Logic and Logical Philosophy (2021)
DOI: 10.12775/LLP.2021.002
Published online: February 19, 2021

Matej Drobňák

Lewisian Naturalness and a new Sceptical Challenge

Abstract. The criterion of naturalness represents David Lewis's attempt to answer some of the sceptical arguments in (meta-) semantics by comparing the naturalness of meaning candidates. Recently, the criterion has been challenged by a new sceptical argument. Williams argues that the criterion cannot rule out the candidates which are not permuted versions of an intended interpretation. He presents such a candidate — the arithmetical interpretation (a specific instantiation of Henkin's model), and he argues that it opens up the possibility of Pythagorean worlds, i.e. the worlds similar to ours in which the arithmetical interpretation is the best candidate for a semantic theory. The aim of this paper is a) to reconsider the general conditions for the applicability of Lewis's criterion of naturalness and b) to show that Williams's new sceptical challenge is based on a problematic assumption that the arithmetical interpretation is independent of fundamental properties and relations. As I show, if the criterion of naturalness is applied properly, it can respond even to the new sceptical challenge.

Keywords: David Lewis; fundamental properties; meaning; metasemantics; naturalness; pythagorean worlds

1. Introduction

How can we determine what expressions mean? What may sound like a simple question turns out to be a hard puzzle when we try to give a systematic account of meaning. One of the prominent approaches that attempts to answer the question is interpretationism [2, 9, 12]. According to the proponents of interpretationism, the way to determine what expressions of a language mean is to create an interpretation (a model) of a language which is in accordance with the use of the expressions/sentences of the language. Broadly speaking, the interpretation is

Received April 3, 2020. *Revised* January 30, 2021

supposed to capture a correlation between the truth values of sentences and states of the world. This is provided by assigning semantic values to sentences or expressions[1] in such a way, that the interpretation will (mostly) match the interpreted sentences with the correct truth values.

However, since the 1960s several philosophers [8] — following [5, 20, 21, 34] have formulated sceptical arguments which indicate that the pattern of use is not enough to determine meaning. The reason is that several different interpretations of a language are always in accordance with the data about the use of the language and so several different meaning candidates[2] are always in accordance with the pattern of use of each particular expression.

An attempt to answer some of those arguments can be found in [13]. David Lewis proposed a criterion of naturalness to decide which meaning candidate for a predicate is the most natural and hence the most eligible. According to Lewis, properties (instantiated by classes of objects) can play the role of semantic values of predicates and some properties are inherently more natural than others. Importantly, the naturalness of properties should be compared by comparing the lengths of their definitions when defined in terms of fundamental properties and relations — the longer definition in fundamental terms a property has, the less natural the property is.[3]

[1] In what follows, I will use the terms 'meaning' and 'semantic value' interchangeably. Notice also that interpretationism does not need to be necessarily linked to any specific approach to semantics. For Lewis, the assigned semantic values were understood as functions from possible worlds to truth values (for sentences) and functions from possible worlds to sets of individuals (for expressions). For Davidson, the assigned semantic values were understood as truth conditions (for sentences). Interpretationism in general is more a methodological approach which states how meaning should be studied than a metaphysical approach which states what meaning is. To avoid a commitment to any particular approach to semantics, I will talk about semantic values as being 'instantiated by classes of objects'.

[2] By 'meaning candidates' I mean the semantic values assigned to the same expression by different interpretations. For example, in accordance with the famous grue paradox, one meaning candidate for the expression 'green' would be instantiated by a class of all green objects and another meaning candidate would be instantiated by a class of all the objects that are green before t, but that are blue otherwise.

[3] In what follows, I will use the expressions 'in terms of fundamental properties and relations' and 'in fundamental terms' interchangeably. By 'fundamental terms' I mean terms used to represent fundamental properties and relations. A detailed explanation of which properties and relations Lewis considered to be fundamental is presented in Section 2.

Lewis's response to the sceptical arguments rests on a dialectical comparison of naturalness of different interpretations/meaning candidates. It is based on the assumption that the alternative interpretations of a language are formulated in terms of permutations of some intended interpretation.[4] In such cases, the permutation clauses needed to create an alternative interpretation extend the length of its definition and so the alternative interpretation can be discarded as less natural than the intended interpretation.

Williams [31] presents a new sceptical challenge which seems to be beyond the force of Lewis's dialectical comparison of naturalness. As Williams shows, there are alternative interpretations of a language which do not rely on permutations. An example is his arithmetical interpretation and its subsequent application in the argument for the existence of Pythagorean worlds. The arithmetical interpretation is a specific instantiation of Henkin's model [6, 7]. In general, Henkin's model provides an interpretation of a language by assigning equivalence classes of constants as semantic values. However, the domain of Henkin's model can be freely chosen and Williams creates his arithmetical interpretation by changing the domain from constants to natural numbers. As a result, Williams's arithmetical interpretation assigns equivalence classes of numbers as semantic values. Since Henkin's model is not a permuted version of some intended interpretation, Williams argues, Lewis's dialectical comparison of naturalness cannot determine whether meaning candidates assigned to expressions by the arithmetical interpretation are less natural than the candidates assigned by an intended interpretation.

A partial aim of this paper is to show that Williams's new sceptical challenge is based on a problematic assumption that the arithmetical interpretation is independent of fundamental properties and relations and because of that it is not an eligible candidate for the comparison of naturalness. On the basis of deficiencies spotted in Williams's argumentation, I will discuss the possibility of comparisons in which a) one meaning candidate/interpretation is not defined in fundamental terms and b) meaning candidates/interpretations are defined in different fundamental terms. The main aim of the paper is to show that Lewis's criterion of naturalness works properly under condition that all the candidates for a comparison of naturalness are defined in terms of the same fundamental

[4] Lewis response is primarily focused on Putnam's model-theoretic argument [19, 20] which is explicitly stated in terms of permutations.

properties and relations. If this "constraint of fundamental uniformity" holds, then we gain a strong tool for disqualifying the whole group of specific formulations of sceptical arguments — including Williams's new sceptical challenge.

