_j^n s.t. $g(V_i^n) = \|Q_j^n\|_{\mathcal{M}^+}$ guarantees that, for every p-variable x_i^n, $g^\tau(x_i^n) = \#\tau(\psi)$, for some formula $\psi \in \mathcal{L}_2^+$ with n free individual variables.

Using this fact, condition 2 can be verified by induction on the complexity of φ, very much as in Lemma 5(3). That condition 3 is satisfied can be shown in an analogous way to Lemma 5(4). □

Lemma 8 establishes that $\mathsf{UTB}[\tau]$ is consistent relative to SOL. Moreover, Lemma 8 entails that $\mathsf{UTB}[\tau]$ does not afford us more inferences between translations than SOL already allows us to draw between the original formulae.

Theorem 5 *Let $\Gamma \subseteq \mathcal{L}_2$ and $\varphi \in \mathcal{L}_2$:*

$$\textit{If } \mathrm{Trsl}_\tau(\Gamma, \varphi), \mathrm{D}_\tau(\Gamma, \varphi), \tau(\Gamma) \vdash_{\mathsf{UTB}[\tau]} \tau(\varphi), \textit{ then } \Gamma \vdash_{\mathsf{SOL}} \varphi.$$

Proof Assume $\mathrm{Trsl}_\tau(\Gamma, \varphi), \mathrm{D}_\tau(\Gamma, \varphi), \tau(\Gamma) \vdash_{\mathsf{UTB}[\tau]} \tau(\varphi)$. By Soundness of FOL, we have that $\mathsf{UTB}[\tau], \mathrm{Trsl}_\tau(\Gamma, \varphi), \mathrm{D}_\tau(\Gamma, \varphi), \tau(\Gamma) \models \tau(\varphi)$. Let $\mathcal{M}'$ be a faithful Henkin $\mathcal{L}_2$-structure and g an assignment for $\mathcal{M}'$ s.t. $\mathcal{M}', g \models \Gamma$. By the downward Löwenheim-Skolem theorem for SOL (cf. Shapiro [29, Theorem 4.18]) there is a *countable* faithful Henkin model $\mathcal{M}$ that is an elementary submodel of $\mathcal{M}'$ s.t.

$\mathcal{M}, g \vDash \Gamma$. By Lemma 8(3), $\mathcal{M}^\tau, g^\tau \vDash \mathsf{UTB}[\tau]$ and, by Lemma 8(2), $\mathcal{M}^\tau, g^\tau \vDash \tau(\Gamma)$. Moreover, it follows from the definition of g^τ that both $\mathcal{M}^\tau, g^\tau \vDash \mathrm{Trsl}_\tau(\Gamma, \varphi)$ and $\mathcal{M}^\tau, g^\tau \vDash \mathrm{D}_\tau(\Gamma, \varphi)$. Hence, $\mathcal{M}^\tau, g^\tau \vDash \tau(\varphi)$ and, by Lemma 8(2), $\mathcal{M}, g \vDash \varphi$. Since $\mathcal{M}$ is an elementary submodel of $\mathcal{M}'$, also $\mathcal{M}', g \vDash \varphi$. Since $\mathcal{M}'$ was an arbitrary faithful Henkin structure, we have that $\Gamma \vDash_{\mathsf{SOL}} \varphi$. Thus, by the completeness of SOL, $\Gamma \vdash_{\mathsf{SOL}} \varphi$. □

Corollary 7 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_2$:*

$$\textit{If } \tau(\Gamma) \vdash_{\mathsf{UTB}[\tau]} \tau(\varphi), \textit{ then } \Gamma \vdash_{\mathsf{SOL}} \varphi.$$

Put together, Corollaries 6 and 7 entail the following:

Corollary 8 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_2$:*

$$\Gamma \vdash_{\mathsf{SOL}} \varphi \textit{ iff } \tau(\Gamma) \vdash_{\mathsf{UTB}[\tau]} \tau(\varphi).$$

Corollary 7 shows that each theory $\tau(\Gamma) + \mathsf{UTB}[\tau]$ formulated in $\mathcal{L}_\mathrm{T}$ is proof-theoretically conservative over $\Gamma + \mathsf{SOL}$ *modulo* τ, as $\tau(\Gamma) + \mathsf{UTB}[\tau]$ is an extension of $\tau(\Gamma + \mathsf{SOL})$ (cf. Corollary 6). Thus, since τ is a natural translation, we conclude from Corollary 8 that the addition of a uniformly disquotational truth predicate for a given class of sentences affords us the same inferential power as predicate quantifiers (provided that the base language has a canonical term denoting each object in the first-domain of its intended interpretation). The generalising function of a truth predicate governed by this kind of disquotation is no more and no less than the simulation of quantification into predicate position in a first-order setting.

