INVARIANCE AND LOGICALITY IN PERSPECTIVE*

Gila Sher

Although the invariance criterion of logicality first emerged as a criterion of a largely mathematical interest (Mostowski 1957, Lindström 1966, Tarski 1966), it has developed into a criterion of considerable philosophical significance. As a philosophical criterion, invariance has been studied and developed from several perspectives. Two of these are the natural-language perspective and the theoretical-foundational perspective, centered on logic's role in knowledge. My own work (Sher 1991 to 2016) has focused on the second perspective. I have argued that the invariance criterion of logicality makes important contributions to the development of a theoretical foundation for logic focused on its contribution to knowledge – a dual, normative-descriptive foundation centered on (i) the veridicality of logic and (ii) its strong modal force. Those who focus on the natural-language perspective concentrate on the descriptive adequacy of this criterion for the study of natural language. Here we have on the one hand philosophers and linguists who study the criterion's contributions to linguistic semantics (see Peters & Westerståhl 2006 and references there). On the other hand, there are critics of the criterion who base their criticisms on its purported linguistic and intuitive inadequacy (see, e.g., Hanson 1997, Gómez-Torrente 2002, McFarlane 2005/2015, and Woods 2016). Thus, Woods opens his nuanced criticism by saying:

> I argue that in order to apply the most common type of criteria for logicality, invariance criteria, to *natural language*, we need to [require] both invariance of content ... and invariance of character ... . If we do not require this, then old objections ... suitably modified, demonstrate that content invariant expressions can display *intuitive* marks of *non-logicality*. [2016: 778, my emphases]

---

* I would like to thank the participants in the conference "Model Theory: Philosophy, Mathematics, and Language" (Munich Center for Mathematical Philosophy, LMU, 2017), and in particular Gil Sagi and Jack Woods for very helpful comments.

These critics commonly focus on natural-language inferences whose logical validity is allegedly sanctioned by the invariance criterion but challenged by speakers' intuitions (either raw or theory-laden intuitions). Some criticisms are directed at the prevalent version of the invariance criterion, while others are directed at the very idea of an invariance criterion. Still others are directed at the more general idea of a precise, systematic criterion of, or necessary-and-sufficient condition for, logicality, regardless of whether it involves invariance. Among the latter, some opt for a purely pragmatist approach to logicality.

Naturally, there is room for misunderstandings between philosophers who evaluate the invariance criterion of logicality on different grounds and from different perspectives. In particular, there is room for misunderstandings between (i) those who evaluate this criterion on theoretical grounds and those who evaluate it on intuitive grounds, and (ii) those who evaluate it from the point of view of its contribution to a philosophical foundation of logic focused on logic's veridicality and role in knowledge and those who evaluate it from the point of view of its descriptive adequacy with regard to natural language. In this paper I will try to remove a few misunderstandings concerning the theoretical-foundational perspective on the invariance criterion of logicality. To avoid repetition, I will focus on certain aspects of invariance that I have not expanded on in the past as well as on certain points concerning the theoretical approach to invariance and logicality that have led to misunderstandings. I hope that the clarification of these points will help alleviate the tensions between the theoretical-foundational approach to logicality and the natural-linguistic approach.

1. The General Idea of Invariance

Invariance in general is a relation of the form "X is invariant under all variations Y" (where "variations" can be understood as "changes", "transformations", "replacements", and similar expressions, and "Y" can be read as "in Y", "of Y","of type Y", "of type Y in Z", etc.). Invariance, in this general sense, is a very fruitful notion. Three examples (on different levels) of claims involving an invariance relation, taken from logic, mathematics, and physics, are:

(1) A sentence is logically true iff (if and only if) its truth is invariant (preserved) under all replacements of one model by another.

(2) The different geometries can be characterized in terms of the transformations of space under which their concepts are invariant.

(3) The laws of physics are invariant under all changes of inertial frames of reference.

The first example is a reformulation of the standard semantic (model-theoretic) definition of logical truth. Spelled out in more detail, it says that a sentence is logically true iff it is true (in the/a model representing the actual world or even just true in some model) and its truth is preserved under all variations in models (replacements of any model by another). The second example is based on Klein's 1872 *Erlangen* program of classifying geometries and explaining the relations between them in terms of the transformations of space under which their characteristic notions are invariant. Thus, the notions of "rigid-body" geometry are invariant under all transformations of space that preserve *distance* between points, while the notions of Euclidean geometry are invariant under all transformations of space that preserve *ratios of distance* between points. Since the latter condition involves invariance under more transformations than the former, Euclidean geometry is more general than rigid-body geometry. One of the most general geometries is topology, whose notions are invariant under all transformations that preserve closeness (open sets). And in principle, geometry G_1 is *more general* than geometry G_2 iff the notions of G_1 are *invariant* under *more transformations* of space than the notions of G_2. The third example is taken from special relativity, whose laws are invariant under all variations in inertial frames of reference.

What does invariance mean? What is its significance? What does it amount to? We may say that when X is invariant under all variations Y, X "does not notice", "does not pay attention to", "is blind to" changes in Y, "is immune" to changes of type Y, or "is not affected" by changes in Y and "cannot be undermined" by discoveries concerning features that vary from one Y to another. Thus, if we regard models as portraying all possible ways the world could have been (in some relevant sense of "possible"), then we may say that logical truths "do not pay attention" to whether the world is as portrayed by one model or by any other. In a similar way, the property of being a Euclidean triangle "is blind" to transformations of space that change distances between points so long that they preserve ratios of distances. (The image of any Euclidean triangle under such transformations is also a Euclidean triangle.) The laws of physics "are immune to changes" in inertial frames, or "are not affected" ("cannot be undermined") by discoveries concerning the distinctive features of given inertial frameworks, those that vary from one inertial framework to another. And so on.