In Section 2, I present Lewis's views on naturalness. In Section 3, I present Williams's new sceptical challenge. In Section 4, I critically examine Williams's argumentation and I reconsider the general conditions for the applicability of Lewis's criterion of naturalness.

2. David Lewis on naturalness

Lewis introduced the idea of the naturalness of properties in [13]. One of his key motivations was to find a way to trace objective distinctions and similarities between objects and use it as a ground for a comparison of different ways of categorizing objects. The idea that some ways of categorizing objects are "somehow better" is intuitively appealing. Most people would agree that the class consisting of all pieces of copper is "somehow better" than the class consisting of a piece of copper, a monkey, an aeroplane and a revolution. The question is, however, how this intuitive idea should be spelled out in detail.

According to Lewis, any categorization of objects into a class instantiates a property, but not all the properties are equally good in matching objective categories/joints in the world — some properties are more natural than other. Because of that, for Lewis, the answer to the question of what makes some categorization better than other lies in answering the question of how to compare the naturalness of properties.

As we can see, the properties which we have to take into account in comparisons can be instantiated by classes of radically different kinds of objects — we can imagine classes involving ordinary spatial objects (such as aeroplanes) and living creatures (such as monkeys), but also events (such as revolutions) and the list may go on. The question is how we can find some common ground for the comparison of naturalness of properties if they can be instantiated by such a heterogeneous group of objects. For Lewis, the most reasonable candidate for such a common ground is physical reality. What monkeys, aeroplanes, copper and even revolutions have in common is that they are part of the same physical world.[5] In fact, physics tells us that, at some level, everything in the world consists of *the same physical particles* — quarks and electrons.

Lewis's reliance on physics in this matter has not been instrumental. It was a result of his general persuasion that physics has a privileged access to our world and its structure and it served as a ground for his metaphysics.[6] He summarized his view in [16] under the label of Humean supervenience. Two backbones of his view are that a) the only fundamental properties/relations in our world are microphysical and spatiotemporal (related to points or point-sized occupants of points) and b) all other properties and relations supervene on the fundamental ones.[7] The point (a) led Lewis to a consequence that the only fundamental properties/relations in our world are those related to quarks and electrons as the smallest (point-sized) occupants of our world.[8] In Lewis's terminology fundamental microphysical properties and relations are perfectly natural [see 13, pp. 357–368].

The point (b) led Lewis to his reductive account of properties. All the properties which are not fundamental, for example macroscopic properties such as being human or being wooden, can be "reduced" to fundamental ones and can be ranked according to their degree of naturalness. Perfectly natural properties are at the bottom of the ranking. All other properties are ranked on the basis of their distance from perfectly natural ones. The distance of properties from fundamental ones is measured by the length of their definitions in terms of fundamental properties and relations. A good example of a simplified definition of a hydrogen atom is provided in [26, p. 143]: "$\exists x \exists y (Ex \wedge Py \wedge Rxy)$", which is a formal

[5] It is questionable if, e.g., abstract objects are part of the physical world. At this point I suggest understanding 'be part of' in a very broad sense. The topic of the relation between abstract objects and the physical world will be discussed later.

[6] See Nolan [18] for a discussion.

[7] Weatherson [30] characterizes (a) by the label 'spatiotemporalism' and (b) by the motto 'Truth supervenes on being'.

[8] Notice that, according to Lewis, properties/relations which we consider to be fundamental can change with possible developments in physics. However, it is not clear whether the condition of spatiotemporality can be compatible with current quantum mechanics. As Nolan [18] argues, this possible incompatibility is not a consequence of ignorance, but of Lewis's erudite criticism of quantum mechanics (as known and discussed in the nighties). Unfortunately, it is not clear whether and how Lewis would adapt his views.

notation of "There exist an electron and a proton, the first of which orbits the second".[9]

In one way or another, we should be able to find a chain of definitions which links the property being defined to some fundamental properties and relations. Writing down such definitions could be laborious for macroproperties such as being human, but I accept as an initial assumption of this paper that, theoretically, there is no obstacle which could prevent us from defining any property in fundamental terms.[10]

Since the length of the definitions of some properties can be enormous, we could state the degree of naturalness of a property as the number of connectives in its definition stated in fundamental terms and so we can talk about the syntactic complexity of definitions.

Lewis's criterion for ranking properties is really simple: the longer the definition, the less natural the property. For example, we can stipulate that the definition of the property of being a molecule of water is considerably shorter than the definition of the property of being wooden. To define the property of being a molecule of water we need only to double the definition of the property of being a hydrogen atom, to add the definition of the property of being an atom of oxygen, and to add the definition of their relation (chemical bond). In contrast to that, the definition of being wooden consists of definitions (microphysical descriptions) of the properties and relations of all the atoms and molecules which form what we call wood. Therefore, the property of being a molecule of water is more natural than the property of being wooden.

2.1. From properties to predicates

Beside metaphysics, Lewis found the use for the idea of natural properties also in philosophy of language. According to Lewis, properties can play the role of semantic values for predicates. Thanks to this step, the application of the criterion of naturalness in language is quite straightforward. If we have a tool for comparing the naturalness of properties, we can apply it to predicates as well.

[9] It is really only a simplified version of the definition — terms such as 'proton' are not fundamental and a hydrogen atom is not a property. But we can accept this simplification, since physics can give a definition of protons in terms of quarks and their properties/relations and we can talk about the property of being a hydrogen atom.

[10] A similar approach is generally accepted and its discussion can be found in [23].

Lewis believed that the criterion of naturalness could be used to answer some of the so-called sceptical arguments. Since the 1960s we have been able to find several sceptical arguments with a similar pattern of argumentation: For each expression you can find at least two meaning candidates which fit its use perfectly well. Since they fit the use perfectly well, there is *no fact of the matter* which candidate is the correct one. Kripke [8] used the example of addition/quaddition, Goodman [5] used the example of green/grue, and Putnam [19, 20] presented his model-theoretic argument.[11] Intuitively, we know that a predicate of a natural language can mean green, but hardly something like grue. But finding a satisfactory answer to sceptical arguments turned out to be more problematic than it might seem.[12] Lewis therefore suggested using the criterion of naturalness to answer at least those sceptical arguments which focus on (or which can be restated in such a way that they focus on) predicates: To decide which meaning candidate is the best for a predicate means to decide which candidate is the most natural, i.e. which candidate has the shortest definition when stated in fundamental terms. In Lewis's words, the best meaning candidate is a relatively natural property [see 13, p. 372]. The word 'relatively' in this context indicates the ranking of the meaning candidate — it says that the meaning candidate is more natural *relative* to other meaning candidates which fit the use of the predicate equally well.