As in the previous section, we note the following two consequences of Lemma 8, where $\mathrm{Tr}(\mathbb{N})$ is the set of sentences of $\mathcal{L}_{\mathsf{PA}}$ that are true in the standard model.

Corollary 9 $\mathsf{UTB}[\tau] + \mathrm{Tr}(\mathbb{N})$ *is conservative over* SOL *modulo* τ.

A fortiori, $\mathsf{UTB}[\tau]$ combined with *any* sound "syntax theory" (e.g. PA) formulated in $\mathscr{L}_{\mathsf{PA}}$ is also conservative over SOL modulo τ. As the following corollary shows, Lemma 8 also entails that, if $\Gamma \subseteq \mathcal{L}_2$ is consistent, then $\tau(\Gamma) + \mathsf{UTB}[\tau]$ together with *any* sound "syntax theory" formulated in $\mathscr{L}_{\mathsf{PA}}$ has a model in which the arithmetical vocabulary is interpreted standardly.

Corollary 10 *If $\Gamma \subseteq \mathcal{L}_2$ is consistent, then $\tau(\Gamma) + \mathsf{UTB}[\tau]$ has a model in which the arithmetical vocabulary is interpreted by the standard model of arithmetic.*

Again, note that a simple modification of our argument shows that we can replace $\mathsf{UTB}[\tau]$ with $\mathsf{UTB}[\tau] + \mathsf{UTB}[\mathsf{PA}]$ in Corollaries 9 and 10, where $\mathsf{UTB}[\mathsf{PA}]$ denotes the result of extending PA with axioms of uniform disquotation for the language of $\mathcal{L}_{\mathsf{PA}}$: one simply needs to add all $\mathcal{L}_{\mathsf{PA}}$-sentences that are true in $\mathbb{N}$ to the extension of the truth predicate $\|\mathrm{T}\|^\tau_{\mathcal{M}}$ in the proof of Lemma 8.

5 Satisfaction and Predicate Quantification

In Section 4 we showed that quantification into predicate position can be mimicked by means of a truth predicate, provided that the base language has a name for each object in its intended first-order domain. This restriction was in place to allow for the formulation of uniform disquotation in terms of a truth predicate. One can define 'true of' in terms of the standard truth predicate with the aid of the substitution function, i.e. $\varphi(\mathbf{u})$ is true of $\mathbf{d}$ if and only if the sentence that results from replacing in $\varphi(\mathbf{u})$ each free variable u_i with a name for d_i is true. Thus, to define truth-of in terms of truth every object in the range of the quantifiers that lead uniform disquotation must have a name.

The goal of this section is to point to results analogous to the ones established in the previous section that apply to base languages without restrictions. For this purpose, instead of using a truth predicate and defining truth-of in terms of it, we appeal directly to a primitive truth-of or satisfaction predicate. Since the proofs are very similar to the ones offered in the previous section, we mostly state our results in this section without proof, to avoid repetition.

There is one difference though. In order to talk about satisfaction of predicates by sequences of objects, we need to be able to talk about sequences of objects. In the previous section, we assumed the domain of objects to be countable. This enabled us to talk about sequences of objects within PA, because (given a coding of those objects by numbers) one can code finite sequence of objects by numbers. This strategy is no longer available as we now allow our domain of objects to be uncountable. Thus, in the present section, we use a fragment of set theory as our syntax theory. The basic set theory BS described in Devlin [30, p. 36] is more than sufficient.[39,40]