Accordingly, one of the ways in which invariance is highly significant is that the *stronger* the invariance conditions a given notion satisfies (or the characteristic notions of a given field satisfy), the *stronger* or more *stable* the notion (field of knowledge) is, in relevant respects. "Stronger", in the cases we consider here, can be characterized as follows: Invariance condition I_1 is stronger than invariance condition I_2 if the class of transformations associated with I_1 properly includes the class of transformations associated with I_2. But if the stronger the invariance conditions satisfied by X, the stronger (in relevant respects) X is, then it is to be expected that *if X satisfies especially strong invariance conditions, X is especially strong (in relevant respects)*. It would thus not be surprising if we could explain the fact that, and the way in which, logical truths and consequences are *stronger* than other truths and consequences based on their strong invariance. And as we shall see below, it is indeed possible to explain the exceptional modal force of logical truths and consequences based on the fact that they, and/or some of their constituents, satisfy certain *especially strong invariance conditions*.

2. The Theoretical Challenges of Logicality and Veridicality

I. *The Logicality Challenge*

The logicality challenge is the challenge of establishing the theoretical viability of a system of genuine *logical consequences* and explaining how it might be structured. Philosophers may have less and more demanding conceptions of *genuine* logical consequence. Here I am interested in a relatively *demanding* conception, associated with logic's role in knowledge. This role, as I understand it, is to devise a powerful, universal method or system for extending knowledge in any field by moving us from truths − robust, correspondence-like truths − that we may already know to truths (of the same kind) that we may not yet know. In this spirit, I require that a genuine logical consequence satisfy the following strong conditions:

(T) A logical consequence *transmits truth* from premises to conclusion (where *truth* is a demanding notion: correspondence in a broad yet robust sense, rather than mere coherence, pragmatic justification, disquotation, etc.).[1]

(M) The transmission of truth is *guaranteed* with an *especially strong modal force*.

II. *The Veridicality Challenge*

The veridicality challenge is the challenge of truth and justification of the logical theory (system) itself. To be adequate, a logical theory has to say *true* things about logical truths and consequences. It should *not* say that a sentence S follows logically from a set of sentences Γ *unless* S *in fact* follows logically from Γ, i.e., unless the sentences of Γ in fact transmit correspondence-truth – truth *in the world* – to S and do so with an especially strong modal force. It is not sufficient that our intuitions tell us, or give us the impression, that this is the case; this has to *be* the case, and we need to theoretically *justify* the claim that it is the case.

Now, ideally, there would be no need to treat the *veridicality* challenge as a separate

---

[1] (i) I understand "disquotational truth" in this paper as exemplifying the view that truth in general takes into account only facts (such as disquotation) concerning language. I understand "robust" as involving demanding requirements concerning the world (generally, the extra-linguistic world).

(ii) In this broad sense, correspondence is free from its traditional association with the naive and simplistic idea of copy, mirror-image, or direct isomorphism. For further explanation of this broad (yet robust) conception of correspondence see, e.g., Sher (2016).

challenge. It would go without saying that an adequate system of logical consequence satisfying the logicality challenge produces consequences that *truly* or *in fact* transmit truth from premises to conclusion with an especially strong modal force. But in contemporary philosophy, as we have noted above, philosophers sometimes focus on intuitive rather than theoretical justification.[2] So it is important to indicate that this is not sufficient. An adequate account of logicality must *show* that the requisite conditions on logical consequence are in fact satisfied, and this "showing" must be *theoretical* rather than *merely intuitive* in the everyday sense of the word.

The critical question concerning the invariance criterion of logicality, as a theoretical-foundational criterion, is, then, whether it enables us to establish, theoretically, the viability of a system of consequences that affirms all and only patterns of consequence that in fact transmit truth from premises to conclusions with an especially strong type of necessity.

3. Preliminaries

I. *Methodology.* The challenges of logicality and veridicality are foundational challenges, challenges that have to do with fundamental philosophical questions concerning logic. But the attempt to deal systematically with such foundational questions raises methodological problems that have to be treated with care. Traditionally, philosophers assumed that the only methodology for dealing with foundational questions is the *foundationalist* methodology. But the foundationalist methodology makes a theoretical foundation of logic impossible. I have discussed some of the problems it raises and proposed an alternative methodology elsewhere (Sher 2013, 2016), so here I will be very brief. One problem with the foundationalist methodology is its requirement that in giving a foundation for a field of knowledge K we limit our epistemic

---

[2] For example, Hanson rejects "modal and formal accounts" of logicality on the ground that they "fail to satisfy our *intuitions* about logical consequence" (1997: 386, my emphasis). He denies the logicality of a term alleged to be logical by the invariance account on the ground that "it seems *bizarre*", i.e., *counter-intuitive*, "to treat" it as logical (*ibid.:* 392, my emphasis). And so on.

resources to those produced by *more fundamental* fields than K (fields lying lower than K in the foundationalist hierarchy). But no basic field of knowledge can be given a theoretical foundation under these conditions. Since logic is classified by foundationalists as a basic field, this problem applies to logic. In the literature, many philosophers focus on a particular aspect of this problem: due to the basicness of logic, we cannot provide a theoretical foundation for it without *circularity* or *infinite regress*. Since all forms of circularity and infinite regress violate the foundationalist strictures, we cannot provide a theoretical foundation for logic at all.

To investigate logicality theoretically, therefore, we need a different methodology. The methodology I will use here is a *holistic* methodology of a *special kind*, called "foundational holism" (see Sher 2016). This methodology is holistic rather than foundationalist, but it differs from various other types of holism in being *geared toward foundational* studies. Thus, this holistic methodology is *world oriented* rather than coherentist, it emphasizes the *inner complexity of structures* rather than totalities or wholes, and so on. Its holistic nature is reflected in its attentiveness to large and open-ended networks of connections between diverse elements. It recognizes that there are many ways to reach the world cognitively, both on the level of discovery and on the level of justification. In particular, both discovery and justification may exhibit multiple patterns, some hierarchical, others not. Accordingly, not all forms of circularity are forbidden: some occurrences of circularity are innocuous, and some are even constructive. The paradigmatic metaphor of foundational holism is Neurath's Boat. In trying to meet the logicality challenge we go back and forth between various kinds of considerations on various levels, using whatever resources are available to us at the moment.

