Canadian Journal of Philosophy (2023), 53: 1, 21–37
doi:10.1017/can.2023.21

ARTICLE

Akaike and the No Miracle Argument for Scientific Realism

Alireza Fatollahi

Bilkent University Department of Philosophy Ankara, Turkey
Email: alirezafatollahi@bilkent.edu.tr

Abstract
The "No Miracle Argument" for scientific realism contends that the only plausible explanation for the predictive success of scientific theories is their truthlikeness, but doesn't specify what 'truthlikeness' means. I argue that if we understand 'truthlikeness' in terms of Kullback-Leibler (KL) divergence, the resulting realist thesis (RKL) is a plausible explanation for science's success. Still, RKL probably falls short of the realist's ideal. I argue, however, that the strongest version of realism that the argument can plausibly establish is RKL. The realist needs another argument for establishing a stronger realist thesis.

Keywords: Scientific realism; No Miracle Argument; Akaike Information Criterion; model selection

1. Introduction

Probably no other topic in philosophy of science has been as widely discussed as scientific realism. Despite widespread disagreement on the topic, there is surprisingly near consensus that the strongest argument for scientific realism is the "No Miracle Argument" (NMA) or the "Ultimate Argument." Here is Hilary Putnam's formulation of NMA, often considered to be the first presentation of it in its current form.[1]

> The positive argument for realism is that it is the only philosophy that doesn't make the success of science a miracle. That terms in mature scientific theories typically refer (this formulation is due to Richard Boyd), that the theories accepted in a mature science are typically approximately true, that the same term can refer to the same thing even when it occurs in different theories. (1975, 73)

Scientific realism is usually understood to involve at least two claims. First, a semantic thesis: that scientific terms refer. And second, an epistemic thesis: that the best scientific theories are approximately true. The present essay focuses on the epistemic thesis.

NMA challenges antirealists to offer a plausible explanation for science's marvelous success other than the theory's being truthlike. The argument rests on four assumptions.

(I) Remarkable Success: that the best scientific theories have been remarkably predictively successful.

[1]See also Psillos (1999, 70–75). The "success-to-truth" inference has gone through many modifications in light of serious antirealist challenges. Vickers (2019) contains a nice discussion of some of these developments. Such developments are largely orthogonal to my discussion. My main contentions are equally applicable to Putnam's original argument and, say, the "Qualified Realist Statement" defended in Vickers (2019).

© The Author(s), 2023. Published by Cambridge University Press on behalf of Canadian Journal of Philosophy. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

(II) Remarkable Success requires an explanation.
(III) A good explanation for Remarkable Success is the Realist Thesis: that scientific theories with a track record of predictive success are close to truth.
(IV) Uniqueness: the Realist Thesis is the only plausible explanation for Remarkable Success.

Assessing NMA's plausibility is difficult because of the reliance of premises (III) and (IV) on the notion of approximate truth. As Larry Laudan writes, "the notion of approximate truth is presently too vague to permit one to judge whether a theory consisting entirely of approximately true laws would be empirically successful" (1981, 47). Now, there's been considerable progress in this area since Laudan's (1981) influential paper.[2] However, to the best of my knowledge, we still don't have any account of this notion *that would render NMA plausible.* Let me explain. Following Popper (1963, 234), one can distinguish between two sorts of enterprises. In the semantic enterprise, one tries to offer a satisfactory account of 'closeness to truth,' while in the epistemic enterprise, one tries to decide whether our current theories are indeed close to truth. We have pursued these projects largely independently of one another—a suboptimal methodology in my view. We have a number of options for how to understand approximate truth, but still don't know whether the Realist Thesis understood in terms of any of them constitutes the only plausible explanation of science's success.

To assess NMA's plausibility, we must combine the semantic and the epistemic projects. This is my goal in this essay. I offer a conception of 'approximate truth' that I argue yields the strongest necessary (or the weakest sufficient) version of the Realist Thesis that can explain science's predictive success. Weaken the thesis and success becomes a mere coincidence. Strengthen it and you've added elements that you don't need for a *thorough* explanation of such success.