2.2. Williams on David Lewis

Unfortunately, the definition of the criterion of naturalness is the only part of Lewis's theory which is shared by current interpretations of Lewis. What role the criterion of naturalness should play in Lewis's broader views on language as stated in [9] or [12] is still an open question. Sider [24, 25, 26], Weatherson [28], and Stalnaker [27] argue that the criterion of naturalness determines the meaning of particular predicates. Schwarz [22] and Weatherson [29] argue that the criterion determines the mental content and Williams [31, 33] presents a holistic interpretation of naturalness.

[11] Each of the authors stated the argument in slightly different ways. Goodman presented it as a problem of induction and Putnam as a problem of reference. But the main point is still the same.

[12] The absence of a satisfactory solution compelled philosophers, e.g., [3, 21], to accept the results of sceptical arguments.

With respect to the topic of this paper, I will focus on Williams's interpretation in detail.[13] The most distinctive feature of Williams's interpretation of Lewis's view is a holistic approach. Williams believes that the criterion of naturalness is primarily meant to (partially) determine the simplicity of theories — the best theory is the one which best fits the data and if there are more theories which fit the data equally well, then the criterion applies. From this perspective, finding the best meaning candidate is only a consequence of finding the best semantic theory for a language in general — the best meaning candidate for a predicate is the one which is assigned to a predicate by the best semantic theory.

When applied to our current topic of semantics, the best semantic theory (interpretation) should be the one which fits the data (the use of a language) better than others and if more semantic theories fit the data equally well, then the best semantic theory is the one which has the lowest syntactic complexity when stated in fundamental terms. Now the question is how we can measure the syntactic complexity of a semantic theory. First of all, Williams takes into account semantic theories as stated by Lewis [10, p. 35]. According to Lewis, the major part of a semantic theory consists of clauses assigning semantic values to expressions. Because of that we can assume that the syntactic complexity of a semantic theory (roughly) equals the syntactic complexities of the definitions of the assigned semantic values.[14] In other words, the best semantic theory is the one which assigns, relatively to the use, the most natural semantic values (properties) to the expressions (predicates) of some language.

But according to Williams, Lewis's strategy is a kind of two-step interpretationism. The first step is to specify the data — in general, some true statements about the states of the world. The second step is to find a semantic theory which is able to interpret statements of these facts as (mostly) true. What I presented in the last paragraph is the second step. Clauses assigning semantic values are clauses which interpret expressions

[13] Notice that deciding which interpretation is correct is notoriously problematic. After 1983, Lewis applied the idea of naturalness beyond the context of sceptical arguments several times, e.g., in [15, 16, 17], and it is not clear how all those applications hook together. In what follows, I refrain from making claims about the correctness of particular interpretations.

[14] Since the semantic values of predicates are properties, the naturalness of semantic values is stated by the definitions of properties in fundamental terms.

in such a way that statements of facts in which those expressions are included are rendered true.

The only question to be answered is what this 'data' is for semantic theories. Lewis introduced the so-called global folk theory in [14], which is supposed to play the role of sentences to be interpreted. Since Lewis is quite imprecise as to what this global folk theory should be I follow Williams's definition. Global folk theory is "the sum total of all the platitudes gathered from every walk of life — all the sentences that we take to be too obvious to question" [31, p. 367]. In other words, we can understand global folk theory as a set of true sentences such as 'Grass is green'; 'Penguins cannot fly'; 'If you are hurt, you feel pain' and so on.

3. Williams against naturalness

If we accept Williams's holistic approach to naturalness, then we can state a sceptical argument in terms of theories: For any global folk theory we can imagine at least two interpretations (semantic theories) which fit global folk theory equally well. Since they fit global folk theory equally well, there is *no fact of the matter* which one is the best/correct interpretation (semantic theory). What we obtained now is a simplified version of Putnam's model-theoretic argument. According to Putnam's argument [20], the only constraint which decides if something is an interpretation of global folk theory (a language in general) is the criterion of truth-preserving. But the criterion of truth-preserving is too general to determine the reference of subsentential items. If the reference of subsentential items is underdetermined, then we can easily formulate candidate interpretations. We only need to take some interpretation of a language and add to it a permutation clause which permutes the initial assignments of names to referents and which permutes predicates in such a way that the overall truth values of sentences stay unchanged.

When answering sceptical arguments, Lewis builds on the assumption that permuted interpretations consist of some intended interpretation + permutation clauses. So, if the intended interpretation requires n symbols to be defined in fundamental terms then any permuted interpretation needs at least $n + x$ symbols, where x stands for the symbols which are needed to state the syntactic complexity of permutation clauses and $x > 0$. Therefore, Lewis argues, the intended interpretation is the best interpretation of global folk theory and we have a way out of this sceptical argument. We can call this *a dialectical comparison of naturalness* — as opposed to a comparison which would be based on the actual syntactic complexities of compared definitions.

3.1. The arithmetical interpretation

As Williams argues, Lewis's response to sceptical arguments is insufficient, because it is possible to find alternative interpretations which are not only permuted versions of some intended interpretation. To show that there really are such interpretations, Williams relies on Henkin's model. According to Henkin's model-existence theorem [6, 7], for any consistent set of sentences (a language) we can construct a model which serves as an interpretation of the set. If we want to construct Henkin's model for a certain language, first of all we have to create a domain of the model which consists of equivalence classes of all the constants of the language. The equivalence classes are then assigned to predicates as their extensions and therefore present an interpretation of the language. As we noted, global folk theory is a set of sentences. If we are willing to follow Williams and assume that the set is consistent, then it must be possible to construct Henkin's model for it as well.