II. *Philosophical Theory and Mathematical Background-Theory.* In studying logicality theoretically from a philosophical point of view we are faced with a special problem. On the one hand, we aim at a *philosophical* rather than a mathematical account, and in particular, we wish to avoid commitment to any particular mathematical background-theory. On the other hand, using the resources of some mathematical background-theory may have considerable benefits:

expressing philosophical ideas using precise terms-of-art, bringing clear examples and counter-examples, answering questions that are difficult to answer without mathematical resources, and so on. The usefulness of a mathematical background-theory is especially significant in the philosophies of logic and mathematics, due to the formality of the disciplines they study. But using a specific mathematical theory as a background theory might introduce complications. Whereas our philosophical ideas are devoid of problematic mathematical commitments, using the resources of a specific mathematical theory to express them can easily create the false impression that they do carry such commitments. To avoid such false impressions, I prefer to divide my discussion of logicality into two parts. In Sher (2016) I started by formulating and explaining my ideas philosophically, without using mathematical terms-of-art. Once this account was completed, I presented a *precisified* version of the account, helping myself to the resources of a specific mathematical theory, ZFC. Throughout the discussion I stressed that in principle one could use a different mathematical background-theory, with different mathematical commitments, so ZFC's commitments are *not inherent* in the account.

Due to limitations of space I will not be able to be as thorough in separating the two accounts here. But to avoid misunderstandings, it is important to be aware of this point. In particular, it is important to realize that the explanation of invariance and logicality given in the present paper is *philosophical* rather than set-theoretical. It is not committed to ZFC; nor does it carry its commitments.

4. Two Invariance Principles of Logicality.

In the philosophical literature on logicality, talk of *invariance* is usually directed at one use of invariance – demarcation of *logical constants* – and accordingly, at one type of invariance. But in fact, there is another use, and another type, of invariance in logical semantics as well. This invariance principle appears as my first example of general invariance above. It concerns the use of *models* for demarcating *logical truths and consequences*. I will call it "the first invariance principle of logicality", or "the model-theoretic invariance principle" (I-M).

I. *The First Invariance Principle of Logicality: Invariance under (changes in) models (I-M)*

The first invariance principle of logicality underlies the standard semantic definition of logical consequence, whose roots go back to Tarski (1936). Consider a collection Γ of sentences of a given language L and a sentence S of L. The standard semantic definition of logical consequence can be formulated as:

(LC) S is a logical consequence of Γ (in L) iff in every model (for L) in which all the sentences of Γ are true S is also true,

without commitment to a specific mathematical construal of models. To capture the requirement that the truth in question is of a robust kind, i.e., a broadly correspondence-truth, we can reformulate LC as:

(LC') S is a logical consequence of Γ (in L) iff in every model (for L) in which all the sentences of Γ are correspondence-true S is also correspondence-true.

Now, although people rarely think of LC as a definition of logical consequence in terms of *invariance*, the idea of invariance (the same idea as in Section 1 above) is implicit in it. We can make this idea explicit by reformulating LC as *Invariance-under-Models*, I-M:

(I-M) S is a logical consequence of Γ iff the transmission of (correspondence-) truth from Γ to S is *invariant* under all variations in (replacements of) models,

Three questions concerning LC, or its reformulation, I-M, concern language, models, and logical constants:

(a) *Language.* What kind of language is assumed by LC/I-M? Since we are interested in a *theoretical* account of logicality, we need to think of this language, which we may identify with L above, as a *theoretical language*, rather than as a natural language. As a theoretical language, L abstracts from those features of language in general that are deemed irrelevant for understanding logicality.[3]

(b) *Models.* How shall we understand models, philosophically? To capture the conception of logical consequence as transmitting correspondence-truth from sentences to sentences with an especially strong modal force, models have to satisfy certain conditions: (i) models should represent all and only ways the actual world could have been, given a relevant understanding of possibility,[4] (ii) there should be a model that represents the way the world actually is in relevant respects,[5] and (iii) the totality of models should be especially large, i.e., the conception of possibility involved should be especially broad, broader than that of physical and even metaphysical possibility. By focusing on the world – the way it is and the ways it could have been – (i) and (ii) ensure that logical consequence transmits the right kind of truth, namely correspondence-truth (truth-in-the-world), and that the transmission of truth occurs in all relevant situations, actual and counterfactual. (iii) ensures that logical consequences have an especially strong modal force, i.e., the modal force of logic is greater than that of physics and even metaphysics. What the relevant conception of possibility is will become clear shortly.

(c) *Logical Constants.* To achieve transmission of truth and exceptional modal force, logical consequence is dependent on a special feature of sentences. This feature is commonly called "logical form", but in fact it could also be called "logical content". Logical consequence takes into account only the logical form or content of the sentences involved, not their non-logical form/content. The logical form/content of sentences has to do with the identity and

---

[3] In principle, L can be either an extensional or an intensional language. But for reasons explained in Sher (1991), the logical constants of L are extensional.

[4] As we will see below, however, my conception of models as *representational* is subject to constraints that distinguish it from the "representational" conception discussed, and rightly rejected, by Etchemendy (1990). For a more detailed explanation, see Sher (1996).

[5] Some philosophers (e.g., Field 2009) argue that models based on standard set theory as a background theory are incapable of adequately representing the actual world. Whether they are right or wrong, the fact that our conception of models is not tied up to this particular background theory (or, indeed, to any other) exempts it from this argument.

distribution of constants of a certain kind: constants that, due to their special character, support especially strong and universal consequences. Logical consequence holds fixed the content or denotation of these constants while treating the content or denotation of the non-logical constants as variable (in effect, treating these constants as schematic letters or variables). When it comes to models, logical constants have a fixed denotation (content, satisfaction conditions) in all models, while the denotation (content, satisfaction conditions) of the non-logical constants varies (vary) from model to model.[6]

Given the conditions T (transmission of truth) and M (especially strong modal force) on *logical* consequence, it is crucial that we set specific requirements on admissible logical constants. This was already noted by Tarski (1936). If, for example, we treat the material conditional as a non-logical constant, changing its denotation from model to model, Modus Ponens will come out logically invalid. And if we treat "Tarski", "Frege", and "is a logician" as logical constants, "Tarski is a logician; therefore, Frege is a logician" will come out logically valid. Tarski himself did not arrive at any principled criterion for (characterization of, requirements on) logical constants in his 1936 paper, "consider[ing] it quite possible that investigations will bring no positive results in this direction" (*ibid.:* 420). From the present perspective, the challenge is to find a criterion for logical constants that satisfies, and perhaps even maximizes the satisfaction of, the two conditions on logical consequence, T and M.

