



Indistinguishability as a constraint on priors

Ron Avni¹

Received: 18 January 2024 / Accepted: 13 April 2024

© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

Invoking metaphysical naturalness is “perhaps the most popular proposed solution” to the problem of grue (Hedden in *Can J Philos* 45:716–743, 2016). Accordingly, Bradley (*Mind* 129:179–203, 2020) develops a “Lewisian” method for constraining priors based on the syntactic simplicity of descriptions of possible worlds in a language whose predicates correspond to natural properties. The Lewisian method therefore requires a solution to the arguably unsolved problem of measuring syntactic simplicity. But this paper argues that *given* a solution to this problem, there is a better alternative to the Lewisian method: a novel “Neo-Carnapian” method for constraining priors based on the indistinguishability of experiences.

Keywords Induction · Naturalness · Lewis · Carnap · Kolmogorov complexity

1 Introduction

According to a Bayesian, what should one’s priors be? Bayesians typically accept that priors should be probabilistically coherent, but to what extent should one’s priors be constrained further? “Subjective” Bayesians may require *some* constraints on priors, such as that propositions should be assigned priors of one or zero iff the propositions are logical truths or fallacies, respectively. But Subjective Bayesians typically reject significant additional constraints on priors. Arguably, a cost of the Subjective Bayesian view is that the view deems permissible some inductive inferences that intuitively seem bad. According to a Subjective Bayesian, it is possible for a rational agent to observe a large number of green emeralds and inductively infer that all emeralds are grue¹ (that is, green if observed before time t and otherwise blue) rather than green. Intuitively, such an inference is a bad inference. But, arguably, the Subjective Bayesian must accept that such an inference could be rational. Such considerations motivate “Objective” Bayesians to find a principled basis on which to significantly constrain priors. The

¹ As in Goodman (1955).

✉ Ron Avni
ron.avni@gmail.com

¹ The University of Texas at Austin, Austin, TX, USA

hope is that correctly constrained priors will render intuitively satisfying verdicts regarding the goodness of inductive inferences.

Invoking metaphysical naturalness is “perhaps the most popular proposed solution” to the problem of *grue* (Hedden, 2016, p. 724). Accordingly, Bradley (2020) develops a “Lewisian” Objective Bayesian method for constraining priors using naturalness. Briefly, the Lewisian method is to a priori identify the natural properties, describe every possible world in terms of these natural properties, and assign higher priors to worlds with syntactically simpler descriptions.

Bradley does not address a significant challenge for the Lewisian method. As discussed further herein, the challenge is that measuring syntactic simplicity is arguably an unsolved problem. Perhaps the most sophisticated measure of syntactic simplicity, commonly known as “Kolmogorov Complexity”,² faces difficulties when used to constrain priors.

But suppose we charitably assume that the problem of measuring syntactic simplicity can be solved. This paper argues that *given* a solution to this problem, there is a better alternative to the Lewisian method: a novel “Neo-Carnapian” method for constraining priors based on the indistinguishability of experiences. Briefly, the Neo-Carnapian method is to describe an agent’s experiences in terms of whether each experience is indistinguishable from each other experience, and to assign higher priors to hypotheses that the experiences have syntactically simpler descriptions.

As argued herein, the Neo-Carnapian method requires little by way of metaphysical or epistemological commitments. The Neo-Carnapian method is therefore an elegant way to parlay a solution to the problem of measuring syntactic simplicity into a solution to the problem of constraining priors. Put differently, the Neo-Carnapian method does the heavy lifting of constraining priors using just syntactic simplicity. By contrast, the Lewisian method requires much more than a solution to the problem of measuring syntactic simplicity: the Lewisian method additionally requires arguably onerous metaphysical and epistemological commitments to the existence and a priori identifiability of natural properties. The upshot is that the Neo-Carnapian method for constraining priors is more attractive than the Lewisian method.

This paper proceeds as follows. Section 2 summarizes Kolmogorov Complexity. Section 3 summarizes the Lewisian method. Section 4 presents the novel Neo-Carnapian method. Section 5 argues that the Neo-Carnapian method is more attractive than the Lewisian method. Section 6 concludes.

2 Kolmogorov complexity

This section briefly introduces Kolmogorov Complexity. The purpose of introducing Kolmogorov complexity is to provide a formal common ground for the Lewisian and Neo-Carnapian methods, to facilitate giving worked examples of each method, and to provide a concrete illustration of the challenge of measuring syntactic simplicity in the context of constraining priors.

² Although the name “Kolmogorov Complexity” is “well-entrenched” (Li and Vitányi 2019, p. viii), the first person to discover the idea was actually Ray Solomonoff (*ibid.*, p. 95).