What is more, Henkin's model has one specific feature — what equivalence classes consist of can be freely chosen on the condition that when assigned, equivalence classes do not contradict the truth values of the sentences being interpreted. In other words, Henkin's model can interpret global folk theory as referring to anything, if the interpretation does not contradict the truth values of sentences in global folk theory.

Williams's idea is that if we construct Henkin's model for global folk theory and replace the constants in its domain with natural numbers, we can create an arithmetical interpretation of global folk theory [see 31, pp. 382–385]. In practice, if the arithmetical interpretation turns out to be the best candidate for global folk theory, then any sentence uttered in the "language of global folk theory" should be interpreted as some kind of mathematical claim. So when someone is talking about a barking dog or about the broken arm of her brother, she should be actually interpreted as making some complicated mathematical claims.

As Williams argues, the problem for Lewis is that the arithmetical interpretation cannot be dismissed by his dialectical comparison of naturalness because the arithmetical interpretation is not formulated as a permuted version of the intended interpretation.

3.2. The argument for Pythagorean worlds

However, Williams tries to push his argument further — towards the existence of Pythagorean worlds; that is to say, worlds in which the arithmetical interpretation is more natural than the intended interpretation.

Williams puts forward the following consideration: imagine a world which is a perfect duplicate of our world from quarks up, but the level of quarks is not its fundamental layer. In the duplicate world, quarks are composed of even smaller spatiotemporal microphysical particles with different properties and relations. Otherwise the duplicate world is the same as ours, not only on the microphysical level of quarks, but also on the level of macrophysical properties. If the duplicate world has the same macroproperties, then the same global folk theory applies in both worlds (because the same sentences are true in both worlds). If the same global folk theory applies in both worlds, then the same arithmetical interpretation applies in both worlds. This leads Williams to the conclusion that the arithmetical interpretation "enable(s) us to provide an interpretation whose fit with total theory is quite independent of what might or might not be going on at extreme microscopic levels" [31, p. 388]. In other words, the arithmetical interpretation is independent of microphysical settings of the worlds. Let us call this *an argument for microphysical independence*.

From this point, it is only a small step to Pythagorean worlds: imagine a string of possible worlds. Each world in the string is a perfect duplicate of ours from quarks up, but each subsequent world in the string has one more microphysical layer below the level of quarks. With added layers in successive worlds, the level of fundamental properties moves down as well. Since the same global folk theory applies in these worlds, we can assume that something similar to our intended interpretation applies in these worlds as well. As the fundamental layer moves down, the syntactic complexity of the intended interpretation increases, because the chain of definitions from macroproperties to fundamental properties is longer. Since the arithmetical interpretation is independent of the microphysical settings of the worlds, its syntactic complexity stays unchanged. If the string of worlds is long enough, the syntactic complexity of the intended interpretation must exceed the syntactic complexity of the arithmetical interpretation at some point, hence all the worlds beyond this point would be Pythagorean.

As Williams concludes, acceptance of the arithmetical interpretation as the best semantic theory in any macroduplicate of our world is counterintuitive. What is more, we do not know the actual syntactic complexities of interpretations in our world and so there is a chance that our world is also Pythagorean. Because of that, Williams argues, we should abandon Lewis's criterion of naturalness in its current form.

4. Pythagorean worlds reconsidered

In the rest of this paper, I will argue that the argument for Pythagorean worlds is based on a problematic assumption and that there is a way to respond to the new sceptical challenge and to save the criterion of naturalness. In this subsection, I will show why the conclusion of the argument for microphysical independence is not able to serve as a premise for the subsequent argument for Pythagorean worlds. To sum up, Williams noticed that the arithmetical interpretation can fit global folk theory in all the worlds which are duplicates of ours at the macro level, regardless of which microphysical properties and relations are fundamental in these worlds. I must say that this is correct — if an interpretation can fit data in our world, then it can fit the same data in any world.

Subsequently, Williams used the conclusion that the arithmetical interpretation is independent of microphysical fundamental properties and relations as a premise in the argument for the existence of Pythagorean worlds. The whole argument for the existence of Pythagorean worlds is based on the assumption that the degree of naturalness (the syntactic complexity) of the arithmetical interpretation stays fixed — because it is independent of microphysical fundamental properties and relations — while the degree of naturalness of the intended interpretation increases with added layers.

The problem is that when Williams starts the argument for the existence of Pythagorean worlds, he presupposes that it is *the degree of naturalness* of the arithmetical interpretation which is independent of microphysical fundamental properties and relations. But the conclusion of the argument for microphysical independence states that *the fit* of the arithmetical interpretation is independent of microphysical fundamental properties and relations.

The question is whether the independence of the fit of the arithmetical interpretation guarantees the independence of its degree of naturalness. And the answer to this question is negative. Fitting global folk theory equally well equals interpreting the same sentences of global folk theory as true. The only factor which influences sentences being interpreted as true by the arithmetical interpretation is a redistribution of members (constants/numbers) in its equivalence classes. On the other hand, the degree of naturalness depends on the definitions of all the assigned semantic values (equivalence classes) in fundamental terms. But the redistribution of members in equivalence classes and the definitions of equivalence classes are mutually independent — we can change the way we define equivalence classes and preserve the same redistribution of their members (and vice versa). If the fit of the arithmetical interpretation has no influence on the way we define the arithmetical interpretation, then the argument for the microphysical independence gives us no reason to suppose that the degree of naturalness of the arithmetical interpretation is independent of the microphysical fundamental properties and relations.

In summary, the argument for microphysical independence is not able to support the premise on which Williams relies in the subsequent argument for the existence of Pythagorean worlds. Without the premise that the degree of naturalness of the arithmetical interpretation is independent of microphysical fundamental properties and relations, the whole idea of Pythagorean worlds collapses.[15]

4.1. Naturalness of the arithmetical interpretation

The outcome of the previous subsection is that the plausibility of the argument presented by Williams depends on the plausibility of the assumption that the arithmetical interpretation has a degree of naturalness and at the same time is independent of microphysical fundamental properties and relations. However, the discussion whether the assumption is plausible uncovers much broader issue which spans well beyond the plausibility of Williams's argument — it leads to a discussion of an appropriate application of the criterion of naturalness.