These considerations leave the theoretical philosopher of logic with three major tasks:

(A) Construct a theoretical criterion for logical constants.

---

[6] (i) See Sher (1991). This amounts to another important constraint on models.

(ii) As explained in Sher (*ibid.*), the fixity of logical constants does not mean that they have the same extension in all models (the extension of the universal quantifier in a model with 8 individuals is a set of 8 individuals, while its extension in a model with 9 individuals is a set of 9 individuals). What it means that their extension is determined for all models in advance, by a fixed principle. (In the case of the universal quantifier, this principle says that its extension in any model is the whole domain of that model).

(iii) For an interesting discussion of the fixity of logical constants in the context of current model theory (the current mathematical theory of models), see Sagi (2018).

(B) Specify a type of possibility suitable for logical consequence (and underlying the totality of models).

(C) Explain how (A) and (B) satisfy T and M.

In other words, the theoretical philosopher's task is find, or develop, a theoretical criterion for (or characterization of) logical constants and identify a type of possibility that, together, render LC/I-M an adequate criterion of logical consequence. This brings us to the second invariance principle of logicality and the discussion of formal possibility.

II. *The Second Invariance Principle of Logicality: Invariance under isomorphisms (I-I)*

The second invariance principle of logicality is a criterion for logical constants. This criterion is often referred to as the "invariance under isomorphisms" criterion (I-I).[7]

The Invariance-under-Isomorphisms criterion for logical constants (I-I) that I will discuss here is the criterion developed in Sher (1991) based on earlier mathematical criteria due to Mostowski (1957) and Lindström (1966).[8] I-I has two parts, an objectual part and a linguistic part. The latter concerns the constants of the language L, the former − their objectual denotations, and more generally, objects (in particular, extra-linguistic objects).

A. *Objectual Part of I-I.* The objectual part of I-I applies to objects of a certain kind. One can think of these objects in various ways. Given the present goal, I prefer to think of the relevant objects as properties, where properties include proper properties, relations, and functions of any

---

[7] It is also often referred to as the "invariance under bijections" criterion and the "Tarski-Sher thesis". A related criterion is the invariance-under-automorphisms/permutations criterion (Mostowski 1957, Tarski 1966), but depending how one understands it, this criterion is significantly different from, and inferior to, the invariance-under-isomorphisms/bijections criterion. See McGee (1996) and Sher (1991, 2016).

[8] The 1991 criterion was developed in the mid-80's, before Tarski's 1966 lecture was published. But it can also be construed as a development of the criterion proposed by Tarski.

level and any arity.[9] I-I divides properties into two types: those that do and those they do not satisfy it. Adherents of I-I regard the former as admissible denotations of logical constants, the latter as inadmissible. The formulations of I-I by Mostowski, Lindström, and Tarski are limited to its objectual part.

B. *Linguistic Part of I-I.*[10] The linguistic part of I-I does two things:

(a) It tells us that a logical constant must denote a property that satisfies the objectual part of I-I.

(b) It sets additional conditions on logical constants, intended to ensure that logical constants satisfying (a) are adequately integrated into a syntactic-semantic system of logical consequence incorporating LC/I-M.

In this paper I will focus on the objectual part of I-I. (For the linguistic entries of I-I see Sher 1991.) In accordance with my second preliminary comment in Section 3, I will offer two versions of I-I: one that is not and one that is couched in a mathematical background-theory. I will call the non-mathematical version of the criterion "Invariance under 1-1 replacements of individuals", and I will use the abbreviation "I-R" for this version. I-R is intended to be understood in a way that does not involve specific mathematical (including set-theoretical) commitments. Depending on context, "I-I" will name either the mathematical version of the criterion or the broader conception (I-R).

The non-specifically-mathematical version of I-I, *invariance-under-replacements-of-individuals*, or I-R, can be presented as follows:

(I-R) An n-place property, $\mathcal{P}$, of level m, is invariant under all 1-1 replacements of individuals iff for any domain of individuals, D, and any argument, β, of $\mathcal{P}$ (in D), β has the property

---

[9] I think of objects in general as divided into individuals (objects of level 0) and properties (objects of level >0). The use of properties in the present discussion does not assume any specific theory of properties, and various theories of properties are compatible with this account. For the purpose of the present discussion we can for the most part disregard current controversies concerning properties.

[10] "Linguistic", here, is a theoretical adjective applicable to languages in the sense of Section 4(I)(a) above.

𝒫 (in D) iff the image of β under any 1-1 replacement ℝ of the individuals in D has the property 𝒫 (in D', the image of D under ℝ).[11]

Consider the 2-place 1st-level property *x-loves-y*.[12] It is quite clear that this property does not satisfy I-R. But the 2-place 1st level property *x=y* does satisfy I-R. Consider the 1-place 2nd-level property P-IS-A-PROPERTY-OF-HUMANS, where P is a 1-place 1st-level property. This property does not satisfy I-R, but the 1-place 2nd-level property P-IS-NON-EMPTY – the existential-quantifier property – does.

The mathematical version of I-R, (I-I below) is thought of as a precisification of I-R:

(I-I) An n-place property, 𝒫, of level m, is invariant under all isomorphisms iff for any domains D, D' and any arguments β, β' of 𝒫 in D,D' respectively: if <D,β> is *isomorphic* to <D',β'>, then β has the property 𝒫 (in D) iff β' has 𝒫 (in D').[13]

As noted above, various variants of (I-I) can be introduced using various background mathematical theories. One version of I-I will use ZFC as its background theory,[14] another may use Russell's theory of types as background theory,[15] and still others may have other mathematical background theories.

Although in the historical order of discovery I-I was prior to I-R, in the order of

---

[11] (i) D is any collection of individuals, actual or counterfactual. Since I-R is not formulated in any specific mathematical background theory, D does not have to be identified as a set, a proper class, or an entity of any other specific mathematical type. For the sake of simplicity we assume that D is non-empty.