My suggestion employs a prominent notion of divergence in statistics, called Kullback-Leibler (KL) divergence.[3] The KL divergence of the probability density, g, from another probability density, f, is defined by

$$DKL(f\,||g) = df\,Ef\left[\log\frac{f}{g}\right]. \tag{1}$$

'$E_f[.]$' denotes expectation with respect to f. The KL divergence of g from f can be understood as a measure of how far, on average, g's predictions are from data generated by f. This is the version of the Realist Thesis (hereafter, RKL) that I think plausibly explains Remarkable Success: predictively successful scientific theories are close to truth in the sense that their predictive elements are close to truth in KL divergence.[4] (For simplicity, I sometimes talk about the theory being close to truth in KL divergence or just 'KL-close' to truth as shorthand for its *predictive elements* having this feature.) By the 'predictive elements' of a theory, I mean those general claims within the theory that most immediately result in predictions. Typically, these are obtained from big conjunctions of various parts of the theory (from core theoretic commitments to auxiliary hypotheses and background knowledge about observational apparatus, etc.) together with approximations and estimations. The predictions can be represented by probability distributions for various possible experimental outcomes. It's sensible to envisage them as probability distributions, because observation is subject

[2]See Niiniluoto, Cevolani, and Kuipers (2022) for an overview.

[3]Kullback and Leibler (1951); The idea of using KL divergence as a measure of truthlikeness is advocated in Rosenskrantz (1980), Niiniluoto (1987) and more recently in Niiniluoto (2021). My proposal is different from these in two respects. First, I'm not primarily concerned with the semantic project explicating the meaning of truthlikeness. Second, as I'll explain momentarily, my proposal only asserts that the *predictive elements* of successful theories are KL-close to truth.

In section 4, I discuss an inherent limitation of this idea: strictly speaking, RKL asserts that the theory is close to truth in KL divergence given a known or estimated probability distribution over the independent variable(s).

[4]In section 4, I discuss an inherent limitation of this idea: strictly speaking, RKL asserts that the theory is close to truth in KL divergence given a known or estimated probability distribution over the independent variable(s).

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

to error. Even if we are examining deterministic theories, obtaining any particular observational value is not guaranteed but is more or less likely. This motivates using a probabilistic notion of divergence.

Three important clarifications. First, the debate over scientific realism has often centered around whether inference to the best explanation (IBE) is a plausible form of inference, especially when applied to "unobservables." I don't contribute to this debate. I argue that even granting that IBE is a plausible form of inference and that science can in principle tell us about unobservables, NMA can at most establish RKL. But, as I'll explain, the predictive elements of a theory can be arbitrarily KL-close to truth without its mathematical form, ontological commitments or fundamental ideas remotely resembling the truth. (Ontological claims like "electrons exist" and fundamental ideas like "all physical phenomena that are not due to gravitation, electricity and magnetism can be explained in terms of short-range intermolecular forces," which was the core of what Eugene Frankel has called "the short-range force paradigm" in the beginning of the nineteenth century [1976, 144].) This isn't a defense of antirealism. All I say is that *NMA* isn't a particularly strong argument for realism.

Second, van Fraassen's constructive empiricism claims that "science aims to give us theories which are empirically adequate," where "a theory is empirically adequate exactly if what it says about the observable things and events in this world, is true" (1980, 12). Consider the related (but different) thesis that what best scientific theories say about observables *are* approximately true and call it Realism about Observables (RO). RO, just like the Realist Thesis, relies on the notion of "approximate truth," which is left unspecified. However, even if one understands this notion in terms of KL divergence, the resulting thesis (call it ROKL) wouldn't be equivalent to RKL. Because predictive elements of a theory are *not* (necessarily) its claims about observables. It is true that insofar as the predictive elements tend to be about entities that are closer to the observable end of the observable-unobservable spectrum, RKL *tends to be* about aspects of the theory that concern observables. However, this is mere correlation. Predictive elements can make predictions about, say, phenomena that are too small to detect or outside of our "light cone" to be detectable or simply unobservable given our current observational capabilities. Instead of the observable/unobservable dichotomy, what distinguishes predictive elements of a theory from its fundamental ideas is their level of specificity.[5] The fundamental ideas of a theory typically consist of a few guiding insights that leave many details unspecified. For example, in the early modern period, the idea that all physical phenomena are explainable by sizes, shapes, and motions of the corpuscles of matter involved in them constituted the core of what was called "mechanistic philosophy." Various posits and theoretic commitments needed to be added to this idea for it to lead to any meaningful prediction. The upshot of my argument in this essay is that what explains the success of science is the KL-closeness of such maximally specific claims (predictive elements), but the move from that fact (by an IBE) to the approximate truth of fundamental/ontological ideas of the theory is unwarranted.

