Logic and **L**ogical **P**hilosophy
Volume 32 (2023), 495–510
DOI: 10.12775/LLP.2023.012

Sebastian G. W. Speitel

Logical Constants and Arithmetical Forms

Abstract. This paper reflects on the limits of logical form set by a novel criterion of logicality proposed in (Bonnay and Speitel, 2021). The interest stems from the fact that the delineation of logical terms according to the criterion exceeds the boundaries of standard first-order logic. Among 'novel' logical terms is the quantifier "there are infinitely many". Since the structure of the natural numbers is categorically characterisable in a language including this quantifier we ask: does this imply that arithmetical forms have been reduced to logical forms? And, in general, what other conditions need to be satisfied for a form to qualify as "fully logical"? We survey answers to these questions.

Keywords: logical constants; logical form; criterion of logicality; formality

1. Introduction

The notion of logical consequence relies on the idea of *logical form.* Sentences and arguments possess such forms and it is they that are responsible for the particular type of necessary truth-transmission that grounds logical truth and consequence. The logical form of an argument or sentence is determined by a special class of expressions, the logical constants of the relevant language. What is a logical constant, however, and thus which forms are logical, has proven difficult to satisfactorily answer.

Bonnay and Speitel (2021) outlined an account according to which a constant should count as part of the logical lexicon if it possesses a denotation of a certain kind and its inferential behaviour fixes this denotation in the right way. The resulting criterion rendered several novel constants logical and thus expanded the collection of logical forms

Special Issue: Perspectives on Logical Form. Guest Editor: Pavel Arazim

Published online June 23, 2023

grounding claims of 'logically following from'. This paper reflects on the expanded boundaries of logic that resulted from applications of the criterion. In particular, it focuses on the impact the quantifier "there are infinitely many" has on the difference between logical and arithmetical forms.

The structure of the paper is as follows: Section 2 introduces the idea of logical form in more detail and, very cursorily, situates the criterion proposed in (Bonnay and Speitel, 2021) in the general theoretical landscape. Section 3 reflects on the idea that logicality must be closed under definability and investigates the underlying notion of definability at work in such a claim. Section 4 then considers the relationship between intuitively mathematical notions that have been moved closer to logical notions by the criterion and the accompanying demand of closure under definability. Finally, Section 5 concludes with a short set of observations.

2. Logical Form(s)

The notion of *logical form* is essential for an account of logical validity: an argument is *logically valid* iff all arguments of the same logical form are truth-preserving. A sentence is *logically valid* (i.e., *logically true*) iff every sentence of the same logical form is true. Sentences and arguments are logically true/valid in virtue of their *form* alone, independent of any 'material content', i.e., independently of what they might be about.

Logic's formality, its reliance on truth and truth-transmission in virtue of form alone, is meant to account for its great generality and topic-neutrality: since logic 'does not care' what its statements are about, since it does not distinguish between different topics and fields of inquiry, it is *topic-neutral*. Since its pronouncements apply to all domains and fields of knowledge equally, its truths possess a high, discipline-transcendent degree of *generality*.

The generality and topic-neutrality of logical statements finds expression in the fact that the non-logical vocabulary occurs *non-essentially* in the statements of a logical language. The truth of a particular logical truth, such as, for example, $p \vee \neg p$, is independent of the occurrence of any particular p — semantic ascent is needed to express the logical law in full generality as $\varphi \vee \neg\varphi$, where φ ranges over all object-language sentences.

This non-essentiality of the non-logical expressions can be captured in a variety of ways — either by treating expressions of this type as fully schematic, i.e., replaceable by any other expression of the same grammatical category, or through permitting any re-interpretation consistent with their type in the evaluation of a statement's status. Regardless, though, what determines the form of a statement are those expressions that occur non-schematically, whose meanings are held fixed in the evaluation of the truth or falsity of the statements in which they occur.

To capture the notion of *logical* consequence, of necessary truth-preservation in virtue of form alone, the choice of which expressions are to be held fixed cannot be arbitrary:

> Underlying our whole construction is the division of all terms of the language discussed into logical and extra-logical. This division is certainly not quite arbitrary. If, for example, we were to include among the extra-logical signs the implication sign, or the universal quantifier, then our definition of the concept of logical consequence would lead to results which obviously contradict ordinary usage.
>
> (Tarski, 1983, 419)

Rather, it must be such that the truths rendered logical according to the relevant choice of constants are purely formal, general, topic neutral. The idea that a choice of logical constants, determining a notion of logical form, must be consonant with the resulting relation of logical consequence being formal, general, topic-neutral, etc., has given rise to a search for *criteria of logicality*, both proof- and model-theoretic, delineating a class of expressions ensuring these properties of the relation of consequence they ground (see MacFarlane, 2015, for overview).