Let $\mathcal{L}_{\mathrm{Sat}}$ expand $\mathcal{L}$ with (a fresh copy of) the signature of $\mathscr{L}_{\in}$ (the language of set theory) and a binary predicate Sat, for satisfaction, instead of truth. On its intended meaning, $\mathrm{Sat}(u, v)$ holds of the code u of a formula φ and the code v of a sequence $\langle d_1, \ldots, d_n\rangle$ just in case φ has exactly n free variables and is true of $\mathbf{d}$. As before, we (recursively) split the variables of $\mathcal{L}_{\mathrm{Sat}}$ into p-variables, $x_1^1, x_2^1, \ldots, x_1^2, x_2^2, \ldots$, and non-$p$-variables, $y_1, y_2, \ldots$.

Let $\mathrm{Seq}(v)$ be a formula of $\mathcal{L}_{\in}$ that says that v is a finite sequence. Moreover, let $\ln(v)$ be a function symbol for the function that maps each finite sequence to its length and $u^\frown v$ for the function that takes two sequences and returns the result of concatenating them in order. Finally, let $\mathrm{Sat}^*(u, v)$ be short for

$$\exists w\,(\mathrm{Seq}(w) \wedge \mathrm{Sat}(u, w^\frown v) \ \wedge\ \ln(w) + \ln(v) = \mathrm{nvff}(u)).$$

[39] BS comprises the axioms of Extensionality, Pairing, Union, Fnfinity, Cartesian Products, the Σ_0-Comprehension schema, and a certain induction schema. This theory is weaker than Kripke-Platek set theory.

[40] As with the previous cases, the axioms of the syntax theory are not really required for our results. However, in the present case, we decided to include them into the truth theory as their disentanglement would require more word space than in the previous cases.

In plain words, $\mathrm{Sat}^*(u, v)$ holds just in case there is a sequence w such that, concatenated with v, satisfies u, provided that the length of w plus the length of v is exactly the number of non-p-variables that occur free in u.

As before, we define a translation $\sigma : \mathcal{L}_2 \rightarrow \mathcal{L}_{\mathrm{Sat}}$ and formulate a disquotational theory $\mathsf{UTB}[\sigma]$ in terms of it. Whilst we deploy $\mathrm{Sat}^*(u, v)$ in the definition of σ, the disquotational axioms are given for $\mathrm{Sat}(u, v)$. This is to deal with possible *first-order parameters* in the derivation of the translation of (Comprehension) in $\mathsf{UTB}[\sigma]$ (cf. footnote 41 in Theorem 6).

By Kleene's Recursion Theorem we obtain a pr function $\sigma : \mathcal{L}_2 \rightarrow \mathcal{L}_{\mathrm{Sat}}$ such that

$$\sigma(t) = \begin{cases} y_i, & \text{if } t := v_i \\ c, & \text{if } t := c \end{cases}$$

and

$$\sigma(\varphi) = \begin{cases} F(\sigma(t_1), \ldots, \sigma(t_n)), & \text{if } \varphi := F(t_1, \ldots, t_n) \\ \mathrm{Sat}^*(x_i^n, \langle \sigma(t_1), \ldots, \sigma(t_n) \rangle), & \text{if } \varphi := V_i^n(t_1, \ldots, t_n) \\ \mathrm{Sat}^*(\ulcorner \sigma(\psi) \urcorner, \langle \sigma(t_1), \ldots, \sigma(t_n) \rangle), & \text{if } \varphi := \{\lambda \mathbf{v}.\, \psi\}(t_1, \ldots, t_n) \\ \neg \sigma(\psi), & \text{if } \varphi := \neg \psi \\ \sigma(\psi) \wedge \sigma(\chi), & \text{if } \varphi := \psi \wedge \chi \\ \forall y_i\, (\mathrm{D}(y_i) \rightarrow \sigma(\psi)), & \text{if } \varphi := \forall v_i\, \psi \\ \forall x_i^n\, (\mathrm{Trsl}_\sigma^n(x_i^n) \rightarrow \sigma(\psi)), & \text{if } \varphi := \forall V_i^n\, \psi \end{cases}$$

where $\mathrm{Trsl}_\sigma^n(v)$ is short for $\exists u\, (\mathrm{For}_{\mathcal{L}_2}(u) \wedge \varphi_\sigma(u, v) \wedge \mathrm{nvff}(v) \geqslant \overline{n})$ and $\varphi_\sigma(u, v)$ is a formula of $\mathcal{L}_\in$ expressing σ.