Consider the sequence s_1 , which consists of all zeros, and the sequence s_2 , which consists of twenty zeros followed thereafter by ones:

$$s_1 : < 0, \dots >$$

$$s_2 : < 0, 1, 1, 1, \dots >$$

The Kolmogorov Complexity of a sequence is defined as the length of the shortest computer program written in a given programming language whose output is that sequence.³ Let us calculate the Kolmogorov Complexity of s_1 and s_2 using a “pseudocode” language that resembles commonly used programming languages such as C. In our pseudocode language, we may suppose that the shortest computer program whose output is s_1 is written (using 19 characters) as follows:

```
repeat {output ("0") }
```

The above program repeats endlessly the process of outputting the numeral “0”. By contrast, we may suppose that the shortest program whose output is s_2 is written (using 55 characters) as follows:

```
i=1; repeat { if (i<21) {output ("0") } else {output ("1") } ; i++; }
```

The above program begins by setting the variable i to equal one. Then, the program repeats endlessly the following process: first, the program checks whether i is less than 21; if it is, the program outputs the numeral “0”; otherwise, the program outputs the numeral “1”; either way, the variable i is then incremented by one; then the endlessly repeating process begins again—by checking whether i is less than 21, and so on.

Arguably, the logic of the first program is simpler than the logic of the second program since the latter program must switch from outputting zeros to outputting ones after outputting twenty ones. Whereas the shortest s_1 -generating program is 19 characters long in our pseudocode, the shortest s_2 -generating program is 55 characters long in our pseudocode. Thus, in this example, the Kolmogorov Complexity of s_1 is 19, or in shorthand, $K(s_1) = 19$, and $K(s_2) = 55$.

As mentioned above, Kolmogorov Complexity is perhaps the most sophisticated measure of syntactic simplicity. However, using Kolmogorov Complexity to constrain priors arguably faces difficulties. A full exposition of this topic is beyond the scope of this paper, but let us briefly consider what are perhaps the three greatest difficulties, in order of increasing severity.

The first difficulty is that the Kolmogorov Complexity function $K(s)$ is not computable: briefly, no computer can be programmed to compute the function. This fact may be taken to call into question whether an agent could know $K(s)$ a priori. On the other hand, it is standardly assumed that Bayesian ideal agents know all the logical truths of first-order logic, and these truths are also not computable. Whether there is a deep problem here may perhaps turn on the connection between computability and ideal reasoning.

³ For ease of exposition, Kolmogorov Complexity is defined herein by reference to a *programming language*. Typical expositions of Kolmogorov Complexity define it by reference to a *Universal Turing Machine*. For our purposes, these definitions are equivalent.

The second difficulty is that using Kolmogorov Complexity to constrain priors may implicitly assume that worldly processes are computable, but perhaps some processes are uncomputable.⁴ This problem is standardly handled by assigning a prior of zero to uncomputable processes.

The third difficulty is arguably the most severe. The difficulty is that there might not be a canonical choice of a programming language for the purpose of parameterizing Kolmogorov Complexity. To bring out this difficulty, recall that we calculated $K(s_1)$ and $K(s_2)$ by choosing a “pseudocode” language with a particular syntax. But where did this pseudocode language come from? It resembles commonly used programming languages, but that fact merely reflects contingent human interests and cognitive limitations. We could have chosen a different programming language, such as a programming language which includes a command “twentyzeros ()” which outputs twenty zeros followed thereafter by ones. Using *this* alternative programming language, the shortest program whose output is s_2 is simply “twentyzeros ()”, and therefore $K(s_2) = 13$, in which case s_2 is syntactically simpler than s_1 . So, a lack of a canonical choice of programming language threatens the objectivity of Kolmogorov Complexity as a measure of syntactic simplicity.

Whether these difficulties can be resolved is beyond the scope of this paper. But as we are about to see, the Lewisian method *requires* a solution to the problem of measuring syntactic simplicity. Therefore, to be charitable to the Lewisian, let us suppose that this problem can indeed be solved, and that we have at our disposal a good measure of syntactic simplicity.

3 The Lewisian method

This section summarizes Bradley’s proposed Lewisian method for using naturalness to constrain priors based on descriptions of possible worlds. Bradley describes this method briefly, so this summary of Bradley’s description will perforce be doubly terse.