[15] Williams [32] argues that in some worlds all properties can be independent of the structure of these worlds. He calls such properties emergent. Even if we are willing to accept the possibility of "emergent worlds", it is questionable whether the arithmetical interpretation can have any degree of naturalness in such worlds. This topic will be discussed in the next subsection.

If we assume, as Lewis did, that the only fundamental properties and relations in our world are those proposed by physics, then to be independent of microphysical fundamental properties/relations equals to be independent of fundamental properties/relations, full stop. As far as I can see, if an interpretation is independent of fundamental properties and relations, then it cannot have a degree of naturalness.

A degree of naturalness is defined as a syntactic complexity of definitions stated in fundamental terms. To find out the syntactic complexity of an interpretation, we first need to state its definition in fundamental terms. In order to state a definition in fundamental terms, we have to find a definitional chain between what we are defining and some fundamental properties and relations. If we accept Lewis's assumption that the only fundamental properties and relations in our world are those proposed by physics and we accept Williams's conclusion that the arithmetical interpretation is independent of these properties and relations, we should also accept that it is not possible to state the definition of the arithmetical interpretation in fundamental terms and, subsequently, to state the degree of naturalness of the arithmetical interpretation.

This does not mean that we cannot state a definition of the arithmetical interpretation in some non-fundamental terms. But if we do so, then we are not able to state its degree of naturalness. In my view, this makes the arithmetical interpretation an ineligible candidate for the comparison of naturalness.

Notice a parallel with a comparison of scientific explanations or theories. There are, for example, several neuroscientific theories that try to explain consciousness — the global neuronal workspace, recurrent processing theory, higher order theory, and information integration theory.[16] All four theories are built on the assumption that consciousness can be traced to specific biochemical properties and mechanisms of brains and all four theories lead to specific predictions regarding these mechanisms. Beside these four theories, we can imagine a theory of consciousness based on some esoteric properties according to which consciousness is a result of a presence of the divine spark in our soul. All four theories mentioned above are empirically testable and the future research can show that one of the theories is better or worse than the other three. In contrast to that, the esoteric theory is not only a worse theory. The fact that the esoteric theory is based on properties and mechanisms which

[16] For a survey of different theories see Wu [35].

are not empirically testable makes it an ineligible candidate for a comparison because it fails to comply with the basic standards for a scientific explanation and scientific theories.

Similarly, if the arithmetical interpretation is not linked to any fundamental properties and relations, then it is an ineligible candidate for a comparison of naturalness because it fails to comply with the basic standards for a semantic theory/an interpretation of a language. Ruling out ineligible candidates is the main reason, I believe, why Lewis characterizes the comparison of naturalness not only as a comparison of syntactic complexities of definitions, but as a comparison of syntactic complexities of definitions *when defined in terms of fundamental properties and relations.* The condition that interpretations must be defined in fundamental terms secures that only eligible candidates are considered for the comparison of naturalness.

This point can be generalized to any case where the criterion of naturalness is in use: in order to have a degree of naturalness, a property/an interpretation must be related by a definitional chain to some fundamental properties and relations. Let us call this *the constraint of fundamental interconnectivity.* If an interpretation is independent of fundamental properties and relations, then it has no degree of naturalness and this makes it an ineligible candidate for the comparison of naturalness.

The acceptability of the constraint can be supported by considerations about the general function of language as a representational tool. If the primary function of language is to represent the world around us, then any interpretation which interprets language as representing something else misses the point and we have no reason to consider it as a serious candidate for a semantic theory. The constraint of fundamental interconnectivity helps us to exclude all such interpretations even before the criterion of naturalness is in use.[17]

4.2. Plurality of fundamental properties

If the arithmetical interpretation has a degree of naturalness, as Williams claims, then there must be some way to define it in fundamental terms. I see two ways in which the interpretation could be defined in fundamental terms:

[17] Interestingly, in contrast to Williams's arithmetical interpretation, Putnam's permuted interpretation satisfies the constraint.

(A) defined in terms of the microphysical fundamental properties and relations;
(B) defined in terms of some other fundamental properties and relations.

At first, I will discuss option (B) when applied to the case of the arithmetical interpretation. Although it will be discussed through a particular example, I believe that it can lead us to a general principle for the application of the criterion of naturalness.

The only assumption Williams needs to support his argument for the existence of Pythagorean worlds is the assumption that the syntactic complexity of the arithmetical interpretation is independent of the microphysical fundamental properties and relations. To achieve this, he can suppose that the degree of naturalness of the arithmetical interpretation is fixed for all worlds in the string by some different fundamental properties and relations — let us call them mathematical.[18] As a matter of fact, this really seems to be the way in which Williams thinks about the arithmetical interpretation, though he discusses the option only in a short footnote:

> This supposes the mathematical vocabulary involved to be "perfectly natural." This might be questioned, but so long as there is some finite specification of mathematical vocabulary in perfectly natural terms, which does not vary from world to world, the overall point will not be affected. [31, p. 388].

If we define the arithmetical interpretation in terms of mathematical fundamental properties and relations and apply it to our world, we have two candidate interpretations for our global folk theory. The first candidate is the intended interpretation defined in terms of the microphysical fundamental properties and relations. The second candidate is the arithmetical interpretation defined in terms of the mathematical fundamental properties and relations.

The attractiveness of examples in which two candidates are defined in different fundamental terms is understandable. If we are liberal enough in our understanding of fundamental properties and relations, we can

[18] A possibility of such fundamental properties is discussed in [26]. In his book, Sider argues that besides microphysical fundamental properties/relations we should also accept mathematical and logical fundamental properties/relations. Unfortunately, he does not focus on questions 1) which mathematical properties and relations are fundamental and why and 2) how we could use them to state the syntactic complexities of interpretations/properties.

formulate many candidate interpretations with minimal effort and none of them will have the form of a permutation. But is it possible to compare the degrees of naturalness of two rival interpretations if their definitions are stated in different fundamental terms?