Objection: RKL is a version of realism in name only. Most, if not all, scientific realists maintain that science's success is due to the truthlikeness of its theoretic commitments or claims about unobservables, but RKL doesn't entail that. Reply: although I think RKL is a (weak) version of the Realist Thesis, whether one categorizes it as a form of realism is unimportant. My claim is that RKL successfully meets NMA's challenge of explaining science's success and anything stronger than RKL isn't needed for such an explanation. The plausibility of my argument doesn't rest on the name I've chosen for the thesis.

Third, much of the recent debate on scientific realism has focused on various forms of "selective realism," where in response to serious antirealist challenges, realists have offered various criteria for distinguishing between the "idle" and the "working" parts of the theory. The idea is that successful

[5]The exact relation between RKL and ROKL depends on how one disambiguates a technical ambiguity in RKL. I will discuss this in 4.a.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

novel predictions only confirm those elements of the theory responsible for producing those predictions.[6] My argument is orthogonal to that debate. Suppose we grant the selective realist that there is a nonproblematic criterion by which to distinguish between the idle and the working parts. By definition, only the latter "fuel the derivation" of the predictive elements of the theory. In that event, my argument concerns the relation between the predictive elements and the *working* ontological/fundamental ideas of the theory. And the upshot of my argument is this: the success of the predictive elements of a theory doesn't require as an explanation the truthlikeness of its *working* fundamental/ontological ideas.

I appeal to the results in the Akaikean model selection framework in statistics. Model selection is the study of families of statistical hypotheses (called 'models'). Akaike (1973) presented interesting results on how to estimate the comparative KL divergences of models from truth. Since then, there has been considerable progress in this area, making it a powerful tool for a study of the predictive success of theories.[7] Section 2 briefly introduces this framework.[8]

In section 3, I specify the main features of Remarkable Success and describe the strongest plausible version of Remarkable Success I can think of. Thereafter, I assume that this version of Remarkable Success is true in order to show that even this version doesn't require anything stronger than RKL as an explanation. I end section 3 by arguing that RKL explains Remarkable Success and that if RKL is false, Remarkable Success becomes an improbable coincidence.

In section 4, I discuss how strong a version of the Realist Thesis RKL is and conclude that it isn't very strong. I talk about a number of ways in which a theory can satisfy RKL's demands but still significantly differ from the truth.

Section 5 contains the most important part of my argument, where I consider whether NMA can justify a stronger version of the Realist Thesis than RKL. I argue that the most promising realist rejoinder against my account is to argue that RKL *itself* requires an explanation and the only plausible explanation for it is the truthlikeness of the theory's fundamental commitments. I then argue that this is mistaken. I offer an explanation for RKL that doesn't require the truthlikeness of core commitments, which is based on three ideas. The most important of them, which is borrowed from the Akaikean framework, contends that simpler theories that have better fit with the extant

[6]Variants of this idea are championed in Psillos (1999), Kitcher (2001) and Harker (2013). See Chang (2003), Stanford (2006), and Lyons (2006) for antirealist replies.