Bradley assumes that there are “perfectly natural properties” upon which everything supervenes (p. 180). Hereafter let us drop the word “perfectly” for ease of exposition; Bradley does not develop the distinction between perfectly natural properties and other natural properties, and nothing herein turns on this distinction. Rather than defining natural properties, Bradley refers primarily to Lewis (1983), although Bradley does mention that he considers “plausible candidates” for natural properties to be the properties of fundamental physics (ibid.). Beyond this metaphysical commitment, Bradley takes on the following further epistemological commitment:

“A PRIORI: Given a set of properties instantiated in some world w , it is a priori which of the properties are perfectly natural in w .”⁵ (Ibid.)

How does the existence and a priori identifiability of natural properties help constrain priors? Bradley proposes a method inspired by Lewis for using naturalness to constrain priors based on eligible descriptions of possible worlds. The method may be illustrated

⁴ For example, Copeland (2002, pp. 481–482) considers the possibility that the sequence of digits of the magnitude of a physical quantity is measurable but uncomputable.

⁵ The text in small caps is in italicized proper case in the original.

as follows. Consider worlds in which an agent observes emeralds. Let the symbols “g”, “b”, “ γ ”, and “ β ” be shorthand for statements that an emerald is, respectively, green, blue, grue, and bleen. Here, grue means green if observed among the first twenty observations and otherwise blue, and bleen means blue if observed among the first twenty observations and otherwise green. We can then describe a world w_1 with the sequence s_{w_1} , which consists of twenty grue emeralds followed thereafter by bleen emeralds, and we can describe a world w_2 with the sequence s_{w_2} , which consists of twenty green emeralds followed thereafter by blue emeralds:

$$s_{w_1} : \langle \gamma, \beta, \beta, \beta, \dots \rangle$$

$$s_{w_2} : \langle g, b, b, b, \dots \rangle$$

In w_1 , all the emeralds are green, and in w_2 , the first twenty emeralds are green, and subsequent emeralds are blue. So, the Lewisian’s goal is to constrain priors so that w_1 is assigned a higher prior than w_2 . Note that s_{w_1} and s_{w_2} are syntactically isomorphic, so how will the Lewisian tell them apart? By itself, a measure of syntactic simplicity such as Kolmogorov Complexity cannot help. The Lewisian’s answer is as follows.

Let us suppose that, in accordance with A PRIORI, the Lewisian determines that green and blue are natural properties, and that grue and bleen are not natural properties. The Lewisian method requires that descriptions of possible worlds are written using only predicates denoting natural properties. Such descriptions are *eligible*, whereas descriptions that use non-natural properties are *ineligible*. By enforcing the rule that descriptions must be eligible, the sequences s_{w_1} and s_{w_2} are rewritten as follows:

$$s_{w_1} : \langle g, \dots \rangle$$

$$s_{w_2} : \langle g, b, b, b, \dots \rangle$$

In rewriting the description of w_1 , the only change is that the description now only uses predicates denoting natural properties.

The final step is to assign priors to eligible descriptions in proportion to their syntactic simplicity.⁶ At this juncture, Bradley in effect merely asserts that the rewritten s_{w_1} is syntactically simpler than s_{w_2} . Bradley does not provide justification for this claim, but we charitably assumed (in Sect. 2) that we have at our disposal a good measure of syntactic simplicity—and any such measure would yield the verdict that s_{w_1} is syntactically simpler than s_{w_2} . Given this verdict, we thereby arrive at the desired result that w_1 is assigned a higher prior than w_2 .

3.1 Two worries about the Lewisian method

One might worry: is the Lewisian method *feasible*? This worry may arise because no ordinary human could complete the Lewisian method, and yet many ordinary humans

⁶ Descriptions of equal syntactic simplicity are therefore assigned equal priors, which arguably embeds a principle of indifference in the Lewisian method. Thanks to an anonymous reviewer for drawing attention to the role of principles of indifference in this context.

seem to perform many justified inductive inferences. But this worry is not specific to the Lewisian method: it applies to any theory of “ideal epistemology”⁷ which entails that humans are not rational. For present purposes, let us bracket generic worries about theories of ideal epistemology. Plausibly, if the Lewisian method, or any method, succeeds in constraining priors, further research in “non-ideal epistemology”⁸ could focus on whether or how such a method bears on human rationality.

One might also worry: *in virtue of what* does the Lewisian method succeed in constraining objective priors? Bradley claims that it is *primitive* that rationality is grounded in having priors that privilege simpler eligible hypotheses. No further fact explains this grounding principle.

Bradley makes a helpful distinction between metaphysical fundamentality and epistemic fundamentality as a basis for privileging hypotheses:

- “(a) Epistemic fundamentality. Hypotheses that postulate uniformity with respect to epistemically fundamental concepts are privileged;
- (b) Metaphysical fundamentality. Hypotheses that postulate uniformity with respect to metaphysically fundamental properties are privileged...” (p. 185)

The Lewisian method is based on metaphysical fundamentality. We now turn to a rival method for constraining priors based on epistemic fundamentality.