It may seem that I am trying to find a problem where clearly no problem exists. We can easily imagine that we find out the syntactic complexity of the arithmetical interpretation defined in mathematical fundamental terms as well as the syntactic complexity of the intended interpretation defined in microphysical fundamental terms. Let us say that the syntactic complexity of the arithmetical interpretation is 2000 and the syntactic complexity of the intended interpretation is 2500.[19] Seemingly, there is no problem in comparing them and stating that the degree of naturalness of the arithmetical interpretation is higher and therefore we live in a Pythagorean world.

My question now is why we should compare syntactic complexities based on different fundamental properties and relations in a 1:1 fashion. How can we know that the value of one connective in the definition stated in microphysical fundamental terms is the same as the value of one connective in the definition stated in mathematical fundamental terms?

The problem of such comparisons is that the definitional chain from macroproperties to fundamental layers starts to be a variable in comparisons and distorts the comparisons. Such a comparison of naturalness is like trying to find out who is better at the high jump, but the highest point overleaped by the first high jumper is measured in centimetres and the highest point overleaped by the second high jumper is measured in inches. Can we say that the first high jumper is better than the second one simply by comparing the measured numbers? Such a statement would lose any entitlement to objectivity because the result would be distorted by different conditions of the respective measurements (in particular, by the difference in the systems of measurements employed). Similarly, we can state the syntactic complexities of interpretations defined in different fundamental terms. But we cannot pretend that the results can be used to objectively compare the naturalness of interpretations because the results are distorted by different conditions in which we stated the syntactic complexities (in particular, by the difference

[19] The numbers here represent the number of connectives in definitions of interpretations in fundamental terms.

in the definitional chains needed to link macroproperties with relevant fundamental properties and relations).

The problem of vanished objectivity is the problem of sneaked arbitrariness. If we allow comparisons based on definitions stated in different fundamental terms, and we are liberal enough with respect to which properties and relations we consider to be fundamental, then anyone can imagine anything and any reasonable discussion of naturalness fades away. We can always imagine a case in which the initial setting[20] of definitional chains influences a comparison to such an extent that the comparison always has an obvious winner — even before we state the actual definitions. We can always imagine that the arithmetical interpretation is defined in such fundamental terms that its definitional chain includes much more layers from macroproperties to fundamental properties and relations than the one related to our intended interpretation and the difference in the number of layers is so big that the intended interpretation will always be a better candidate for our semantic theory. On the other hand, we can always imagine that the arithmetical interpretation is defined in such fundamental terms that its definitional chain includes far fewer layers and the difference in the number of layers is so big that the intended interpretation will never be a better candidate for our semantic theory.

If the factor of definitional chains related to different fundamental properties and relations influences the comparisons of naturalness, then the factor should be taken into account in the comparisons of interpretations. One way how to do it is to state syntactic complexities in different units corresponding to different fundamental properties and relations. For example, we can say that the syntactic complexity of the arithmetical interpretation is 2000 CMaF (connectives in a definition stated in mathematical fundamental terms) and the syntactic complexity of the intended interpretation is 2500 CMiF (connectives in a definition stated in microphysical fundamental terms).

If we do so, then the acceptability of comparisons based on different fundamental terms depends on the possibility to state "rules of conversion" between different kinds of fundamental terms. Unfortunately, as far as I know, there are no viable attempts available so far. Williams restricts his discussion of this matter to a short footnote and even Sider

[20] By 'initial setting' I mean the number of layers/steps from macroproperties to fundamental properties and relations.

[26], as the strongest proponent of mathematical fundamental properties and relations, tries to argue only that there are *some* mathematical fundamental properties and relations, but he is not able to say which mathematical properties and relations actually are fundamental (the unit); not to mention how to compare them with microphysical fundamental properties and relations (the rules of conversion). And I doubt that any such attempt can be successful. Of course, we can stipulate *some* rules of conversion. But how do we know that those rules reliably represent the relation between different fundamental properties and relations? How can we find independent non-arbitrary criteria for comparing different fundamental properties and relations?

If we want to secure objective comparisons of interpretations (or meaning candidates), we must accept a stronger constraint than the constraint of fundamental interconnectivity — a constraint which does not state only that candidates must have definitional connections to some fundamental properties and relations, but which states that all candidates for the comparison of naturalness must be defined in terms of the same fundamental properties and relations. Let us call this *the constraint of fundamental uniformity*.

There is a grain of textual evidence in favour of this constraint in [13]. At the beginning of the paper, Lewis discusses possible work for his later criterion of naturalness. According to Lewis, a theory of naturalness should be able to make objective claims about the duplication of individuals. His suggestion is to define the duplication of individuals in terms of the arrangement of their microphysical fundamental properties. Two individuals are perfect duplicates iff they share all their fundamental properties. With regard to the definition of duplication Lewis claims:

> It presupposes the physics of our actual world; however physics is contingent and known a posteriori. The definition does not apply to duplication at possible worlds where physics is different, or to duplication between worlds that differ in their physics. [13, p. 356].

In other words, we cannot make a comparison of fundamental properties and so determine if two individuals are perfect duplicates if the individuals are part of worlds with different physics designating different fundamental properties and relations. This example indicates that Lewis was aware that if we try to compare candidates based on different fundamental properties and relations, the result will be distorted. The idea that he assumed the same constraint in the case of comparisons of naturalness seems reasonable despite the fact that he has never stated it explicitly.

4.3. The arithmetical interpretation and microphysical fundamental properties

To fulfill the constraint of fundamental uniformity, it should be possible to state the definition of the arithmetical interpretation in terms of microphysical fundamental properties and relations.[21] This means that we are forced to look for a definitional chain between the equivalence classes consisting of natural numbers and the microstructure of our world and so we are forced to find microphysical descriptions of numbers in the equivalence classes.[22] This is a rather strange requirement. There is no reason to assume that there should be such definitions and surely there are more intuitive ways how to define numbers. However, we are not looking for the best way to define the arithmetical interpretation; we are looking for the best way to compare naturalness and the fact that some interpretation is defined in a strange way does not have to mean that the result of the comparison will not be correct. On the contrary, the fact that it is hard to find a definition of an interpretation in microphysical fundamental terms can be seen as an indication that the interpretation is a very unnatural candidate for our semantic theory.