[7]Forster and Sober (1994) introduced some of these results to the philosophical literature. Sober (1999, 2002) has employed the Akaikean framework to argue for methodological instrumentalism—that science aims at instrumental success. He convincingly argues that if the goal of science was obtaining truth or maximizing probability of truth, certain scientific practices wouldn't make sense. Scientists sometimes don't reject predictively successful theories, despite being almost certain of their falsehoods. I find Sober's argument quite convincing at what it tries to achieve, but I think it wouldn't bother any actual realist. This is because Sober sets up the debate between a version of methodological realism that claims science aims at *exact* truth (or maximizing the probability of being *exactly* true) and a methodological instrumentalism that claims science aims at predictive accuracy. There is no place in his argument for *approximate* truth as the aim of science. I think even the staunchest realist would agree with Sober if these were the only options.

[8]The Akaikean framework recommends maximizing a score called Akaike Information Criterion (AIC). Although I use results about AIC, my argument relies very minimally on the particular form of this score. All my argument needs is that the model which does better in terms of (i) fit with the extant data and (ii) simplicity (in the sense of having fewer adjustable parameters) is KL-closer to truth. (And as I'll explain, in cases with which I am concerned [where data size is very large], when [i] and [ii] are in conflict, [i] has lexical priority.) This claim is applicable to a wide range of cases where AIC itself isn't. For example, apart from methods that are equivalent to AIC for large data (such as cross-validation and AIC_c), even the very different Bayesian method based on Bayesian Information Criterion (BIC) concurs with this claim. (Notice that here I'm only claiming that a wide range of model selection criteria are compatible with the idea that I need for my argument. This is not to say that such criteria are rooted in an a priori conception of simplicity [as paucity of adjustable parameters] or that they balance between goodness-of-fit and simplicity [so understood] for a priori reasons. Each of these criteria specifies a goal and certain assumptions about the inference problem for which it offers a solution. The conception of simplicity as paucity of parameter and the particular trade-off with goodness-of-fit that is recommended by each of these criteria follows from the nature of the inference problem it is designed to solve. I am thankful to an anonymous referee for pointing out the need for this clarification.)

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

data are KL-closer to truth. Therefore, if scientific methodology typically involves selecting simpler theories that have better fit with the available data (I will offer a few reasons that it does), then the methodology of science will typically result in theories that are KL-closer to truth than their predecessors. This idea must be conjoined with two other ideas to establish a full explanation for RKL. (A) that our theory pool doesn't contain too many theories (it's not the case that we are getting KL-closer to truth but never meaningfully close). And (B) that our theory pool contains candidate theories that are meaningfully KL-close to truth. I argue that both (A) and (B) are true about mature scientific disciplines like physics. I then consider the next realist rejoinder: What explains (A) and (B) themselves? After exploring a few possible explanations, I conclude that there are no plausible answers to *this* question that would support anything stronger than RKL. This gets me to the conclusion that NMA can at most give the realist RKL.

2. Akaike and predictive accuracy

Suppose you have a body of data concerning two quantities, X and Y. You know X and Y are linearly related and you wish to find the particular linear function that captures their relation. There is a standard solution to this problem on which there is consensus among statisticians. You must find the linear function with maximum likelihood relative to the data, where the likelihood of a hypothesis H given data D, $\mathrm{L}(H, D)$, is defined as the probability (or probability density) of obtaining D conditional on H being true; that is, $\mathrm{L}(H, D) =^{\mathrm{df}} \mathrm{P}(D|H)$.

What if the inference problem involves choosing between members of different models, for example between the family of linear functions and the family of parabolic functions? (Hereafter, I'll refer to families of hypotheses as 'models,' and I'll reserve the term 'hypothesis' for particular members of models.) Here you cannot simply maximize likelihood. If you do, you will almost always choose a parabolic function (or generally the most complex model), even if the true relation is linear. This is due to the existence of noise in data. The more complex a model is, the more freedom it has to fit the noise, and, therefore, the more likely it is to fit the noise instead of the main pattern in the data. This is called 'overfitting' the data. Akaike's framework offers a solution for how to avoid overfitting if one's aim is to estimate (and maximize) the predictive accuracies of models—or equivalently to estimate (and minimize) their relative KL divergences from truth.