(ii) Given a 𝒫 and a D: If 𝒫 is a 1-place 1st-level property, its arguments in D are individuals in D. If 𝒫 is a 2-place 1st-level property, its arguments in D are pairs of individuals in D. And so on. If 𝒫 is a 1-place 2nd-level property of 1-place properties, its arguments in D are 1-place 1st-level properties whose arguments in D are individuals in D). And so on.

[12] I use italics for 1st-level properties and small capital letters for 2nd-level properties.

[13] <D,β> is *isomorphic* to <D',β'> iff there is a bijection *f* from D to D' such that β' is the image of β under *f*.

[14] In this version, D, D' will be proper sets.

[15] In fact, Russellian type-theory was one of the two background theories used by Tarski for his 1966 version of I-I.

philosophical explanation and justification I-R is prior to I-I. This calls for a methodological clarification: My goal in this paper is to explain how the foundational theorist approaches the question of logicality and how invariance enters into her eventual account. To that end, the explanation I provide has the character of a *rational reconstruction* (in a quasi-Carnapian sense). It does not seek to trace the history of the invariance criterion; instead it explains how it is rational to reconstruct it.

I-I as presented so far is, strictly speaking, a criterion for *properties* and *predicates (including quantifiers)*. What about *sentential operators and connectives*? I-I can be generalized to an invariance criterion of logicality for such operators/connectives in several ways. If we assume bivalence, the sentential version of I-I (given in Sher 1991, 2016) coincides with the usual truth-functionality criterion for logical connectives. For the purpose of the present discussion, however, it is sufficient to focus on I-I as a criterion for properties/predicates.

We are now ready to explain why the *Invariance-under-Isomorphisms* criterion is an appropriate criterion for logical constants and to specify the type of possibility that must be represented by models – the models used in logic, which I will call "logical models". This will enable us to explain how the two invariance conditions, I-M (invariance under models, or LC) and I-I, satisfy the two conditions on an adequate notion (system, method) of logical consequence – T (transmission of correspondence-truth) and M (especially strong modal force).

5. Invariance-under-Isomorphisms, Formality, and Modal Force

One of the distinctive characteristics of the invariance-under-isomorphism criterion – a characteristic that distinguishes it from other criteria for, and accounts of, logical constants[16] – is that it captures a certain *especially fruitful philosophical idea.* This idea is *formality*. Not formality in the traditional syntactic sense, or the schematic semantic sense, or the substitutional

---

[16] From Feferman's (1999, 2010) invariance-under-homomorphisms criterion to pragmatist, non-invariance accounts (see below).

semantic sense, but formality in an *objectual* semantic sense. Objects – specifically properties – satisfying I-I are *formal* in this sense; objects that do not satisfy I-I are *not formal* (in this sense). Any constant can be formal in the syntactic, schematic, or substitutional sense, i.e., be treated as a fixed, distinguished element, partaking in the "form" of sentences (see, e.g., Etchemendy 1990). But only constants that denote properties satisfying I-I are *formal* in a sense that is relevant to the two conditions on logical consequence noted above, T (transmission of correspondence-truth) and M (especially strong modal force).

I-I is connected to formality both extensionally and intensionally. Extensionally, properties satisfying I-I are mathematical and all mathematical objects – individuals, properties, and structures – either satisfy I-I or are systematically correlated with properties that satisfy I-I. Among the mathematical properties that satisfy I-I are identity, the 2nd-level Boolean properties corresponding to the standard logical connectives, the existential- and universal-quantifier properties (NON-EMPTINESS, UNIVERSALITY), ONE, TWO, ..., FINITELY MANY, INFINITELY MANY, IS-REFLEXIVE/SYMMETRIC/ TRANSITIVE, IS-WELL-ORDERED, etc. Mathematical structures, such as the structure of the natural numbers, are systematically correlated with quantifier-properties satisfying I-I. Mathematical individuals such as the number 1 are correlated with 2nd-level cardinality properties – ONE, ... – which satisfy I-I. The 1st-level 1-place property *x-is-even* satisfies I-I when construed as a 3rd-level property of 2nd-level cardinality properties, and so on. In contrast, all paradigmatic non-mathematical objects and properties (such as *Archimedes, is-red* and IS-A-PROPERTY-OF-HUMANS) do not satisfy I-I.[17]

Intensionally, I-I captures the idea of formality as *strong structurality*. Take any property $\mathcal{P}$ of any level, any domain D, and any argument β of $\mathcal{P}$ in D. Now take the pair <D,β> and take any pair <D',β'> that has the exactly the *same structure* as <D,β>. Such a structure can be obtained from <D,β> by a 1-1 replacement of the individuals of D, and if it does, then β satisfies

---

[17] To apply I-I to Archimedes, we identify Archimedes with a property, such as *is-Archimedes*. Clearly, this property does not satisfy I-I.

$\mathcal{P}$ in D iff β' satisfies $\mathcal{P}$ in D'. I.e., $\mathcal{P}$ satisfies I-I iff it pays attention *only* to *highly structural features* of its arguments, iff it is blind to all features of its arguments but (some of) their *highly structural* features. Speaking in terms of invariance, we may say that most properties abstract from some features of their arguments, and as such they satisfy some invariance condition and have some degree of invariance. In this sense, they are at least weakly structural. But I-I is an especially strong invariance condition. Paradigmatically biological, physical, and other properties do not satisfy this condition; only highly-structural properties do. Such highly structural properties are *formal*.

One ramification of I-I is that the transmission of truth from premises to conclusion by a logical consequence is due to formal relations between the contents of its premises and the content of its conclusion. Semantically, the transmission of truth is due to formal relations between the truth conditions of it premises and conclusion. Objectually, the transmission of truth is due to formal relations between the situations corresponding to its premises and conclusion, or more precisely, between the formal structures of these situations. For example, the logical consequence

(4) $(\exists x)(Ax \vee Bx)$, $\sim(\exists x)Ax$; therefore: $(\exists x)Bx$

is based on a relation between two formal structures: a structure of a non-empty union of two properties, P_1 and P_2, the first of which, P_1, is empty, and a structure in which the second property, P_2, is non-empty. This relation is itself formal, so (4) is based on a formal relation between two formal structures, or on a formal relation between formal features of the situations that make (or would make) the premises and conclusion of (4) true. It is due to this relation that (4) transmits (correspondence-) truth from its premises to its conclusion.