Among criteria attempting to delineate the class of logical expressions of a language a certain type of approach has reached almost canonical status: invariance criteria (see Tarski, 1986; Sher, 1991; Feferman, 1999; Bonnay, 2008; Griffiths and Paseau, 2022). Common to these accounts is the idea that what makes logical truths and consequences formal, and therefore accounts for their generality and topic-neutrality, is that the logical constants do not distinguish the identity of objects. In the context of a Tarskian model-theoretic apparatus this idea can be implemented by means of an *invariance-constraint*, ensuring the insensitivity of the model-theoretic objects serving as denotations for the logical symbols to the identity of individuals (see MacFarlane, 2000, for an investigation of the notion of *formality* in logic).

Let M be a domain and $\pi\colon M \to M$ a function from M to M. π is a *permutation* if π is *one-to-one* and *onto*. An object o from the type-hierarchy over M such as, for example, the universal quantifier-on-M, i.e., $\forall^M = \{M\}$, is *permutation-invariant* if $\pi[o] = o$, i.e., if the action of π on M leaves o undisturbed. TARSKI'S THESIS is the claim that a notion is logical iff it is permutation-invariant.[1]

TARSKI'S THESIS remains highly local, assessing the logicality of a notion on the basis of its behaviour over each domain separately. This has made it vulnerable to complaints that it allows non-logical elements to re-enter the picture due to lack of cross-domain comparisons (see, e.g., McGee, 1996). Based on considerations independent of those in (Tarski, 1986) and incorporating elements from the mathematical treatment of logical notions (see Mostowski, 1957; Lindström, 1966), Sher advanced[2] what has become known as the TARSKI-SHER THESIS, the claim that a constant is logical iff it denotes a *bijection-invariant* object.[3] This criterion adopts a more global perspective on logical notions, thereby circumventing many of the shortcomings affecting TARSKI'S THESIS.[4]

Despite evading many of the objections that plagued TARSKI'S THESIS due to its highly local character, the TARSKI-SHER THESIS has not been immune to criticism. A complaint that has frequently been directed against it concerns its encompassing nature: by rendering many more notions logical than traditionally counted as logical it *overgenerates*, "assimilat[ing] logic to mathematics, more specifically to set theory" (Feferman, 1999, 37) (see, esp., Feferman, 1999; Bonnay, 2008). Logic, the criticism continues, thereby oversteps its bounds and loses one of its essential characteristics, its topic-neutrality, as it now includes mathematical and, more specifically, numerical content.

I won't assess the force and motivation of this objection here.[5] What I want to point out are two possible ways to address the perceived shortcoming of bijection-invariance in an invariance-framework. The first consists in *modifying* the invariance relation to reduce the number of no-

[1] The view originates in (Tarski, 1986), though the picture is much more complicated than suggested here.

[2] Starting with (Sher, 1991) and developed further in subsequent works. See (Sher, 2022) for a recent account.

[3] A *bijection* $\beta\colon M \to N$ is a *one-to-one* and *onto* function between M and N.

[4] I am oversimplifying both the motivations and the precise details of the TARSKI-SHER THESIS here, see (Sher, 1991, 2016) for detailed treatments.

[5] But see (Sher, 2016; Griffiths and Paseau, 2022) for responses.

tions rendered logical and to thereby minimize reliance on 'problematic' mathematical content.[6] Another consists in *supplementing* invariance-constraints with further conditions. These can either pertain to further demands on the structure of the model-theoretic denotations themselves[7] or to demands on the way logical symbols come to denote logical denotations.

An approach of the latter kind was advanced in (Bonnay and Speitel, 2021), inspired by ideas from (Feferman, 2015; Bonnay and Westerståhl, 2016). There, the idea was that a logical constant was such that (i) it possesses a logical, i.e., bijection-invariant, interpretation; and (ii) among all *acceptable* denotations the inferential features of the constant under consideration determined a unique one.

A denotation was deemed *acceptable* if it was (a) type-appropriate, (b) logical, i.e., bijection-invariant, and (c) consistent with the inferential features of the constant under consideration. The *inferential features* of a constant were taken to be captured by means of a consequence relation $\vdash$ which encodes the relevant inferential facts of a constant c. A denotation $\mathcal{C}$ was then consistent with the inferential features of c iff interpreting c by $\mathcal{C}$ made $\vdash$ sound w.r.t. the model-theoretic consequence relation yielded by $\mathcal{C}$. In reference to (Carnap, 1943) a notion satisfying the criterion was termed *Carnap-categorical*.

Quantifiers rendered logical by the criterion include the usual universal and existential quantifiers of first-order logic,[8] but also the non-first-order definable quantifiers "there are (in)finitely many" ($\mathcal{Q}_0^M = \{A \subseteq M \mid \aleph_0 \leq |A|\}$ and $\mathcal{Q}_{\text{fin}}^M = \{A \subseteq M \mid |A| < \aleph_0\}$).[9] On the other hand, the quantifier "there are uncountably many" ($\mathcal{Q}_1^M = \{A \subseteq M \mid \aleph_1 \leq |A|\}$) was excluded from the class of logical notions (see Bonnay and Speitel, 2021, for details).