A few notational conventions. I will denote the truth by f. By *the truth*, I mean the true distribution of observed values. If the object of inquiry is the relation between two variables, X and Y, whose (unknown) true relation is $y = D(x)$ with an observational error, E, f is captured by the equation, $y = D(x) + E$.[9] Notice that whether or not $D(x)$ is deterministic, f is always probabilistic because E is probabilistic. Consider a model, $F(\theta)$, defined over a parameter space Θ, ($\theta \in \Theta$). For example, if F is the family of linear functions with a normally distributed error with the unknown variance, σ, F can be characterized by $F{:}\{y = ax + b + \mathrm{N}(0,\sigma),\ a,b \in \mathrm{R},\ \sigma \in \mathrm{R}^{+}\}$. Here the parameter space of F is three dimensional. Since F is understood as a family of probability functions over various possible values of X and Y, we can associate a likelihood function to it, F: $\mathrm{L}(\theta) = \mathrm{L}(\theta,\mathbf{z}) = g(\mathbf{z}|\theta)$, where $g(\mathbf{z}|\theta)$ is the probability density of obtaining the data set $\mathbf{z} = \{(x_1, x_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ conditional on θ being the true parameter value. The value of θ for which $\mathrm{L}(\theta)$ is maximized is called the Maximum Likelihood Estimate (MLE) of F and I denote it by $\widehat{\theta}$.

[9]There is a sense in which $y = D(x)$ may be called the truth. I use the term differently (as is customary in model selection literature), because I'm interested in the *true distribution of observed values.* Nothing substantial hinges on this terminological issue.; Notice that f is defined as the *true* distribution of observed values. When people talk about "error," they often mean the difference between the true value and the value predicted by our theory. In that common usage of error, not every error is observational. It can be due to the fact that the theory is different from the true hypothesis. But here, since f just *is* the (unknow) true distribution, that kind of error doesn't arise. I'm thankful to an anonymous referee for asking me to clarify this point.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

Equation (1) defines the KL divergence of a fully specified probability distribution from another. The KL divergence of a *model* from the truth is defined as the average (with respect to data) KL divergence of the MLE of the model. That is,

$$\{DKL(f\,||F(\theta))\} = df\,Ey\left\{DKL\left(f\,||F\left(\widehat{\theta}(\mathbf{y})\right)\right)\right\} = EyEx[\log(f(\mathbf{x}))] - EyEx\left[\log\left(g\left(\mathbf{x}|\widehat{\theta}(\mathbf{y})\right)\right)\right]. \tag{2}$$

Here $g(x|\widehat{\theta}(\mathbf{y}))$ is the likelihood (relative to data set $\mathbf{x}$) of the MLE of F with respect to the data $\mathbf{y}$. The last equality holds because of (1) and the fact that $\log(f/g) = \log(f) - \log(g)$. $E_{\mathbf{y}}E_{\mathbf{x}}[\log(f(\mathbf{x}))]$ is an unknown constant that depends solely on f. When we are interested in comparing contender models, this term can be ignored, since it is equal for all models. Therefore, a quantity of interest is:

$$\mathrm{A}(F) = df\,E\mathbf{y}E\mathbf{x}\left[\log\left(g\left(\mathbf{x}|\widehat{\theta}(\mathbf{y})\right)\right)\right]. \tag{3}$$

Malcolm Forster and Elliott Sober have called this quantity the predictive accuracy of F.[10] Suppose you obtain data, $\boldsymbol{y}$, generated by f and use it to determine $\widehat{\theta}(\mathbf{y})$. Then you obtain a new data set, $\boldsymbol{x}$, and measure the fit of $\widehat{\theta}(\mathbf{y})$ relative to $\mathbf{x}$, where fit is measured in terms of the logarithm of likelihood (hereafter, *log-likelihood*). The predictive accuracy of F is the average value—with respect to both $\mathbf{x}$ and $\mathbf{y}$—of this log-likelihood. The predictive accuracy of f is more than any other hypothesis or model. And the KL-closer to f a model is, the better its predictions are.