Another ramification of I-I is that the transmission of truth from premises to conclusion by a logical consequence has an especially strong modal force. This ramification arises from the fact that the invariance associated with the properties denoted by logical constants – invariance under isomorphisms – is connected to an especially strong type of necessity. The connection between invariance under isomorphisms and strong necessity is based on the fact that properties invariant under all isomorphisms cannot distinguish between individuals of any kinds, actual or counterfactual, and therefore the laws governing such properties cannot distinguish between actual-counterfactual individuals of any kind either. Since the space of such actual-counterfactual individuals is especially large (larger than the space of individuals that physical and even metaphysical properties do not distinguish), the actual-counterfactual scope of the laws governing them is especially large. In other words, these laws have an especially strong modal force. Since logical consequences are grounded in such laws, they have an especially strong modal force. This result has two parts: 1. Logical consequences are grounded in laws governing the properties denoted by their logical constants, namely formal properties. 2. Since these properties have an especially strong degree of invariance, their laws – formal laws – hold in an especially large space of possibilities, hence have an especially strong modal force.

We can finally understand the notion of possibility represented by logical models: logical models represent the totality of formal possibilities, namely, all the ways the world could have been when only formal structure is taken into account. This is the reason invariance under all replacements of models is an adequate criterion of logical consequence.

We have seen how the two invariance criteria of logicality, invariance-under-isomorphisms (I-I) and invariance under models (I-M) establish, theoretically, the viability of an adequate system of logical consequence. Elsewhere (Sher 2016 and works mentioned there) I showed that the formality of logical consequences (in the sense of I-I) also explains their other properties: their considerable generality, topic neutrality, basicness, certainty, and normativity, as well as their quasi-apriority.[18]

---

[18] Concerning generality, Bonnay (2008) interprets Tarski as saying that I-I is associated with *utmost generality* rather than with *formality*. But for reasons presented both in Bonnay (*ibid.*) and in Sher (2008), I-I does not really capture the idea of utmost generality. It captures the idea of formality which, in turn, is associated with *considerable*, yet not *utmost*, generality. For discussion see *op. cit.*

This theoretical account of logic employs the foundational holistic methodology. The account is developed in a stage by stage (step by step) manner, going back and forth in a Neurath-boat style. While in earlier stages we did not have sufficient resources for explaining the relation between logical and metaphysical possibility, at this point we do. The degree of invariance of metaphysical properties is smaller than that of formal properties. Hence the space of logical/formal possibilities is greater than that of metaphysical possibilities. Consider the metaphysical impossibility of being all-red and yellow at the same time. This impossibility is not formal. The property of being both all-red and yellow is not invariant under all isomorphisms. That is to say, the combination of being all red and yellow is not ruled out on formal grounds. Formal possibility abstracts from most features of individuals, including color and color relationships. Therefore, an individual that is both all red and yellow is formally possible and as such belongs in the domain of some logical models. There are models that represent individuals that are both all red and yellow, individuals that are both dead and alive, individuals that do not satisfy the regularities of biology or the laws of physics or the principles of metaphysics. This is the reason the scope of logical possibility is broader than that of other types of possibility and the modal force of preserving truth in all (logical) models is exceptionally high.[19]

Many of the alleged counter-examples to I-I neglect the difference between formal possibility and other kinds of possibility, which is crucial for understanding the philosophical significance of both the invariance-under-models criterion (I-M or LC) and the invariance-under-

[19] Regarding the comparison between logical and metaphysical necessity/possibility, however, it is important to note that metaphysics is a highly heterogeneous discipline, dealing, on the one hand, with very basic ontological issues, such as what makes something an object, and on the other hand, with less basic issues, such as causality, free will, observable vs. unobservable objects, physical vs. mathematical objects, abstract vs. concrete objects, and so on, including issues like color incompatibilities. These less basic (but still quite basic) issues occupy a much larger space in contemporary metaphysics than the more basic ones, and my references to metaphysics in this paper concern metaphysical possibilities and impossibilities of the less basic kinds. (I leave the relation between logic and the most basic parts of metaphysics to another paper.)

isomorphisms criterion (I-I). These alleged counter-examples often assume an intuitive or a metaphysical notion of possibility, which is weaker than the notion relevant for I-M and I-I. Therefore, they are not genuine counter-examples. These examples are also usually presented as natural-language examples.

This brings us to the relation between the theoretical, philosophical-foundational, perspective on logicality and the natural-linguistic perspective.

6. The Natural-Linguistic and Foundational Perspectives.

So far we have discussed the two invariance criteria of logicality − invariance-under-models (I-M) and invariance-under-isomorphisms (I-I) − as criteria designed to satisfy theoretical conditions on logical consequence: transmission of (correspondence-) truth (T) and especially strong modal force (M). We have seen that, from this perspective, the combination of the two invariance criteria, I-M and I-I, fares well: it ensures the satisfaction of T and M, thereby establishing the viability, in principle, of an adequate system of logical consequence. How does this combination, and in particular I-I, fare from a natural-linguistic perspective? Is I-I a descriptively adequate criterion of logicality from this perspective?

To answer this question we need, first, to understand what it means. What, exactly, does *descriptive adequacy* amount to in this case? How do we establish it in principle? It is hard to find a detailed answer to these questions in the critical literature on I-I.