4 The Neo-Carnapian method

This section presents the Neo-Carnapian method for constraining priors. Like the Lewisian method, the Neo-Carnapian method ends with measuring syntactic simplicity, but it begins at a different starting point. The Neo-Carnapian method is so named because it is similar to a “Carnapian” approach to induction, as discussed later in Sect. 5.1.

The key ideas behind the Neo-Carnapian approach are as follows. Bayesian methods update credences based on evidence. The notion of evidence that is relevant for the discussion that follows is the evidence which one gains directly from experience. It will be useful to refer to an agent’s *sequence of experiences*, which is the agent’s finite set of experiences in chronological order through a given time and which may be referred to as $\langle e_1, e_2, \dots, e_n \rangle$.

Some experiences may be *indistinguishable* for an agent. There are many ways to precisify the relevant notion of indistinguishability. For present purposes, rather than commit to any particular precisification, let us rely on an intuitive gloss. Roughly speaking, two of an agent’s experiences are indistinguishable iff were the experiences to be swapped, the agent’s subjective state at any given time would be unchanged.

The *pattern* of an agent’s sequence of experiences is a sequence of numbers which can be generated from the agent’s sequence of experiences as follows. Record the agent’s first experience with the number zero. For the n^{th} experience, where $n > 1$, consider whether the n^{th} experience is indistinguishable from any past experience. If

⁷ As in Carr (2022).

⁸ Ibid.

yes, record the same number that was recorded for that past indistinguishable experience. If no, record the largest number yet recorded plus one.

Let us consider how this works for the case of the emeralds discussed in Sect. 3. First, suppose that an agent's sequence of experiences consists entirely of experiences as of seeing a green emerald. Then the first entry of the pattern of this sequence of experiences is zero, and since each of the subsequent experiences are indistinguishable from this first experience, it follows that each of the subsequent entries of the pattern of this sequence of experiences is also zero. We may therefore record this pattern as follows:

$$p_1 : \langle 0, \dots \rangle$$

Next, suppose instead that an agent's sequence of experiences consists of twenty experiences as of seeing a green emerald, followed thereafter by experiences as of seeing a blue emerald. Then the first entry of the pattern of this sequence of experiences is zero; each of the following nineteen entries of the pattern is also zero, since the second through twentieth experiences are indistinguishable from the first experience; the 21st entry of the pattern is one, since this experience is not indistinguishable from any previous experience, and so we record the largest number yet recorded (namely, zero) plus one; and each subsequent entry of the pattern is one, since the 22nd and subsequent experiences are indistinguishable from the 21st experience. We may therefore record this pattern as follows:

$$p_2 : \langle 0, 1, 1, 1, \dots \rangle$$

The final step is to constrain priors based on the syntactic simplicity of the patterns. At this juncture, the Neo-Carnapian may proceed in different ways, one of which is simply to follow the Lewisian's lead. The Lewisian assigns priors in proportion to the syntactic simplicity of eligible descriptions. Accordingly, the Neo-Carnapian may assign priors in proportion to the syntactic simplicity of patterns. Plainly, p_1 and p_2 are symbolically isomorphic to the Lewisian's re-written s_{w1} and s_{w2} , respectively. We thereby arrive at the desired result that the hypothesis that the agent's sequence of experiences is encoded by p_1 is assigned a higher prior than the hypothesis that the agent's sequence of experiences is encoded by p_2 .⁹

4.1 The problem of Grue

Let us consider potential worries about whether the Neo-Carnapian method handles the problem of grue. An immediate worry might arise from the observation that in the

⁹ In following the Lewisian's lead, the Neo-Carnapian arguably embeds a principle of indifference as described in footnote 7. Does using a principle of indifference to allocate priors to experiences with equally simple patterns require a priori knowledge of all possible experiences? If so, such a priori knowledge may be required of both the Lewisian and Neo-Carnapian agents, since each countenances distinct experiences associated with equally simple descriptions. Whether such a priori knowledge is required, and whether this depends on how syntactic simplicity is used to constrain priors, is beyond the scope of this paper. Thanks to an anonymous reviewer for drawing attention to this topic.

case of the emeralds described above, the experiences were characterized as of seeing green or blue emeralds. But couldn't the experiences have been characterized just as well as of seeing grue or bleen emeralds?

The answer is yes. But this fact is irrelevant to the Neo-Carnapian method, since this method does not rely on such characterization. The only relevant facts about experiences are which experiences are indistinguishable from which other experiences, and these facts are not a function of how each experience is characterized. In the case of the emeralds, *prima facie*, a human's experience as of seeing a green emerald is indistinguishable from a human's experience as of seeing a grue emerald before the relevant time t , and this fact does not depend on whether the human characterizes such emeralds as green or grue.