The goal in the Akaikean framework is to estimate predictive accuracy, which, by equations (2) and (3), is equivalent to the KL divergences of the model from truth modulo a constant, which is equal to the unknown predictive accuracy of f. Akaike offered the Akaike Information Criterion (AIC) of model F as an estimator of its predictive accuracy.

$$\mathrm{AIC}(F,\mathbf{y}) = df\log\mathrm{L}\left(\widehat{\theta}(\mathbf{y})\right) - k \tag{4}$$

Here $\log\mathrm{L}(\widehat{\theta}(\mathbf{y}))$ is the log-likelihood of F's MLE and k is its number of adjustable parameters. Akaike showed that if certain rather nonrestrictive conditions are met, $\mathrm{AIC}(F)$ is an unbiased estimator of $\mathrm{A}(F)$ for large data sets. An estimator is unbiased just in case its average value equals the value it estimates. That is,

$$E\mathbf{y}[\mathrm{AIC}(F,\mathbf{y})] = \mathrm{A}(F) \tag{5}$$

[11]

Equation (5) captures Akaike's main result.

3. Precisifying the realist thesis

Three features of the empirical success of scientific theories are pertinent to NMA. First, our best theories have sometimes fit *large* bodies of data *very precisely*. Second, such theories are not exceedingly complicated— they do not achieve such marvelous fit with data by making increasingly complicated (or ad hoc) modifications. Third, and importantly, scientific theories have been able to predict surprising facts that have nothing to do with the body of observations for the

[10] Actually, they divide this quantity by n, the number of data points to make it independent of data size, which makes sense. For my purposes $\mathrm{A}(F)$ as defined by equation (3) works well enough.

[11] This equation is approximately true. However, the difference between the two sides of the equation quickly goes to zero as data size increases.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

accommodation of which the theories were originally formulated. For example, Fresnel's wave theory of light entailed that if an opaque disk is placed in front of a point source of light, there will be a bright spot at the center of the disk's shadow. That such a bright spot exists wasn't part of the evidence for the accommodation of which Fresnel designed his theory—indeed it was pointed out as an "absurd" consequence of the theory by Poisson. And this makes the theory especially well-confirmed by the observation of the bright spot. Or so it is argued by proponents of predictivism: the thesis that *ceteris paribus*, successful predictions provide stronger evidence for a theory than accommodations.

I believe a very strong version of predictivism can be established using Akaike's results. Accordingly, other things being equal, the model that predicts a body of data has a few units higher AIC score than the model that accommodates that same body of data. (A difference of a few units of AIC amounts to a *significant* difference in estimated predictive accuracy.)[12] However, since I'll argue that NMA cannot establish anything stronger than RKL, *even if such a strong version of predictivism is true*, establishing the truth of this version of predictivism is unnecessary for my argument in the present essay. The reader who maintains a weaker version of predictivism (or rejects it altogether) should find no problem with my assumption of this version of predictivism here.

Remarkably Successful theories are simple and fit the data excellently. Moreover, they have had successful novel predictions. Therefore, they enjoy an amazing balance of goodness-of-fit and simplicity: we can be very confident of their predictive accuracies. What explains this? RKL. By equations (2) and (3), if a theory is KL-close to truth, its predictive accuracy is close to that of f—the highest possible predictive accuracy.[13] If T is KL-close to f, the *expected* (meaning, average) accuracy of T's predictions is high. Thus, it is reasonable to *expect* T to have a track-record of empirical success.

The above explanation might appear tautological, but it's not. RKL concerns a general tendency of a theory. Remarkable Success concerns its actual history. Compare: that the coin I am tossing is fair explains the observation that in the 1,000 times I tossed it, the ratio of heads to tails was nearly 1. However, admittedly the above explanation doesn't offer *much* by way of explaining Remarkable Success. Surely, we'd want a deeper explanation of Remarkable Success. I offer such an explanation by offering an explanation for RKL itself in section 5.

Before ending this section, I'd like to point out that RKL is the weakest sufficient explanation of Remarkable Success in the sense that if RKL is false, then Remarkable Success is an improbable coincidence. This comes directly from the fact that if T is not KL-close to f, its predictions aren't likely to be close to data generated by f.[14]

[12]I have defended this view in Fatollahi (2023).