Two co-authors who do raise this question are Peters and Westerståhl (2006). Peters and Westerståhl first formulate this question in a way that is similar to our theoretical question, namely, by asking whether I-I is adequate for a *genuine* logical consequence. Next they ask whether the method commonly used in empirical linguistics, namely, the method of consulting speakers' linguistic intuitions, is appropriate for answering this question. It is widely agreed that this method is appropriate for determining grammaticality; the question is whether it is also appropriate for determining validity and logicality. Peters and Westerståhl are skeptical about a positive answer to this question. While linguistic intuitions have been shown to be reliable with respect to grammaticality, it is easy to see that they are *unreliable* in determining validity and logicality. As a result, Peters and Westerståhl give up the attempt to solve the problem of logicality from a natural-linguistic perspective. They take I-I to be a *necessary* condition on logical constants in natural language, but they do not try to determine whether it is a *sufficient* condition, i.e., whether it is an adequate *criterion* of logicality for natural language.

What they do investigate, instead, is whether the invariance-under-isomorphisms criterion, I-I, enables us to better understand linguistic phenomena that are difficult to understand without it. Their answer to this question is positive. They show that and explain how non-standard logical quantifiers sanctioned by I-I enable us to explain phenomena concerning determiners and complex quantifier-structures in natural language. For example, the non-standard monadic logical operator MOST, sanctioned by I-I, explains the behavior of the determiner "most" in sentences such as "Two critics reviewed *most* films"; the polyadic operator MOST ... -AND-MOST ..., sanctioned by I-I, explains *branching-quantifier* structures in natural language such as "Most of the boys in my class and most of the girls in your class have all dated each other"; and so on.

Peters & Westerståhl's approach is reasonable. On the one hand, studying the ways the invariance-under-isomorphisms criterion, I-I, provides new resources for understanding linguistic structures, both in natural language and in artificial languages, makes good sense. But relying on speakers' intuitions to determine validity and logicality does not. Validity and logicality are significantly different from grammaticality, and employing the same method for both requires careful justification.

Most philosophers, however, do not heed Peters & Westerståhl's warning about the use of linguistic intuitions to determine logicality. Such intuitions are widely used by philosophers as grounds for rejecting I-I without any attempt to justify the use of intuition as an arbiter in this case. In addition, some opponents of I-I appeal to views whose relevance to logicality is questionable. Let me explain these points by reference to two alleged counter-examples to I-I due to Gómez-Torrente (2002, 2003): "unicorn" and "male widow".

Gómez-Torrente claims that the properties *is-a-unicorn* and *is-a-male-widow* are empty "in all possible universes" (*Ibid.:* 2002, p. 18). As such, he says, they satisfy the invariance-under-isomorphisms criterion, I-I. Accordingly, the linguistic expressions "is a unicorn" and "is a male widow" come out logical. This, in turn, implies that "There are no unicorns" and "There are no male widows" are logically true. But these sentences "are intuitively not logically true" (*Ibid,:* 2003, p. 204). Hence, according to Gómez-Torrente, I-I is not an adequate criterion of logicality.

I explained why this criticism is incorrect in Sher (2003). But there I focused on the fact that the linguistic expressions "x is a unicorn" and "x is a male widow" do not satisfy the extended, linguistic, version of I-I, spelled out in Sher (1991). Here I would like to focus on the properties *is-a-unicorn* and *is-a-male-widow*. I would like to point out certain assumptions underlying Gómez-Torrente's use of these properties to criticize I-I and explain how these assumptions lead us to think that these properties satisfy I-I when in fact they do not. Let me begin with *male widow.*

Gómez-Torrente claims that *male widow* is empty in "all possible universes". What is the basis for this claim? My understanding is that this claim is based on our ordinary intuitions. But this approach to the issue *neglects* the fact that the notion of possibility involved in both the invariance-under-isomorphisms criterion (I-I) and the invariance-under-models criterion (I-M) is a specific and especially broad notion of possibility, namely, the notion of *formal possibility*, whereas the notion of possibility employed in the claim that *male widow* is empty in all possible universes is a non-specific notion of possibility, one that is usually understood in a way that makes it weaker than formal possibility. This explains why this example cannot be used to undermine I-I. The incompatibility between being male and being a widow is *not a formal incompatibility*. Therefore, it does not rule out the *formal possibility* of situations in which the property *male-widow,* like the property *both-all-red-and-yellow*, is *not* empty. *Male-widow*, then, does *not* satisfy I-I, and "There are no male widows" is *not* true in all logical models, hence does *not* come out logically true on the invariance account of logicality incorporating I-I and I-M.

What about the property *unicorn*? Why would anyone think that this property is empty in all possible universes? The claim that *unicorn* is empty in all possible universes is, if I understand Gómez-Torrente correctly, based not on natural-linguistic intuitions but on a particular philosophical theory that is naturally viewed as belonging to the philosophy of language or to metaphysics, due to Kripke (1972/80). But this theory does not provide an adequate ground for rejecting I-I. First, this theory is not a theory of *formal* possibility/necessity, but a theory of *metaphysical* possibility/necessity, and as such it is irrelevant to I-I. Second, this theory does not really say that *unicorn* is an empty property in all possible universes. It says that *unicorn*, being a mythological-species "property", is, like all other mythological species "properties", *not* a genuine *property*. I will not go into Kripke's reasons for this claim here. But if one accepts his claim, one cannot bring *is-a-unicorn* as a counter-example to I-I, since I-I does not deal with non-properties.

There are other linguistic/intuitive grounds on which some philosophers have tried to deny I-I. For critical discussions of these grounds see, e.g., Paseau (2013), Sagi (2015), and Sher (1991, 2003, 2016).

7. A Pragmatist Approach to Logicality.

A number of philosophers – e.g., Hanson (1997, 2002) and Gómez-Torrente (2002, 2003) – prefer a pragmatist approach to logicality over a theoretical approach. Two main weaknesses of the pragmatist approach to logicality are: (i) its neglect of *veridicality*, and (ii) its neglect of *theoretical explanation*. These, I believe, are pragmatism's main weaknesses in all theoretical branches of knowledge. If, and so long as, we view the search for knowledge as a search for *truth* (in a robust, correspondence, sense), if we require *veridical justification* and *evidence* for theoretical claims, and if we aim at *genuinely explanatory* theories, then we cannot be content with a pragmatist approach to knowledge. In the philosophy of logic, or those parts of the philosophy of logic that are discussed in this paper, the question of truth arises in multiple places and on multiple levels: What should logical consequence transmit from premises to conclusion given its role in knowledge? Is a given claim of logical consequence true? (Does it in fact transmit correspondence truth from premises to conclusion with an especially strong modal force?) What is (are) the source(s) of truth of logical-consequence claims? Is it true that a system of logic based on I-I and I-M satisfies the requirements of transmission of truth and modal force on logical consequence? Does the formality of logic, articulated in terms of I-I, provide a theoretical explanation of the special features of logic – necessity, generality, topic-neutrality, etc.? And so on. All these are theoretical questions of truth and explanation that, in principle, require veridical theoretical answers rather than pragmatic answers.