We formulate our theory of uniform satisfaction in terms of σ as follows:

Definition 15 $\mathsf{UTB}[\sigma]$ is the theory formulated in $\mathcal{L}_{\mathrm{Sat}}$ extending BS with (D) and the following axiom scheme:

(UTB_σ) $\forall \mathbf{x}_j \forall \mathbf{y}_k\, (\bigwedge_{i \leqslant m} \mathrm{Trsl}_\sigma(x_{j_i}) \wedge \bigwedge_{i \leqslant n} \mathrm{D}(y_{k_i}) \rightarrow (\mathrm{Sat}(\ulcorner \varphi(\dot{\mathbf{x}}_j, \mathbf{y}_k) \urcorner, \langle \mathbf{y}_k \rangle) \leftrightarrow \varphi(\mathbf{x}_j, \mathbf{y}_k)))$, if φ is in the range of σ and only $\mathbf{x}_j$ and $\mathbf{y}_k$ are free in φ

The results we state in what follows are analogous to those we offered in the previous section. Their proofs are analogous too, with the exception perhaps of Theorem 6, which we sketch briefly.

If $\Gamma \subseteq \mathcal{L}_2$, let $\sigma(\Gamma) = \{\sigma(\psi) \mid \psi \in \Gamma\}$, $\mathrm{Trsl}_\sigma(\Gamma) = \{\mathrm{Trsl}_\sigma^n(x_i^n) \mid x_i^n$ occurs free in $\sigma(\Gamma)\}$ and $\mathrm{D}_\sigma(\Gamma) = \{\mathrm{D}(y_i) \mid y_i$ occurs free in $\sigma(\Gamma)\}$.

Theorem 6 *Let $\Gamma \subseteq \mathcal{L}_2$ and $\varphi \in \mathcal{L}_2$:*

$$\text{If } \Gamma \vdash_{\mathsf{SOL}} \varphi, \text{ then } \sigma(\Gamma), \mathrm{Trsl}_\sigma(\Gamma, \varphi), \mathrm{D}_\sigma(\Gamma, \varphi) \vdash_{\mathsf{UTB}[\sigma]} \sigma(\varphi).$$

Proof The proof is just like the proof of Theorem 4, by induction on the length of the derivation of φ from Γ in SOL. The only difference is in the case where the derivation is one-step long and φ is an instance of (Comprehension), i.e. $\varphi := \exists V_i^n \forall \mathbf{v}\, (V_i^n \mathbf{v} \leftrightarrow \psi)$, where V_i^n doesn't occur free in ψ.

Let $\mathbf{v}_k$ be the individual variables distinct from $\mathbf{v}$ that occur free in ψ, and let $\mathbf{V}_j$ be the predicate variables distinct from V_i^n that occur free in this formula. The following is the instance of (UTB$_\sigma$) for $\sigma(\psi)$:

$$\forall \mathbf{x}_j \forall \mathbf{y}_k \forall \mathbf{y}\, (\bigwedge_{i \leqslant m} \mathrm{Trsl}_\sigma(x_{j_i}) \wedge \bigwedge_{i \leqslant r} \mathrm{D}(y_{k_i}) \;\wedge\; \bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow \\ (\mathrm{Sat}(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner, \langle \mathbf{y}_k \rangle ^\frown \langle \mathbf{y} \rangle) \leftrightarrow \sigma(\psi)))$$

Since $\mathbf{x}_j$ and $\mathbf{y}_k$ are exactly the variables occurring free in $\sigma(\varphi)$, by ($\forall v$E) and ($\rightarrow$E), we have that $\mathrm{Trsl}_\sigma(\varphi), \mathrm{D}_\sigma(\varphi)$ imply the following:[41]

$$\forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Sat}(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner, \langle \mathbf{y}_k \rangle ^\frown \langle \mathbf{y} \rangle) \leftrightarrow \sigma(\psi)))$$