[13]Even the predictions of f are not entirely accurate because of observational error.

[14]Strictly speaking, this only shows that the probability of Remarkable Success conditional on RKL is high and the probability of Remarkable Success conditional on RKL being false is (very) low. However, the realist ultimately needs an argument that the probability of Realism (or in this case, RKL) conditional on Remarkable Success is high, which doesn't automatically follow from the above argument. For any accepted scientific theory, T, this would follow only if the prior probability of realism (and here RKL) about T is not significantly lower than other theses such as constructive empiricism about T that are compatible with accepting T. (Everyone in this debate agrees that we must accept T as our theory. The question is what exactly is involved in such acceptance.) Champions of NMA have not neglected this aspect of the problem; see for example, Maxwell (1970, 17–18) and Psillos (1999, 72–73). I am skeptical about talk of things like "the prior probability of realism" (and thus about a fully Bayesian treatment of this issue) but *to the extent that it is meaningful and useful to talk about such things,* I find it plausible that the prior probability of RKL about T isn't significantly lower than the prior probability of other theses compatible with accepting T, roughly for reasons offered by Psillos (1999). I am thankful to an anonymous referee for inviting me to clarify this point.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

4. Is RKL realist enough?

4.a The distribution of the independent variable(s)

So far I've argued that RKL explains Remarkable Success. But will the realist find it strong enough? To decide this issue, in this section I consider the ways in which a theory can be KL-close to truth but still different from it.

Suppose truth is represented by f: $y = D(x) + E$, where $D(x)$ is a deterministic function and E is the noise. Our model is g: $y = d(x, \varepsilon)$, where ε is an error term. Importantly, both f and g are *conditional* probability distributions for Y conditional on X. Therefore, RKL is effectively a claim about two conditional probability distributions. This framework must be relativized (in fact it *is* implicitly relativized) to a probability distribution for X.[15] That is, g is KL-close to f *given a known or estimated* $P(X)$. (Hereafter, by 'P(X)' I mean the probability distribution of X relative to which RKL is understood to be true.) Therefore, in principle, g can diverge from f to an arbitrarily large degree for values of X with zero or very small probability according to P(X). That is, RKL can be true of g, if g's predictions are close to data generated by f for certain values of X that "matter," not necessarily the entire range of X. By expanding the range of X that matters, we make RKL stronger and more interesting, but we also make proving it harder.

What determines P(X)? *Minimally*, P(X) must assign nonnegligible probabilities to the observed values of X, otherwise RKL cannot explain Remarkable Success, which pertains to those values of X. Sometimes we know the probability distribution that generated the *observed* values of X (Hereafter '$P_{observed}(X)$.' $P_{observed}(X)$ need not be identical to P(X)). Indeed, in controlled experiments, we decide which values of X are observed. Other times, we have no control over the observed values of X, but we can estimate $P_{observed}(X)$. A sensible way to do so is to take $P_{observed}(X)$ equal to 1/n (n being the size of data) for observed values of X and zero elsewhere. Importantly, in either case, $P_{observed}(X)$ is effectively nonzero only within a range of values of X (hereafter, '$R_{observed}(X)$'). Therefore, if we take P(X) equal to $P_{observed}(X)$, RKL establishes the KL-closeness of g to f only within $R_{observed}(X)$. In that event, anything can happen outside $R_{observed}(X)$, despite RKL being true. There is a familiar example in the history of physics that illustrates this. An excellent explanation for the Remarkable Success of Newtonian mechanics in the nineteenth century was RKL relativized to $P_{observed}(X)$, which was consistent with the later developments in physics that Newtonian mechanics works only for macroscopic phenomena and in small velocities compared to that of light.