3. Logical Constants and Closure under Definability

It is a commonplace to assume that the class of logical notions is closed under definability. For how, it is asked, could non-logical elements enter

[6] See, e.g., (Feferman, 1999; Bonnay, 2008) for approaches of this type.

[7] Such as, e.g., Feferman's (2010) *absoluteness-constraint*.

[8] Based on a result by Bonnay and Westerståhl (2016).

[9] The Carnap-categoricity of the quantifier "there are infinitely many" was first observed by D. Westerståhl and is stated and generalized in (Speitel, 2020).

when combining logical notions by means of logical constructions?

> Operators which are definable in a purely logical manner are logical. We just do not see how a non-logical element could creep in the logical elements of the definition and make the defined operator non logical. (Bonnay, 2008, p. 50)

Relatedly, McGee (1996) takes the fact that permutation-invariant operations are definable in terms of the simple operations of a highly infinitary language, all of which are 'intuitively' logical, to count in favour of considering the former notions logical themselves.[10] Feferman (1999), on the other hand, motivated by extrinsic reasons concerning the limits of logic, ultimately sees need to restrict the permissible types of definition that generate novel logical operations from old.

The contrast between McGee's permissible and Feferman's restrictive take on which definitions are 'logicality-preserving' points to a potential issue in understanding the claim that logicality is closed under definability. For one can agree on the claim while still reaching different judgements about the extent of the class of logical notions not just because of a different starting point but also because one may disagree on what counts as an admissible definition in the first place.

Bonnay (2008), for example, adopts a notion of *explicit definability* as an appropriate standard for preserving logicality. A notion $\mathcal{C}$ is *explicitly definable* in terms of notions $\mathcal{K}_1, \ldots, \mathcal{K}_n$ if there exists a sentence $\varphi_{\mathcal{C}}$ (of appropriate type) of a language $\mathcal{L}$ whose logical symbols are interpreted by $\mathcal{K}_1, \ldots, \mathcal{K}_n$, s.t. for all $\mathcal{M}$, $\mathcal{M} \models \varphi_{\mathcal{C}}$ iff $\mathcal{M} \in \mathcal{C}$. The quantifier $\exists_{\geq 3} = \{\langle M, P^{\mathcal{M}}\rangle \mid P^{\mathcal{M}} \subseteq M \text{ and } 3 \leq |P^{\mathcal{M}}|\}$, for example, can be explicitly defined in FOL by the sentence $\exists x \exists y \exists z(Px \wedge Py \wedge Pz \wedge x \neq y \wedge x \neq z \wedge y \neq z)$. If there is a formula φ of $\mathcal{L}$, s.t. $Mod(\varphi) = \{\mathcal{M} \mid \mathcal{M} \models \varphi\} = \mathcal{C}$ we say that $\mathcal{C}$ is an *elementary class* (is *EC*) in $\mathcal{L}$.

It is worth examining in more detail why definability is usually taken to preserve logicality. The basic intuition underlying the pronouncement that logicality is closed under definability appears to be that, in an important sense, one is not adding anything one did not already have before. Maybe more things are made explicit, given a concrete symbol to play the role that was previously played by a more complex expression, but this sort of move seems to be more or less notational. Definability, on this understanding, thus merely brings things out, makes things ex-

[10] See (Kennedy and Väänänen, 2021) for refinements of this result and (Bonnay and Engström, 2018) for a more systematic investigation.

plicit, that were already there, but does not introduce anything new or unexpected in the process. It allows us to abbreviate and make concrete things that were implicitly sanctioned but for which we were lacking distinguished symbols.

It is useful to contrast the notion of *explicit* definability with the notion of *implicit* definability. The crucial difference between the two consists of the fact that the sentence(s) of an implicit definition draw from a symbol set that already includes a symbol for what is being defined. One is therefore not saying that the new, defined notion behaves like a particular sentence from the 'old' alphabet, — the new notion is not systematically eliminable, — but, rather, that the use of a symbol for the implicitly defined notion fully fixes what this symbol is supposed to mean in the context of the old vocabulary. That, in other words, the old vocabulary was already sufficient to 'pin down' a further notion even if this notion is not equivalent to anything expressible solely in terms of the old vocabulary alone. Nothing more is needed to 'single out' the novel notion than the interaction of a novel symbol with the notions of the old vocabulary. We are thus not introducing anything new 'from the outside', but merely using resources already available and accepted to constrain a novel meaning through the way the vocabulary interacts and 'carves up' logical space. If the notions of the old vocabulary are logical, it is hard to see how they could carve out content in this way that is not also logical.