By the defining equations of $+$, Seq, ln, and $^\frown$, and existential weakening:

$$\exists u\, (\mathrm{Seq}(u) \wedge \mathrm{ln}(u) + \mathrm{ln}(\langle \mathbf{y} \rangle) = \mathrm{nvff}(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner) \;\wedge \\ (\forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Sat}(\ulcorner \sigma(\psi)((\dot{\mathbf{x}}_j)\urcorner, u ^\frown \langle \mathbf{y} \rangle) \leftrightarrow \sigma(\psi)))))$$

Since u is not free in the antecedent nor in $\sigma(\psi)$, we can push the quantifier and the first conjunct inside and obtain

$$\forall \mathbf{y}\, (\bigwedge\nolimits_{i \leqslant n} \mathrm{D}(y_i) \;\rightarrow (\exists u\, (\mathrm{Sat}(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner, u ^\frown \langle \mathbf{y} \rangle) \;\wedge \\ \mathrm{Seq}(u) \wedge \mathrm{ln}(u) + \mathrm{ln}(\langle \mathbf{y} \rangle) = \mathrm{nvff}(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner)) \leftrightarrow \sigma(\psi)))$$

which, by definition of Sat*, is just

$$\forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Sat}^*(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner, \langle \mathbf{y} \rangle) \leftrightarrow \sigma(\psi)))$$

By (Trsl$_\sigma$),

$$\mathrm{Trsl}_\sigma^n(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner) \wedge \forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Sat}^*(\ulcorner \sigma(\psi)(\dot{\mathbf{x}}_j)\urcorner, \langle \mathbf{y} \rangle) \leftrightarrow \sigma(\psi)))$$

and, by ($\exists v$I),

$$\exists x_i^n\, (\mathrm{Trsl}_\sigma^n(x_i^n) \wedge \forall \mathbf{y}\, (\bigwedge_{i \leqslant n} \mathrm{D}(y_i) \rightarrow (\mathrm{Sat}^*(x_i^n, \langle \mathbf{y} \rangle) \leftrightarrow \sigma(\psi)))$$

This is exactly the translation of the instance of (Comprehension) for ψ, i.e. the translation of φ. Thus, we have that $\mathrm{Trsl}_\sigma(\varphi), \mathrm{D}_\sigma(\varphi) \vdash \sigma(\varphi)$. □

As before, the following theorem is proved using the method of associated models (i.e. by proving an analogue of Lemma 8).

[41] Note that, if the translation σ were given in terms of Sat instead of Sat*, this principle would not entail $\sigma(\varphi)$. This is due to the presence of the free variables $\mathbf{y}_k$, i.e. the translations of the first-order parameters occurring in φ.

Theorem 7 *Let $\Gamma \subseteq \mathcal{L}_2$ and $\varphi \in \mathcal{L}_2$:*

$$\textit{If } \mathrm{Trsl}_\sigma(\Gamma, \varphi), \mathrm{D}_\sigma(\Gamma, \varphi), \sigma(\Gamma) \vdash_{\mathsf{UTB}[\sigma]} \sigma(\varphi), \ \textit{then} \ \Gamma \vdash_{\mathsf{SOL}} \varphi.$$

Corollary 11 *Let Γ be a set of sentences and φ a sentence of $\mathcal{L}_2$:*

$$\Gamma \vdash_{\mathsf{SOL}} \varphi \ \textit{iff} \ \sigma(\Gamma) \vdash_{\mathsf{UTB}[\sigma]} \sigma(\varphi).$$

From these results we conclude that the addition of a disquotational satisfaction predicate for a given class of formulae affords us the same inferential power as predicate quantifiers, *without* the assumption that the base language has a term denoting each object in the first-order domain of its intended interpretation. Just like in the case of truth, the generalising function of a satisfaction predicate governed by this kind of disquotational principle affords us nothing more than the inferential power of (full impredicative) quantification into predicate position in a first-order setting.

Acknowledgements We are grateful to Ben Blumson, Volker Halbach, Carlo Nicolai, and especially Dan Waxman, who gave us great feedback on more than one version of this paper. We would also like to thank two anonymous referees for their valuable objections and suggestions, which helped us improve the paper enormously. This paper was written while Thomas Schindler received support from the European Research Council within the project *Truth and Semantics* (TRUST, grant agreement no 803684).