This is not to say, however, that pragmatic considerations cannot play any role in theoretical knowledge. Where can pragmatic considerations enter into the invariantist account of logicality? They can play a partial role in choosing the overall best background theories for the account. (Such a choice is needed when, e.g., we have no decisive veridical basis for choosing between two candidates for a background theory.) They can play a partial role in deciding which logical system licensed by the invariantist account of logicality to choose in a particular context or given a particular goal. (For example, it is pointless to choose a system that includes high-infinite-cardinality quantifiers if our interest is limited to everyday inferences or even to inferences in physics.) They may be used in deciding on certain details of our system of logical consequence. (For example, the decision whether to limit models to structures with *non-empty* universes.) And so on. But pragmatic considerations should be used alongside, and be balanced by, considerations of veridicality and theoretical explanation, not in lieu of these.

8. Conclusion

Invariance plays a central role in many fields of knowledge. In logic, it plays a central role in a theoretical foundational account of logicality, both on the level of logical constants and on the level of logical consequence. Often, however, the invariance criteria of logicality, and in particular I-I, have been evaluated from other perspectives, and this has led to disagreements based on a misunderstanding of their designated role. In this paper I have tried to put some of these disagreements in perspective. In particular, I have explained the foundational-theoretical perspective on logicality as distinguished from the natural-linguistic intuitive perspective.

The foundational-theoretical perspective starts with a conception of logic's role in the advancement of human knowledge, and proceeds to the requirement that logical consequences transmit (correspondence) truth from sentences to sentences with an especially strong modal force. It shows how the two invariance criteria of logicality, invariance-under-models and invariance-under-isomorphisms, give rise to a logical system that grounds logical consequences in a particular facet of the world, *formal laws*, which have the requisite modal force. A central aspect of this account is the connection between invariance, formality, and modal force. Logical constants represent formal properties, properties that have an especially high degree of invariance and as such do not distinguish between most individuals (including metaphysically possible and impossible individuals). Logical consequences are based on laws governing the relations between such properties, laws that hold in all formally-possible situations, which are represented by the totality of models. As such, their modal force is greater than the modal force of laws and consequences of other disciplines, whose actual-counterfactual scope is narrower.

References

Bonnay, D. 2008. "Logicality and Invariance". *Bulletin of Symbolic Logic* 14: 29-68.

Etchemendy, J. 1990. *The Concept of Logical Consequence.* Cambridge: Harvard.

Feferman, S. 1999. "Logic, Logics, and Logicism". *Notre-Dame Journal of Formal Logic* 40: 31-

54.

----. 2010. “Set-theoretical Invariance Criteria for Logicality”. *Notre Dame Journal of Formal Logic* 51: 3-20.

Field, H. 2009. “What is the Normative Role of Logic?”. *Proceedings of the Aristotelian Society*, Supplementary Vol. 83: 251-68.

Gómez-Torrente, M. 2002. “The Problem of Logical Constants”. *Bulletin of Symbolic Logic* 8: 1-37.

----. 2003. “The ‘Must’ and the ‘Heptahedron’. Remarks on Remarks”. *Theoria* 18: 199-206.

Hanson, W.H. 1997. “The Concept of Logical Consequence”. *Phil. Review* 106: 365-409.

----. 2002. “The Formal-Structural View of Logical Consequence: A Reply to Gila Sher”. *Philosophical Review* 111: 243-58.

Klein, F. 1872. “A comparative review of recent researches in geometry”. Verlag von Andreas Deichert, Erlangen.

Kripke, S. 1972/80. *Naming and Necessity*. Cambridge: Harvard.

Lindström, P. 1966. “First Order Predicate Logic with Generalized Quantifiers”. *Theoria* 32: 186-95.

MacFarlane, J. 2005/15. “Logical Constants”. *Stanford Encyclopedia of Philosophy* (on line).

McGee, V. 1996. “Logical Operations”. *Journal of Philosophical Logic* 25: 567-80.

Mostowski, A. 1957. “On a Generalization of Quantifiers”. *Fund. Mathematicae* 44: 12-36.

Peseau, A. 2013. “The Overgeneration Argument(s): A Succinct Refutation”. *Analysis* 74: 40-47.

Peters, S. & D. Westerståhl. 2006. *Quantifiers in Language and Logic.* Oxford.

Sagi, G. 2015. “The Modal and Epistemic Arguments against the Invariance Criterion for Logical Terms”. *Journal of Philosophy* 112: 159-167.

----. 2018. “Logicality and Meaning”. *Review of Symbolic Logic* 11: 133-59.

Sher, G. 1991. *The Bounds of Logic.* Cambridge: MIT.

----. 1996. “Did Tarski Commit ‘Tarski's Fallacy’?” *Journal of Symbolic Logic* 61: 653-86.

----. 2003. "A Characterization of Logical Constants Is Possible". *Theoria* 18: 189-97.

----. 2008. "Tarski's Thesis". *New Essays on Tarski and Philosophy*, ed. D. Patterson. Oxford: 300-339.

----. 2013. "The Foundational Problem of Logic". *Bulletin of Symbolic Logic* 19: 145-98.

----. 2016. *Epistemic Friction: An Essay on Knowledge, Truth, and Knowledge.* Oxford.

Tarski, A. 1936. "On the Concept of Logical Consequence". In *Logic, Semantics, Metamathematics.* Indianapolis: Hackett 1983: 409-20.

----. 1966. "What Are Logical Notions?" *History and Philosophy of Logic* 7 (1986): 143-54.

Woods, J. 2016. "Characterizing Invariance". *Ergo* 3: 778-807.