A possible way to find definitions of numbers in terms of microphysical fundamental properties and relations, in the context of [11], would be to suppose that numbers are abstract (theoretical) terms and to use the method of Ramsification. This presupposes that the numbers in the domain of the arithmetical interpretation are part of some bigger theory which includes a sufficient number of non-theoretical terms.[23] In the

[21] There is also a specific reason why Williams should frame his argument in such a way. One of the aims of Williams is to show that Lewis should give up his reductive account of properties. In the heart of Lewis's reductive account is the idea that microphysical fundamental properties and relations are somehow privileged. If Williams wants to show that the criterion of naturalness is not able to give satisfactory results, he should not build his argument on the assumption that there are other than microphysical fundamental properties and relations.

[22] All the definitions of interpretations/meaning candidates are produced in this way. For example, the definition of the property being wooden is a microphysical description of all the atoms which form what we normally call wooden.

[23] Most probably some version of mathematical nominalism could be applied. However, I am sceptical as to whether some existing version of mathematical nominalism as stated, e.g., in [4] or as discussed in [1] could provide the required definitions.

first step, we could use Ramsification to define each number by all the sentences of a theory which include the number + other non-theoretical terms. Then the definition of a number in microphysical fundamental terms would consist of microphysical definitions of all the non-theoretical terms which are involved in its definition.

Suppose that the method of Ramsification can provide definitions of all the numbers in the equivalence classes of the arithmetical interpretation in microphysical fundamental terms and so it is possible to state the degree of naturalness of the arithmetical interpretation. Nevertheless, the arithmetical interpretation is not a permuted version of the intended interpretation and so it is not possible to rely on Lewis's dialectical comparison of naturalness. Are we trapped in the position of not being able to compare their degrees of naturalness? I think not.

Notice that the definitions of ramsified semantic values always consist of definitions of several other terms, and only those "other terms" are defined in fundamental terms. For example, using the method of Ramsification, we can define the predicate 'being stressed' in terms of a biological theory that tells us what happens in a body as part of a stress response. A part of such a ramsified definition of 'being stressed' would be 'having increased levels of adrenaline, noradrenaline, and cortisol'. Thus, the final definition of 'being stressed' in fundamental terms would consists of (among other things) the definitions of properties 'being adrenaline', 'being noradrenaline', and 'being cortisol' in fundamental terms.

The point that I would like to stress is that the method of Ramsification multiplies the lengths of definitions in fundamental terms because each ramsified definition consists of several other definitions. Because of that, it is reasonable to assume that the syntactic complexity of an average ramsified semantic value is several times higher than the syntactic complexity of an average semantic value which can be defined directly in microphysical fundamental terms.

Keeping this in mind, let us summarize what we know about the arithmetical and the intended interpretation. Both interpretations interpret the same global folk theory. Therefore, they interpret the same number of predicates and so the number of semantic values they assign is the same as well. Moreover, we know that the arithmetical interpretation assigns only the equivalence classes of numbers as semantic values and to find their definitions in microphysical fundamental terms requires using the method of Ramsification. The intended interpretation, on the other hand, assigns various properties as semantic values — some of them are theoretical and can be defined only by Ramsification (such as being stressed), but some of them can have direct definitions in terms of fundamental properties (such as being aluminous).

If the method of Ramsification multiplies the lengths of definitions in fundamental terms, and all the semantic values assigned by the arithmetical interpretation need to be ramsified, while only some semantic values assigned by the intended interpretation need to be ramsified, then we have a strong reason to believe that the syntactic complexity of the intended interpretation is lower. If the intended interpretation has a lower syntactic complexity and hence it is more natural than the arithmetical interpretation, then we have a reason to believe that we do not live in a Pythagorean world.

The way in which we reached this conclusion may seem peculiar. Trying to find a definition of the arithmetical interpretation — which assigns equivalence classes of numbers as semantic values — in terms of microphysical fundamental properties and relations may look strange. But as regards the comparison, the result given by the criterion of naturalness is the one which we could expect. If all the semantic values of the arithmetical interpretation need to be ramsified in order to be defined in fundamental terms, then the degree of naturalness of the arithmetical interpretation is most probably low. But there is nothing wrong about it. The criterion of naturalness shows only what was clear from the beginning — that the arithmetical interpretation is a theoretical model of our language, not the best candidate for our semantic theory.

5. Conclusion

First of all, there is no reason to believe that Williams's argument for microphysical independence can guarantee that the arithmetical interpretation has a degree of naturalness and at the same time is independent of fundamental properties and relations. In accordance with the constraint of fundamental interconnectivity, if the arithmetical interpretation has a degree of naturalness, then it must be possible to state its definition in fundamental terms.

In this paper, I discuss two ways how to define the arithmetical interpretation in fundamental terms. It can be either defined in microphysical fundamental terms or in terms of some other fundamental properties and relations. The second option leads to a problem as the difference in fundamental terms distorts the comparison and the prospects of finding reliable rules of conversion are grim. The only reasonable option for an objective comparison of naturalness between the arithmetical interpretation and the intended interpretation is the case in which the arithmetical interpretation is defined in terms of the same fundamental properties as the intended interpretation, i.e. in terms of microphysical fundamental properties and relations.

However, in such a case we have a reason to believe that the degree of naturalness of the intended interpretation is higher. The reason is that the syntactic complexity of the arithmetical interpretation consists solely of definitions of semantic values which must be ramsified and the method of Ramsification multiplies the lengths of definitions in fundamental terms. If this is so, then the degree of naturalness of the arithmetical interpretation is most probably low and so we found a way to answer the new sceptical challenge proposed by Williams.

The most important conclusion of this paper, however, came along the way. By reconsidering several different scenarios for the comparison of naturalness, we can conclude that in order to make an objective comparison of naturalness, all the candidates for a comparison must be defined in terms of the same fundamental properties and relations. This constraint of fundamental uniformity serves as an important precondition for the comparison of naturalness — it secures a proper functioning of the criterion of naturalness by eliminating ineligible candidates for the comparison of naturalness.