References

1. Putnam, H. (1978). Meaning and knowledge. In *Meaning and the moral sciences* (pp. 1–80). London: Routledge.
2. Gupta, A. (1993). A critique of deflationism. *Philosophical Topics*, *21*, 57–81.
3. Halbach, V. (1999). Disquotationalism and infinite conjunctions. *Mind*, *108*, 1–22.
4. Azzouni, J. (2001). Truth via anaphorically unrestricted quantifiers. *Journal of Philosphical Logic*, *30*, 329–354.
5. Grover, D. L., Camp, J. L., & Belnap, N. D. (1975). A Prosentential Theory of Truth. *Philosophical Studies*, *27*, 73–125.
6. Parsons, C. (1983). Sets and classes. In *Mathematics in philosophy* (pp. 209–220). Cornell University Press.
7. Picollo, L., & Schindler, T. (2018). Disquotation and infinite conjunctions. *Erkenntnis*, *83*, 899–928.
8. Picollo, L., & Schindler, T. (2018). Deflationism and the Function of Truth. *Philosophical Perspectives*, *32*, 326–351.
9. Stern, J. (forthcoming). Belief, truth, and ways of believing. In C. Nicolai, & J. Stern (Eds.) *Modes of Truth. The Unified Approach to Modality, Truth, and Paradox*. Routledge.
10. Takeuti, G. (1987). *Proof theory*, second. Amsterdam: North Holland.
11. Halbach, V. (2014). *Axiomatic theories of truth*, 2nd edn. Cambridge: Cambridge University Press.
12. Feferman, S. (1991). Reflecting on incompleteness. *Journal of Symbolic Logic*, *56*, 1–49.
13. Halbach, V. (1996). *Axiomatische wahrheitstheorien, Logica Nova*. Berlin: Akademie Verlag.
14. Halbach, V. (2000). Truth and reduction. *Erkenntnis*, *53*, 97–126.

15. Cantini, A. (1990). A theory of formal truth arithmetically equivalent to ID_1. *Journal of Symbolic Logic*, *55*, 244–259.
16. Friedman, H., & Sheard, M. (1987). An axiomatic approach to self–referential truth. *Annals of Pure and Applied Logic*, *33*, 1–21.
17. Halbach, V. (2009). Reducing compositional to disquotational truth. *Review of Symbolic Logic*, *2*, 786–798.
18. Picollo, L., & Schindler, T. (2021). Is deflationism compatible with compositional and tarskian truth theories?. In C. Nicolai, & J. Stern (Eds.) *Modes of Truth. The Unified Approach to Modality, Truth, and Paradox* (pp. 41–68). Routledge.
19. Shapiro, S. (1998). Proof and Truth: Through Thick and Thin. *Journal of Philosophy*, *95*, 493–521.
20. Ketland, J. (1999). Deflationism and Tarski's paradise. *Mind*, *108*, 69–94.
21. Nicolai, C. (2021). Fix, Express, Quantify. Disquotation after its logic. *Mind*.
22. Fujimoto, K. (2010). Relative truth definability of axiomatic truth theories. *Bulletin of Symbolic Logic*, *16*, 305–344.
23. Monk, D. J. (1976). *Mathematical* L*ogic*. New York: Springer.
24. Tarski, A. (1956). The concept of truth in formalized languages. In *Logic, Semantics, Metamathematics* (pp. 152–278). Oxford: Clarendon Press.
25. Fine, K. (1970). Propositional Quantifiers in Modal Logic. *Theoria*, *36*, 336–346.
26. Kaplan, D. (1970). S5 with Quantifiable Propositional Variables. *Journal of Symbolic Logic*, *35*, 355.
27. Hinman, P. (1978). *Recursion Theoretic Hierarchies*. Berlin: Springer.
28. Schindler, T. (2017). Some notes on truth and comprehension. *Journal of Philosophical Logic*, 47, 449–479.
29. Shapiro, S. (1991). *Foundations without foundationalism: A case for second-order logic*. New York: Oxford University Press.
30. Devlin, K. (1984). *Constructibility*. Berlin: Springer.