This point may be expanded as follows. Suppose a human has a sequence entirely of experiences as of seeing green emeralds—or, what is the same thing, a sequence of twenty experiences as of seeing grue emeralds followed thereafter by experiences as of seeing bleen emeralds. In this case, each of the human's experiences would be indistinguishable from each of the human's other experiences. Even if each experience is labeled by the human—as “green” or “grue”, or with numbers, or with names like “Alice” or “Bob”—such labeling does not change the fact that the experiences are indistinguishable in the relevant sense. Therefore, as per above, the Neo-Carnapian method requires that the human have a higher prior in the hypothesis that their future experiences will be indistinguishable from their first twenty experiences than in the hypothesis that their future experiences will not be indistinguishable from their first twenty experiences. After the first twenty experiences, only experiences as of seeing green emeralds and not experiences as of seeing grue emeralds would be indistinguishable from the first twenty experiences. Thus, the human should have a higher prior in the hypothesis that their future experiences will be as of seeing green emeralds than in the hypothesis that their future experiences will be as of seeing grue emeralds.

Another worry might arise from the observation that the foregoing refers specifically to humans. Perhaps green and grue experiences before time t are indistinguishable *for humans*, but shouldn't the Neo-Carnapian method work for *any* epistemic agent? What if, unlike humans, an epistemic agent could, in some sense, *see grue*? Let us consider some extra-terrestrial creatures to reveal that the Neo-Carnapian method succeeds in constraining priors for them.

Suppose that we live in a world in which some objects really are grue. Suppose further that unlike humans, Martians cannot have experiences as of seeing green or blue, but they do have experiences as of hearing a beep or silence, and suppose that Martians have an experience as of hearing a beep iff they encounter an object that is grue, and they have an experience as of hearing silence iff they encounter an object that is bleen. Now suppose that a Martian has twenty experiences as of hearing a beep. Applying the Neo-Carnapian method to this case yields the verdict that the Martian should strongly prefer the hypothesis that each of their future experiences will be as of hearing a beep over the hypothesis that each of their future experiences will be as of hearing silence. But since the Martian's future experiences will be as of hearing a beep iff they encounter a grue object, the Neo-Carnapian method requires that the Martian prefer the hypothesis that their future encounters will be with grue objects.

But, actually, this verdict aligns perfectly with intuition. The Martian has heard twenty beeps, and it seems intuitively correct that the hypothesis that the beeps will continue indefinitely is better than the hypothesis that the beeps will permanently cease. The stipulated fact that the beeps are correlated in some way with experiences had *by humans* is wholly irrelevant.

Is there perhaps a nearby variant of this case which is problematic for the Neo-Carnapian method? Consider a Venusian who has experiences as of seeing green and blue just as humans do and who *also* has experiences as of hearing a beep or silence just as Martians do. Suppose that all emeralds on Venus are grue and presently green, and suppose that the Venusian has twenty experiences as of seeing a green emerald and hearing a beep. Applying the Neo-Carnapian method to this variant yields the verdict that the Venusian should strongly prefer the hypothesis that each of their future experiences will be as of seeing a green emerald and hearing a beep over the hypothesis that each of their future experiences will be as of seeing a blue emerald and hearing a beep. But the Venusian would favor the wrong hypothesis: since emeralds on Venus are grue, each of the Venusian's future experiences will be as of seeing a blue emerald, not a green emerald.

But, actually, as before, this verdict aligns perfectly with intuition. The Venusian has had twenty experiences as of seeing a green emerald and hearing a beep, and it seems intuitively correct that the hypothesis that the same experience will continue indefinitely is better than the hypothesis that a new experience will occur and continue indefinitely. The fact that the Venusian favors the wrong hypothesis is wholly irrelevant; the goal of constraining priors is to yield *justified* inductive inferences, not *infallible* inductive inferences.

4.2 Two worries about the Neo-Carnapian method

Let us consider two other worries about the Neo-Carnapian method. As will become clear, these worries parallel the worries raised in Sect. 3.1 about the Lewisian method.

One might worry about the fact that most if not all ordinary humans have perhaps never had two entirely indistinguishable experiences during their lifetime. The pattern of such humans' experiences is therefore $\langle 1, 2, 3, \dots, n \rangle$. For this pattern, the Neo-Carnapian method presumably strongly prefers the hypothesis that every future experience will continue to be unique. Such a prediction cannot ground common inductive inferences made by humans that seem to be justified. For example, the inductive inference that the sun will rise tomorrow seems justified even though no human has ever experienced two indistinguishable sunrises.