We say that a notion[11] $\mathcal{C}$ is *implicitly defined* over a language $\mathcal{L}_c$ that includes an uninterpreted symbol c of a type appropriate to $\mathcal{C}$, if there is a sentence φ of $\mathcal{L}_c$, possibly including the symbol c, s.t. for all models $\mathcal{M}$, if $\langle \mathcal{M}, \mathcal{C}_{\mathcal{M}} \rangle \models \varphi$ and $\langle \mathcal{M}, \mathcal{C}^*_{\mathcal{M}} \rangle \models \varphi$, then $\mathcal{C} = \mathcal{C}^*$.[12] In other words, relative to φ there is only one possible interpretation of c.

An important category of implicitly defined notions is comprised of the *basic* logical notions, i.e., those taken as given in a language $\mathcal{L}$. Since they are the ones doing the (explicit) defining, they cannot themselves be explicitly defined in terms of something more basic and so their interpretation will have to be fixed by their use, or so the story goes.[13] There are several issues connected with the idea that the use of a sym-

[11] We use 'notion' here in a slightly informal sense, indicating a model-theoretic object that captures, expresses or represents a particular concept. E.g., the object $\{\langle M, A\rangle \mid A \subseteq M \text{ and } \aleph_0 \leq |A|\}$ expresses or captures the concept of 'infinitely many'.

[12] Note that since $\mathcal{C}$ is to serve as a *logical* notion it must be defined for all models $\mathcal{M}$. We denote by $\mathcal{C}_{\mathcal{M}}$ the interpretation of c in $\mathcal{M}$.

[13] See (Murzi and Steinberger, 2017) for an overview of positions of this type.

bol for the logical notions is sufficient to implicitly define them which we do not wish to downplay, but which we will nevertheless ignore in the following.[14] For our purposes here it is important to note that the so-called implicit definitions of the logical constants are different from the characterization of implicit definability we provided above in that what is usually taken to do the implicit defining in this case are *rules of inference*.[15] Conjunction, for example, is often taken to be implicitly defined by the following rules of inference or, equivalently, by the consequence relation presented by them: (i) $\varphi, \psi \models \varphi \wedge \psi$; (ii) $\varphi \wedge \psi \models \varphi$; and (iii) $\varphi \wedge \psi \models \psi$. On this picture, the operation of conjunction, the semantic value of '$\wedge$', is the unique object that makes the inferences (i)–(iii) valid. Similarly for the usual other constants.[16]

Implicit definability, then, is not specifically tied to single sentences. Neither is the justification for the fact that explicit definition in terms of logical notions does not introduce any non-logical content, for that matter. For if a single sentence explicitly defining a novel logical notion does not introduce any non-logical content in doing so, it is difficult to see how a collection of such sentences could. If to explicitly define something amounts to merely introducing a symbol for something that was, in some shape, already present in the language in terms of which the 'novel' notion is defined, then the above characterisation should be extended to encompass, in the very least, notions definable not just by single sentences, but also by entire theories.[17] Concretely, this means that a notion should also qualify as explicitly definable if it is an EC_Δ-class, i.e., if it is of the form $Mod(\Delta) = \{\mathcal{M} \mid \mathcal{M} \models \delta \text{ for all } \delta \in \Delta\}$ for a (possibly infinite) set of sentences Δ.

The quantifier "there are infinitely many", $\mathcal{Q}_0 = \{\mathcal{M} = \langle M, A\rangle \mid A \subseteq M \text{ and } \aleph_0 \leq |A|\}$ is an EC_Δ class: let $\Delta = \{\exists_{\geq n} xPx \mid n \in \mathbb{N}\}$,

[14] The most prominent of which, in the current context, is the above mentioned *Carnap Categoricity Problem*, the underdetermination of the semantic values of the usual logical constants by their usual rules, see (Carnap, 1943).

[15] The idea that rules of inference are implicit definitions has a long history, starting with a remark by Gentzen (1935).

[16] The story is not as simple for the other constants due to a pervasive underdetermination problem discovered by Carnap (1943) and referred to above. However, modifications to the notion of inference can be made so as to reduce this type of underdetermination. See, e.g., (Carnap, 1943) and (Rumfitt, 2000). We will ignore difficulties arising from this issue in the following.

[17] Further generalizations that we will not discuss here are possible.

where each $\exists_{\geq n} xPx$ designates the first-order expressible sentence saying "there are at least n-many x, s.t. Px". Then $\mathcal{Q}_0 = Mod(\Delta)$. Hence, if the above observations are correct, and if the notions of FOL are accepted as logical, no non-logical content is introduced in defining the quantifier $\mathcal{Q}_0$.[18] Furthermore, it can be shown that $\mathcal{Q}_0$ is implicitly defined by the following two patterns of inference:[19]

(a) $\Delta \models Q_0 xPx$
(b) $Q_0 xPx \models \varphi$ for all $\varphi \in \Delta$

This quantifier is therefore, in the terminology of Bonnay and Speitel (2021), *Carnap-categorical*.[20] The above is, in fact, an instance of a much broader phenomenon: *any* notion that is EC_Δ can be implicitly defined analogously to $\mathcal{Q}_0$.[21]

Hence, a conception of logicality building on the notion of *Carnap-categoricity* allows one to make explicit what was already implicit. It renders notions logical in virtue of being (generalized) definable in terms of logical notions, based on the plausible idea that no non-logical content can be introduced this way.