Acknowledgments. The work on this paper was supported by the Specific Research Grant "Empirical grounds of inferentialism" funded by the Philosophical Faculty, University of Hradec Králové (2019–2020). I would like to thank three anonymous reviewers for their valuable comments on the previous version of this paper.

References

[1] Burgess, J. and Rosen, G., *A Subject With No Object*, Oxford: Clarendon Press, 1997.

[2] Davidson, D., "Radical interpretation", *Dialectica* 27, 3/4 (1973): 313–328. DOI: 10.1111/j.1746-8361.1973.tb00623.x

[3] Davidson, D., "The inscrutability of reference", *Southwestern Journal of Philosophy* 10, 2 (1979): 7–19. DOI: 10.5840/swjphil197910229

[4] Field, H., *Science Without Numbers*, Princeton: Princeton University Press, 1980.

[5] Goodman, N., *Fact, Fiction, and Forecast*, Cambridge MA: Harvard University Press, 1983.

[6] Henkin, L., "The completeness of the first-order functional calculus", *Journal of Symbolic Logic* 14, 3 (1949): 159–166. DOI: 10.2307/2267044

[7] Henkin, L., "Completeness in the theory of types", *Journal of Symbolic Logic* 15, 2 (1950): 81–91. DOI: 10.2307/2266967

[8] Kripke, S., *Wittgenstein on Rules and Private Language*, Cambridge: Harvard University Press, 1982.

[9] Lewis, D., *Convention: A Philosophical Study*, Cambridge: Harvard University Press, 1969.

[10] Lewis, D., "General semantics", *Synthese* 22, 1 (1970): 18–67. DOI: 10.1007/BF00413598

[11] Lewis, D., "How to defiìne theoretical terms", *Journal of Philosophy* 67, 13 (1970): 427–446. DOI: 10.2307/2023861

[12] Lewis, D., "Languages and language", pages 3–35 in K. Gunderson (ed.), *Language, Mind, and Knowledge*, Minneapolis: University of Minnesota Press, 1975.

[13] Lewis, D., "New work for a theory of universals", *Australasian Journal of Philosophy* 61, 4 (1983): 343–377. DOI: 10.1080/00048408312341131

[14] Lewis, D., "Putnam's paradox", *Australasian Journal of Philosophy* 62, 3 (1984): 221–236. DOI: 10.1080/00048408412340013

[15] Lewis, D., "Meaning without use: Reply to Hawthorne", *Australasian Journal of Philosophy* 70, 1 (1992): 106–110. DOI: 10.1080/00048408112340093

[16] Lewis, D., "Humean supervenience debugged", *Mind* 103, 412 (1994): 473–490. DOI: 10.1093/mind/103.412.473

[17] Lewis, D., "Reduction of mind", pages 412–431 in S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*, Oxford: Blackwell, 1994.

[18] Nolan, D., "Lewis's philosophical method", pages 25–39 in B. Loewer and J. Schaffer (eds.), *A Companion to David Lewis*, Oxford: Wiley-Blackwell, 2015.

[19] Putnam, H., "Models and reality", *Journal of Symbolic Logic* 45, 3 (1980): 464–482. DOI: 10.2307/2273415

[20] Putnam, H., *Reason, Truth, and History*, Cambridge: Cambridge University Press, 1981.

[21] Quine, W., *Word and Object*, Cambridge: MIT Press, 1960.

[22] Schwarz, W., "Against Magnetism", *Australasian Journal of Philosophy* 92, 1 (2014): 17–36. DOI: 10.1080/00048402.2013.765900

[23] Sider, T., "Sparseness, immanence, and naturalness", *Noûs* 29, 3 (1995): 360–377. DOI: 10.2307/2215604

[24] Sider, T., *Four-Dimensionalism*, Oxford: Oxford University Press, 2001.

[25] Sider, T., "Ontological realism", pages 384–423 in D. Chalmers, D. Manley and R. Wasserman (eds.), *Metametaphysics*, Oxford: Oxford University Press, 2009.

[26] Sider, T., *Writing the Book of the World*, Oxford: Oxford University Press, 2011.

[27] Stalnaker, R., "Lewis on intentionality", *Australasian Journal of Philosophy* 82, 1 (2004): 199–212. DOI: 10.1080/713659796

[28] Weatherson, B., "What good are counterexamples?", *Philosophical Studies* 115, 1 (2003): 1–31. DOI: 10.1023/A:1024961917413

[29] Weatherson, B., "The role of naturalness in Lewis's theory of meaning", *Journal for the History of Analytic Philosophy* 1, 1 (2013): 1–19. DOI: 10.4148/jhap.v1i10.1620

[30] Weatherson, B., "Humean supervenience", pages 101–115 in B. Loewer and J. Schaffer (eds.) *A Companion to David Lewis*, Oxford: Wiley-Blackwell, 2015.

[31] Williams, J. R. G., "Eligibility and inscrutability", *Philosophical Review* 116, 3 (2007): 361–399. DOI: 10.1215/00318108-2007-002

[32] Williams, J. R. G., "The possibility of onion worlds: Rebutting an argument for structural universals", *Australasian Journal of Philosophy* 85, 2 (2007): 193–203. DOI: 10.1080/00048400701380129

[33] Williams, J. R. G., "Lewis on reference and eligibility", pages 367–381 in B. Loewer and J. Schaffer (eds.): *A Companion to David Lewis*, Oxford: Wiley-Blackwell, 2015.

[34] Wittgenstein, L., *Philosophical Investigations*, Oxford: Blackwell, 1953.

[35] Wu, W., “The neuroscience of consciousness”, in E. Zalta (ed.): *The Stanford Encyclopedia of Philosophy*, 2018. Available at: https://plato.stanford.edu/archives/win2018/entries/consciousness-neuroscience

MATEJ DROBŇÁK
Department of Philosophy and Social Sciences
University of Hradec Králové
Hradec Králové, Czech Republic
matej.drobnak@uhk.cz