For present purposes, we may set aside this worry just as we set aside the worry about the feasibility of the Lewisian method. Like the Lewisian method, the Neo-Carnapian method is a theory of ideal epistemology which entails that humans are not rational, and so the Neo-Carnapian is not bothered by what *seem* like justified inductive inferences made by humans. Let us therefore continue bracketing generic worries about theories of ideal epistemology.

One might also worry: *in virtue of what* does the Neo-Carnapian method succeed in constraining objective priors? Here too, the Neo-Carnapian response parallels the

Lewisian response: the Neo-Carnapian may claim that it is *primitive* that rationality is grounded in having priors that privilege hypotheses of simpler patterns. No further fact explains this grounding principle.

5 The Neo-Carnapian method is more attractive than the Lewisian method

We charitably assumed (in Sect. 2) that we have at our disposal a good measure of syntactic simplicity. This section maintains this assumption to present an argument that *given* such a measure, the Neo-Carnapian method is more attractive than the Lewisian method. The argument proceeds in two parts. First, it is argued that the two methods' foundational claims are comparably plausible. Second, it is highlighted that the Lewisian method's metaphysical and epistemological commitments are arguably onerous. Given that the Neo-Carnapian method is unencumbered by the Lewisian method's arguably onerous commitments and given that the two methods are otherwise comparably plausible, the Neo-Carnapian method is more attractive than the Lewisian method.

5.1 The methods' foundational claims are comparably plausible

The foundational claim for each method is that it is *primitive* that rationality is grounded in their method for constraining priors. The Lewisian claims that it is primitive that rationality is grounded in having priors that privilege simpler eligible hypotheses; the Neo-Carnapian claims that it is primitive that rationality is grounded in having priors that privilege hypotheses of simpler patterns; both claim that no further fact explains their grounding principle. *Prima facie*, these two grounding claims are comparably plausible. This is because both claims favor descriptions that are syntactically simpler, and it seems comparably plausible to claim that the relevant descriptions derive from a basic metaphysical notion (of naturalness) or from a basic epistemological notion (of indistinguishability).

However, Bradley presents two foundational objections to a rival "Carnapian" approach to constraining priors which Bradley takes to be an exemplar of approaches which privilege epistemic fundamentality (pp. 187–188). Since the Neo-Carnapian method may be thought of as also privileging epistemic fundamentality, it is worth considering whether Bradley's objections succeed against the Neo-Carnapian method. It is argued here that the Neo-Carnapian method is immune to Bradley's objections because the Neo-Carnapian method differs importantly from the Carnapian approach.

The Carnapian approach described by Bradley can be summarized as follows. (i) The approach favors hypotheses that are syntactically simple. (ii) To get around the problem of syntactically simple hypotheses containing gruesome predicates, the approach rules out gruesome predicates because they are not epistemically "primitive." The idea is roughly that epistemically primitive predicates belong to sensory modalities, whereas gruesome predicates are an illicit mixture of sensory modalities and a

temporal modality.¹⁰ (iii) The approach accepts that objective priors are a function of an agent's sensory modalities. This is because the fact of which predicates are epistemically primitive is a function of the agent. For example, a non-sighted echolocating organism may have epistemically primitive predicates belonging to the echolocation modality but not to the visual modality. Therefore, the Carnapian approach's constraints on objective priors are "intrapersonal" rather than "interpersonal".¹¹

Bradley's objections to the Carnapian approach do not target (i); both the Lewisian method and the Carnapian approach favor syntactically simpler hypotheses. Neither do Bradley's objections target (ii); Bradley takes epistemic primitiveness and metaphysical naturalness to each provide a "plausible" basis for privileging the property green over the property grue (p. 185). Rather, Bradley's objections target (iii). Bradley objects that the Carnapian approach "seems to lead to claims which are somewhat implausible" (p. 187). Bradley's first objection is as follows:

"[The Carnapian approach] requires that we have a priori access to which concepts are epistemically fundamental. And it seems that we have no such access – for example, we cannot learn a priori that we can see colour rather than only seeing black and white. We have to open our eyes and have experiences to discover that we can see colours, and thus discover what our epistemically fundamental concepts are." (Ibid.)

Is this objection applicable to the Neo-Carnapian method? The answer is no. The reason is simply that the Neo-Carnapian method does not require that an agent know a priori which sensory modalities are their own. This point can be verified simply by reviewing the Neo-Carnapian method as described in Sect. 4. There is no step in which the Neo-Carnapian method invokes an agent's a priori knowledge of their sensory modalities. This is a distinguishing feature of the Neo-Carnapian method as compared to the Carnapian approach: no such a priori knowledge is required. Instead, an agent learns over time which sensory modalities they have.