4. Arithmetical Forms

Accepting a principle of closure under definability for logical notions motivates the admission of $\mathcal{Q}_0$, and the forms it gives rise to, as logical. Of course, this also means that everything that is definable in terms of this 'novel' notion ought to qualify as logical as well. This is what we turn to now.

It is well-known that the natural number structure, $\mathbb{N}$, is categorically characterizable in the language of FOL $+ \{Q_0\}$, whereas every theory of

[18] As an anonymous reviewer points out, $\mathcal{Q}_0$ can be seen as expressing a property that several first-order schemas, namely those that are made true only on infinite domains, have in common.

[19] The satisfaction clause for a (type $\langle 1 \rangle$) generalized quantifier $\mathcal{Q}$ is as follows: $\langle \mathcal{M}, P^{\mathcal{M}} \rangle \models QxPx$ iff $\langle \mathcal{M}, P^{\mathcal{M}} \rangle \in \mathcal{Q}$. See, e.g., (Westerståhl, 2019) for details.

[20] Cf. note 9: this fact was first observed by Dag Westerståhl and is stated and generalized in (Speitel, 2020).

[21] See (Speitel, 2020) for proof and details. In fact, Carnap-categoricity is preserved beyond EC_Δ-definability and is closed under further operations that take us out of the class of EC_Δ-definable notions; see (Speitel, 2020) for details.

arithmetic in FOL alone admits non-standard models.[22] That is, the theory of Peano-arithmetic (PA) plus the sentence

$$(*)\ \forall x \neg Q_0 y (y < x)$$

expressing that every element has only finitely many predecessors suffices to uniquely determine the (isomorphism-type of the) intended structure $\mathbb{N}$. Since this structure is uniquely determinable in this way one might wonder whether arithmetical forms have, therefore, been rendered logical. Whether, after all, the logicist project of reducing arithmetic to logic has, in an extended sense, been successful.

Such an assessment of the situation is of course premature since the arithmetical forms in question contain, in addition to logical notions, also arithmetic-specific vocabulary which does not by itself qualify as logical. Thus, the possibility of categorically characterizing $\mathbb{N}$ appears to be perfectly consistent with the claim that not all arithmetical forms are reducible to logical forms.

But the worry associated with the above observation can be formulated in a more subtle way. We consider the language $\mathcal{L}(Q_1, \dots, Q_n)$ of FOL extended with additional quantifier symbols $Q_1, \dots, Q_n$ and interpretations $\mathcal{Q}_1, \dots, \mathcal{Q}_n$. Let $\mathcal{Q}^*$ be a new quantifier of type $\langle k_1, \dots, k_n \rangle$. The quantifier $\mathcal{Q}^*$ is (explicitly) $\mathcal{L}(Q_1, \dots, Q_n)$-definable if there exists an $\mathcal{L}(Q_1, \dots, Q_n)$-sentence $\varphi(R_1, \dots, R_n)$ whose only non-logical symbols are relation-symbols $R_1, \dots, R_n$ (with R_i of adicity k_i), s.t. for all models $\mathcal{M} = \langle M, R_1^{\mathcal{M}}, \dots, R_n^{\mathcal{M}} \rangle$:

$$\mathcal{M} \in \mathcal{Q}^* \text{ iff } \mathcal{M} \models \varphi(R_1, \dots, R_n)$$

Or, equivalenty, if $\mathcal{Q}^* = \mathit{Mod}(\varphi(R_1, \dots, R_n))$. One can now show that explicit definability over $\mathcal{L}(Q_1, \dots, Q_n)$ ensures implicit definability in $\mathcal{L}(Q_1, \dots, Q_n, Q^*)$ (and thus that Carnap-categoricity is preserved under explicit definability). To this end, let $\models_{\mathcal{Q}_1, \dots, \mathcal{Q}_n}$ be the model-theoretic consequence relation of $\mathcal{L}(Q_1, \dots, Q_n)$, where Q_i is interpreted by $\mathcal{Q}_i$, and let $\models_{\mathcal{Q}_1, \dots, \mathcal{Q}_n, \mathcal{Q}^*}$ likewise be the model-theoretic consequence relation of $\mathcal{L}(Q_1, \dots, Q_n, Q^*)$ (with Q^* interpreted by $\mathcal{Q}^*$). Then:[23]

[22] See any introduction to model theory.

[23] See (Speitel, 2020) for details and proof. It is important to note that (model-theoretic) consequence relations as used here are thought of in the tradition of *model-theoretic* or *abstract* logics. That means that the consequence relations are best seen as functions from vocabularies to sets of tuples of sets of sentences and sentences, subject to certain conditions, and it makes model-theoretic consequence relations, in

PROPOSITION. Let $\mathcal{Q}^*$ be $\mathcal{L}(Q_1, \dots, Q_n)$-definable. It follows that $\mathcal{Q}^*$ is implicitly defined relative to $\models_{\mathcal{Q}_1,\dots,\mathcal{Q}_n,\mathcal{Q}^*}$.

This proposition is neither surprising nor particularly revealing but it allows for a neat application in the present context. Consider the type $\langle 2\rangle$-quantifier $\mathcal{Q}_{\mathbb{N}} = \{\langle M, R^M\rangle \mid \langle M, R^M\rangle \cong \langle \omega, <\rangle\}$, i.e., the quantifier encoding the predicate "is isomorphic to the (ordering of the) natural numbers". Now, $\langle \omega, <\rangle$ is categorically described by the following set of sentences:

(i) "R is a linear order", i.e., R is irreflexive, transitive, and connected.
(ii) $\exists x \forall y \neg Ryx$ – "R has a left-minimal element".
(iii) $\forall x \exists y Rxy$ – "R is right-unbounded".
(iv) $\forall x \neg Q_0 y Ryx$ – "every element has finitely many R-predecessors".

Let $\varphi(R)$ be the conjunction of (i)–(iv). Then $\varphi(R)$ $\mathcal{L}(Q_0)$-defines $\mathcal{Q}_{\mathbb{N}}$, i.e., $\mathcal{Q}_{\mathbb{N}} = Mod(\varphi(R))$.

As argued above, $\mathcal{Q}_0$ should be considered logical by the principle of closure under definability. In terms of $\mathcal{Q}_0$ we can then explicitly define $\mathcal{Q}_{\mathbb{N}}$. By the above PROPOSITION, moreover, $\mathcal{Q}_{\mathbb{N}}$ is implicitly defined relative to $\models_{\mathcal{Q}_0,\mathcal{Q}_{\mathbb{N}}}$. Thus, there appear to be reasons to consider the predicate "is isomorphic to the (ordering of the) natural numbers" a *logical* predicate. What should be made of this?

A first reaction, just as above, could be to insist that $\mathcal{Q}_{\mathbb{N}}$ cannot be considered fully logical on the basis of its explicit definition. This is so, it might further be claimed, because the notion used in explicitly defining $\mathcal{Q}_{\mathbb{N}}$, the notion given by $\varphi(R)$, is not itself a logical notion – the constraints imposed by $\varphi(R)$ on the structure of R exceed the limits of logic. Thus, $\mathcal{Q}_{\mathbb{N}}$ has not been defined solely in terms of logical notions, undermining its claim to logicality.

Its implicit definability, however, is more difficult to dismiss. Since we are concerned with implicit definability relative to a consequence relation, rather than by means of a specific sentence/set of sentences, implicit definitions in our sense are able to abstract away from specific vocabularies and (explicit) definability facts. Reliance on a potentially non-logical notion ensuring explicit definability thus becomes less substantial, or at least less transparent.

---

an important sense, language-independent. See (Ebbinghaus, 2017, Definition 1.1.1) for the general idea.

Nonetheless, it might be said that in the application of the proposition to the case of $\mathcal{Q}_{\mathbb{N}}$ we still, clandestinely, relied on the availability of a linear order to establish the implicit definability of $\mathcal{Q}_{\mathbb{N}}$. Thus, there still seems to be some dependency on a non-logical notion potentially affecting its logical status. The question then becomes whether this dependency is of the right kind to establish the non-logicality of $\mathcal{Q}_{\mathbb{N}}$. I don't think that it is; at least, not straightforwardly so.

Implicit definability relative to a consequence relation entails a certain language transcendence: no matter the particular vocabulary, theories and sentences considered, the relevant notion must remain consistent with the patterns licensed by the consequence relation (over that vocabulary). It is, ultimately, the explicit definability that grounds the implicit definability of $\mathcal{Q}_{\mathbb{N}}$, but it is $\mathcal{Q}_{\mathbb{N}}$'s implicit definability that might be taken to ground its logicality. Thus, in the justification of $\mathcal{Q}_{\mathbb{N}}$'s logicality there appears to be no overt reliance on any putatively non-logical concept, but only on an entire consequence relation independently of any specific vocabulary.

We might be accused of hiding certain commitments of the consequence relation in implicitly defining $\mathcal{Q}_{\mathbb{N}}$ here, but note that while explicit definability is sufficient for implicit definability, the reverse need not necessarily hold. Basing logicality-assessments on implicit definability facts is thus, in general, a weaker requirement. What ultimately does the work in *allowing us to show* the implicit definability of $\mathcal{Q}_{\mathbb{N}}$ relative to $\models_{\mathcal{Q}_0,\mathcal{Q}_{\mathbb{N}}}$ is its explicit definability by $\varphi(R)$. But all its implicit definability relative to $\models_{\mathcal{Q}_0,\mathcal{Q}_{\mathbb{N}}}$ allows us to infer is that there is *some* pattern of inference which constrains possible meanings for $Q_{\mathbb{N}}$ tightly enough to uniquely determine $\mathcal{Q}_{\mathbb{N}}$. Without the additional assumption that the mere possibility of being determined in virtue of a non-logical feature (such as that of a (particular) linear order) results in undermining the logical status of an implicitly defined notion, disqualifying $\mathcal{Q}_{\mathbb{N}}$ as non-logical on this basis appears unmotivated.