Bradley's second objection is as follows:

"[The Carnapian approach] implies that if our epistemically fundamental concepts changed, our priors should change. To make this vivid, imagine two agents with no experiences discussing what their priors should be. (Perhaps two fetuses communicate telepathically in the womb.) Initially they have the same sensory organs, and so agree about what their priors should be. Then one of them grows an extra receptor which allows them to see infrared. According to the current theory, they should change their priors. But why should this change to their eyes change their priors about what the world is like?

"We don't really need agents with no experiences to make the point. Imagine that you suddenly grow infrared receptors. According to the current theory, this should alter your priors..." (Ibid.)

Strictly speaking, a fetus cannot have inspected their sensory organs, thereby determined something about their sensory modalities, discussed the same with a fellow

¹⁰ Bradley cites Carnap and Jeffrey (1971, pp. 70–73).

¹¹ Bradley cites Kelly (2014).

fetus, and yet have had “no experiences”. But Bradley accepts, in the second paragraph, that the fetuses’ not having experiences is inessential. Let us therefore accept the case with the amendment that the fetuses have had experiences. Does the objection survive as an objection to the Neo-Carnapian method? The answer is no. A fetus who has evidence on the basis of which they inductively infer that they have grown an infrared receptor does not change their priors. Instead, the fetus conditionalizes on the evidence, thereby increasing their credence in the belief that they will soon have an experience that is not indistinguishable from any previous experience.

An agent who has only had indistinguishable experiences will favor the hypothesis that the next experience will be indistinguishable over the hypothesis that the next experience will be different. Similarly, given the right evidence, the same agent will favor the hypothesis that the next experience will *not* be indistinguishable from any previous experience. Whether such a hypothesis is rationally favored depends on which evidence the agent has. The case of the fetus suggests that the evidence is sufficient to justifiably infer that the next experience will be new. The Neo-Carnapian method allows for this possibility.

It therefore appears that Bradley’s foundational objections to the Carnapian approach do not succeed against the Neo-Carnapian method. The *prima facie* verdict that the foundational claims of both methods are comparably plausible seems to be secure.¹²

5.2 The Lewisian method’s commitments are arguably onerous

Suppose that the foundational claims of the Neo-Carnapian and the Lewisian are comparably plausible, as argued above. Given that we have at our disposal a good measure of syntactic simplicity, the Neo-Carnapian’s work is finished. They do not need to take on any further commitments to constrain priors.¹³

By contrast, the Lewisian has much more work to do. The Lewisian method calls for the existence and a priori identifiability of natural properties upon which *everything* supervenes. These metaphysical and epistemological commitments are arguably onerous, for reasons that include the following. First, the notion of metaphysical naturalness has been criticized as obscure. Writers have variously termed naturalness

¹² It might be viable to constrain priors via indistinguishability while also endorsing a foundational claim of metaphysical fundamentality. For example, a proponent of Hildebrand’s (2016) “qualitative quidditism” could argue that indistinguishability is fundamentally a metaphysical notion: agents discern the indistinguishability of two experiences through an epistemic capacity that detects qualitative natural properties. Such a qualitative quidditist could nonetheless help themselves to constraining priors via indistinguishability exactly as described in Sect 4. If this view is viable, then the Lewisian and Neo-Carnapian proposals are closer than they might appear. Addressing the viability of such a view is beyond the scope of this paper. Thanks to an anonymous reviewer for calling attention to the possibility of a view of the kind described in this paragraph.

¹³ Is a priori knowledge required to enable a process of abstraction for discerning indistinguishable experiences? See Hildebrand (2016, p. 521) for speculations about processes of abstraction that lead to direct acquaintance with certain properties. Thanks to an anonymous reviewer for calling attention to this line of inquiry.

“problematic,” “mystifying,” have sought to systematically replace the notion, and have suggested that it be given up entirely.¹⁴

Second, the Lewisian has an epistemological commitment to the a priori identifiability of the natural properties. But how should an agent go about identifying these properties a priori? As noted in Sect. 3, Bradley takes plausible candidates for natural properties to be the properties of fundamental physics. But it is not obvious how to identify the properties of fundamental physics given a complete description of a world. Bradley in effect concedes that there is a problem here:

A question remains about exactly *how* we ascertain which... properties are natural. I think this question has not received enough attention and I hope to address it elsewhere. (p. 181)

Promissory notes are inevitable in philosophical investigation.¹⁵ However, it is perhaps fair to say that here we may have a hefty *additional* promissory note for which the Objective Bayesian should prefer not to be on the hook.