The upshot of the previous discussion is this: if the principle of closure under definability for logical notions extends to implicit definability relative to a consequence relation, analogous to how the implicit definability of the standard logical constants is often conceived, $\mathcal{Q}_{\mathbb{N}}$ should be deemed logical. However, this leaves room for views on which $\mathcal{Q}_{\mathbb{N}}$ is *less* logical than, say, $\mathcal{Q}_0$ and $\exists$.

This is the case because the implicit definability of $\mathcal{Q}_{\mathbb{N}}$ requires the presence of $\mathcal{Q}_0$ in the language. Without $\mathcal{Q}_0$ it would be impossible to characterize the kind of order ultimately responsible for fixing the semantic value of $Q_{\mathbb{N}}$. Thus, $\mathcal{Q}_{\mathbb{N}}$ depends on $\mathcal{Q}_0$ — it cannot be implicitly defined relative to a consequence relation over the language $\mathcal{L}(Q_{\mathbb{N}})$ alone. $\exists$, on the other hand, is more logical than both $\mathcal{Q}_{\mathbb{N}}$ and $\mathcal{Q}_0$ since it is implicitly definable over a consequence relation involving neither of the other notions (see Bonnay and Westerståhl, 2016), whereas $\mathcal{Q}_0$ (and thus also $\mathcal{Q}_{\mathbb{N}}$) requires $\exists$ for a statement of the patterns of inference rendering it implicitly definable. Notions can thus be compared w.r.t. their logicality, depending on whether they require one another for their implicit definability.

The picture that emerges is this: if implicit definability is a mark of the logical, gradations in terms of logicality are possible.[24] A logical form is, on this view, more basic than another if the non-schematic notions it involves require 'fewer' other non-schematic logical notions to establish their implicit definability in the extended sense outlined above.

5. Concluding Remarks

I will conclude with a brief summary and some observations on the preceding. The paper began by outlining a *basic principle of logicality*, a criterion delineating a class of expressions as logical and as thereby grounding the logical forms that determine the extent of the notion of logical consequence. This was the criterion formulated and defended in (Bonnay and Speitel, 2021) which rendered several notions not traditionally counted as belonging to the logical lexicon logical. I then considered the idea that the class of logical notions must be closed under definability and argued that the notion of definability featuring in this principle ought to be generalized based on the motivations underlying it. This, ultimately, led us to assess a particular candidate, $\mathcal{Q}_{\mathbb{N}}$, for logicality.

Here, we were presented with three options:

(a) *Accept* that $\mathcal{Q}_{\mathbb{N}}$ is logical on the basis of falling within the purview of the closure under definability constraint and therewith that a certain type of arithmetical content is logical.

(b) *Reject* that $\mathcal{Q}_{\mathbb{N}}$ is logical. This could be done on a variety of grounds: on the one hand, the basic principle of logicality operative might be further strengthened through additional requirements, such

[24] Cf. also (Sagi, 2018) for a graded notion of logicality.

as one that disallows non-compact inference patterns in the legitimate determination of a logical denotation. On the other hand, it might be argued that the principle of closure under definability should only be accepted in a much more restricted form. Implicit definability relative to a consequence relation is, on this view, not good enough since it might obscure the real reason or basis that grounds the apparent logicality of a notion (e.g., its definability by means of a non-logical notion) and might thereby permit non-logical elements to enter into the (allegedly) definitional process.

(c) Lastly, one could take a more differentiated view and adopt a *graded conception of logicality*, according to which logical notions can be more or less logical when compared with others. A notion would qualify as less logical than another if it depended on that other notion for its implicit definability. The existential quantifier, on this view, would be more logical than $\mathcal{Q}_0$, as the latter requires a consequence relation over a language possessing existential quantification for its implicit definability, and similarly for the pair $\mathcal{Q}_0$, $\mathcal{Q}_{\mathbb{N}}$.

I hope to take up a more careful assessment of these three options in future work.

Acknowledgments. I would like to thank audiences at the workshop *What Are Logical Forms (Good For)?*, organized by the Department of Logic, Institute of Philosophy, Czech Academy of Sciences, and an anonymous reviewer for helpful feedback and discussion of the contents of this paper. I would like to further thank Dag Westerståhl for the original observation of the Carnap-categoricity of $\mathcal{Q}_0$ and Denis Bonnay for discussion and development of the criterion of logicality on which this paper is based. Lastly, I would like to thank Gila Sher for comments and feedback on material that formed the basis of this article.