5.3 Recap

The Neo-Carnapian and Lewisian methods have comparably plausible foundational claims. And both methods require a solution to the problem of measuring syntactic simplicity to achieve their goal of constraining priors. But the Lewisian method requires much more: they are on the hook to provide a complete theory of the natural properties on which everything supervenes and an a priori methodology for identifying such properties. The Neo-Carnapian method avoids these arguably onerous metaphysical and epistemological commitments of the Lewisian method. In effect, the Neo-Carnapian method does the heavy lifting of constraining priors using just syntactic simplicity. For this reason, the Neo-Carnapian method is more attractive than the Lewisian method.

6 Conclusion

Throughout this paper, we charitably assumed that the problem of measuring syntactic simplicity can be solved. This paper demonstrates that in this case, the Neo-Carnapian method is an elegant way to parlay a solution to the problem of measuring syntactic simplicity into a solution to the problem of constraining priors. This is because the Neo-Carnapian method requires little by way of metaphysical or epistemological commitments. That the Neo-Carnapian method requires few commitments is the reason that it is more attractive than the Lewisian method, as argued above.

If we revoke our charitable assumption and instead deem that the problem of measuring syntactic simplicity cannot be solved, then both the Lewisian and Neo-Carnapian

¹⁴ These four examples correspond, respectively, to Loewer (2007), Witmer et al. (2005), Eddon and Meacham (2015), and Thompson (2016).

¹⁵ See Bradley (2023) for a sketch of a view relating to this question. The extent to which this sketch discharges the promissory note is beyond the scope of the present paper.

methods are sunk. But in this case, the Lewisian method is a fortiori not the most attractive method for constraining priors. It thus follows that *regardless* of whether the problem of measuring syntactic simplicity can be solved, the Lewisian method is not the most attractive method for constraining priors. This by itself is a notable result given the popularity of trying to solve the problem of grue using naturalness.

Acknowledgements For helpful comments and discussion, the author thanks Andrew Dennis Bassford, Josh Dever, Sinan Dogramaci, Robert C. Koons, Miriam Schoenfield, an audience at the University of Texas at Austin, and anonymous referees.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

References

- Bradley, D. (2020). Naturalness as a constraint on priors. *Mind*, *129*, 179–203. <https://doi.org/10.1093/mind/fzz027>
- Bradley, D. (2023). Can we have justified beliefs about fundamental properties? *The Philosophical Quarterly*, *74*, 46–67. <https://doi.org/10.1093/pq/pqad012>
- Carnap, R., & Jeffrey, R. C. (Eds.). (1971). *Studies in Inductive Logic and Probability*. University of California Press.
- Carr, J. R. (2022). Why ideal epistemology? *Mind*, *131*, 1131–1162. <https://doi.org/10.1093/mind/fzab023>
- Copeland, B. J. (2002). Hypercomputation. *Minds and Machines*, *12*, 461–502. <https://doi.org/10.1023/A:1021105915386>
- Eddon, M., & Meacham, C. J. G. (2015). No work for a theory of universals. In J. Schaffer & B. Loewer (Eds.), *A companion to David Lewis* (pp. 116–137). Wiley-Blackwell.
- Goodman, N. (1955). *Fact, fiction, and forecast*. Harvard University Press.
- Hedden, B. (2016). A defense of objectivism about evidential support. *Canadian Journal of Philosophy*, *45*, 716–743. <https://doi.org/10.1080/00455091.2015.1123038>
- Hildebrand, T. (2016). Two Types of Quidditism. *Australasian Journal of Philosophy*, *94*, 516–532. <https://doi.org/10.1080/00048402.2015.1112418>
- Kelly, T. (2014). Evidence can be permissive. In M. Steup & J. Turri (Eds.), *Contemporary debates in epistemology* (pp. 298–311). Wiley-Blackwell.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, *61*, 343–377. <https://doi.org/10.1080/00048408312341131>
- Li, M., & Vitányi, P. (2019). *An introduction to Kolmogorov complexity and its applications* (4th ed.). Springer.
- Loewer, B. (2007). Laws and natural properties. *Philosophical Topics*, *35*, 313–328. <https://doi.org/10.5840/philtopics2007351/214>
- Thompson, N. (2016). Is naturalness natural? *American Philosophical Quarterly*, *53*, 381–396.
- Witmer, G., Butchard, W., & Trogdon, K. (2005). Intrinsicity without naturalness. *Philosophy and Phenomenological Research*, *70*, 326–350. <https://doi.org/10.1111/j.1933-1592.2005.tb00530.x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.