References

Bonnay, D., 2008, "Logicality and invariance", *Bulletin of Symbolic Logic* 14 (1): 29–68. DOI: 10.2178/bsl/1208358843

Bonnay, D., and F. Engström, 2018, "Invariance and definability, with and without equality", *Notre Dame Journal of Formal Logic* 59 (1): 109–133. DOI: 10.1215/00294527-2017-0020

Bonnay, D., and S. G. W. Speitel, 2021, "The Ways of Logicality: Invariance and Categoricity", pages 55–80 in G. Sagi and J. Woods (eds.), *The Semantic Conception of Logic: Essays on Consequence, Invariance, and Meaning*, Cambridge: Cambridge University Press. DOI: 10.1017/9781108524919.004

Bonnay, D., and D. Westerståhl, 2016, "Compositionality solves Carnap's problem", *Erkenntnis* 81: 721–739. DOI: 10.1007/s10670-015-9764-8

Carnap, R., 1943, *Formalization of Logic*, Harvard University Press.

Ebbinghaus, H.-D., 2017, "Extended logics – the general framework", pages 25–76 in J. Barwise and S. Feferman (eds.) *Model-Theoretic Logics*, Cambridge: Cambridge University Press. DOI: 10.1017/9781316717158.005

Feferman, S., 1999, "Logic, logics, and logicism", *Notre Dame Journal of Formal Logic* 40 (1): 31–54. DOI: 10.1305/ndjfl/1039096304

Feferman, S., 2010, "Set-theoretical invariance criteria for logicality", *Notre Dame Journal of Formal Logic* 51 (1): 3–20. DOI: 10.1215/00294527-2010-002

Feferman, S., 2015, "Which quantifiers are logical? A combined semantical and inferential criterion", pages 19–30 in A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthese Library (vol. 373), Springer. DOI: 10.1007/978-3-319-18362-6_2

Gentzen, G., 1935, "Untersuchungen über das logische Schließen I", *Mathematische Zeitschrift* 39: 176–210. DOI: 10.1007/BF01201353

Griffiths, O., and A. Paseau, 2022, *One True Logic. A Monist Manifesto*, Oxford University Press. DOI: 10.1093/oso/9780198829713.002.0003

Kennedy, J., and J. Väänänen, 2021, "Logicality and model classes", *Bulletin of Symbolic Logic* 27 (4): 385–414. DOI: 10.1017/bsl.2021.42

Lindström, P., 1966, "First order predicate logic with generalized quantifiers", *Theoria* 32 (3): 186–195. DOI: 10.1111/j.1755-2567.1966.tb00600.x

MacFarlane, J., 2000, "What does it mean to say that logic is formal?", PhD thesis, University of Pittsburgh.

MacFarlane, J., 2015, "Logical constants", in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2017/entries/logical-constants/

McGee, V., 1997, "Logical operations", *Journal of Philosophical Logic* 25 (6): 567–580. DOI: 10.1007/BF00265253

Mostowski, A., 1957, "On a generalization of quantifiers", *Fundamenta Mathematicae* 44 (1): 12–36. DOI: 10.2307/2964414

Murzi, J., and F. Steinberger, 2017, "Inferentialism", pages 197–224 in B. Hale, C. Wright and A. Miller (eds.), *Blackwell Companion to Philosophy of Language*, second edition, Wiley Blackwell. DOI: 10.1002/9781118972090.ch9

Rumfitt, I., 2000, " 'Yes' and 'No' ", *Mind* 109 (436): 781–823. DOI: 10.1093/mind/109.436.781

Sagi, G., 2018, "Logicality and meaning", *The Review of Symbolic Logic* 11 (1): 133–159. DOI: 10.1017/S1755020317000247

Sher, G., 1991, *The Bounds of Logic: A Generalized Viewpoint*, MIT Press.

Sher, G., 2016, *Epistemic Friction. An Essay on Knowledge, Truth, and Logic*, Oxford University Press. DOI: 10.1093/acprof:oso/9780198768685.001.0001

Sher, G., 2022, *Logical Consequence*, Elements in Philosophy and Logic, Cambridge: Cambridge University Press. DOI: 10.1017/9781108981668

Speitel, S. G. W., 2020, "Logical constants between inference and reference. An essay in the philosophy of logic", UC San Diego Electronic Theses and Dissertations. https://escholarship.org/uc/item/7256r138

Tarski, A., 1983, "On the concept of logical consequence", pages 409–421 in J. Corcoran (ed.), *Logic, Semantics, Metamathematics*, second edition, Hacket Publishing Company.

Tarski, A., 1986, "What are logical notions", *History and Philosophy of Logic* 7: 143–154. DOI: 10.1080/01445348608837096

Westerståhl, D., 2019, "Generalized quantifiers", in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2019/entries/generalized-quantifiers/

SEBASTIAN G.W. SPEITEL
Institute of Philosophy
University of Bonn
Bonn, Germany
sgwspeitel@uni-bonn.de