RKL limited to $R_{observed}(X)$ is what we minimally need to explain Remarkable Success, but what else can we (nondeductively) infer from it? Sometimes we can infer the following stronger thesis (hereafter, 'RKL^+'). Our theories are KL-close to truth over the entire range of X.[16] The truth of RKL (limited to $R_{observed}(X)$) provides *evidence* for RKL^+, but how could we argue for RKL^+ more conclusively? I can think of two ways. First, consider $R_{observed}(X)$ as a function of time. For the best of our theories, this range has been expanding over time, each time the theory having been successful in fitting the data in the expanded range. The realist might argue *inductively* that since RKL has been true for $R_{observed}(X)(t_1)$, $R_{observed}(X)(t_2)$, $R_{observed}(X)(t_3)$, … it will continue to hold for larger ranges of X. Another way is to claim that RKL^+ is a superior explanation to RKL limited to any finite range, because it doesn't appeal to a *mysterious* range within which the theory is KL-close to truth—any finite choice of range invites the question of why *that* particular range and not another?[17] This second argument has the added benefit that it establishes RKL^+ as part of an explanation for Remarkable Success, i.e., within the context of NMA itself.

[15] See Forster (2002, 98) for discussion.

[16] To make the definition of RKL^+ rigorous, consider a series of probability functions, $P_n(X)$, that are uniform distributions over the interval $I_n = (-n, n)$, $n \in N$. RKL^+ asserts that for all values of n above a certain threshold, the theory is KL-close to truth relative to $P_n(X)$.

[17] Notice the metaphysical nature of the question which asks "Why is g close to truth within $R_{observed}(X)$ but not elsewhere?" The epistemic fact that *we* have data only within $R_{observed}(X)$ is, at least prima facie, irrelevant to this question.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

Two cautionary points are in order though. First, we should always be less confident about extrapolation than about interpolation. Regularities observed in a certain range are more likely to show a different character in other ranges of data than in unobserved values within the same range. Second, the deeper or more fundamental a regularity is, the more confident we can be of its validity for other ranges of data. Phenomenological laws tend to be true for certain range of values. Fundamental laws, by contrast, are more likely to hold universally. This is important because the predictive elements of a theory (with which RKL is primarily concerned) are far more intimately related to phenomenological laws than to fundamental ones.

I think although the arguments in favor of RKL^+ justify a modest confidence in it, our confidence cannot be very high (at least in the absence of further evidence) given the above considerations. Indeed, it appears to me that the realist's most plausible move in face of RKL's dependence on P(X) is to argue that one of the appropriate ways in which a theory can approximate truth is by being (nearly) true in certain ranges of X. This is the sense in which people often say that Newtonian mechanics is approximately true under certain circumstances. However, this admission would seriously limit the realist's ability to justify belief in the near truth of fundamental/ontological commitments of the theory based on Remarkable Success or RKL. Our theory, *g*, and the truth, *f*, can produce nearly identical predictions in limited ranges of values, while their mathematical forms are very different. (A familiar example: $y = x$ and $y = \sin(x)$ are very close for values of x near 0.) As I will argue shortly, when mathematical forms of *g* and *f* are (even slightly) different, the fundamental ideas that are suggested/justified by the two can differ significantly. In 4.b, I will mention a few other (more interesting) ways in which a theory can be KL-close to truth but different from it in more fundamental respects.

Before doing so, however, I'd like briefly to discuss the relation between RKL and ROKL (introduced in the introduction). The definition of RKL leaves the nature of P(X) open. Different ways of understanding P(X), lead to different versions of RKL. For example, if one takes the range for which P(X) is nonzero to be the range of observ*able* values of X, then it becomes ROKL. Thus, typically, RKL limited to $R_{observed}(X)$ is weaker than ROKL (because the entire observ*able* range of *X* isn't typically observ*ed*) and RKL^+ (which is itself a form of RKL) is stronger than ROKL (because the entire range of *X* might not be observable). I have a hard time coming up with any convincing motivation for ROKL. I think depending on the case at hand, RKL limited to $R_{observed}(X)$ or RKL^+ are better motivated. At any rate, what I will discuss in the rest of the essay is applicable to all versions of RKL independently of what one takes P(X) to be.

An anonymous referee pointed out that we don't need to take *f* as the true generating function. Since nothing about the mathematical nature of KL divergence depends on *f* being true, everything in the argument will be exactly the same even if we assumed *f* to be merely empirically adequate. This idea has a kernel of truth in it. It is definitely the case that the *mathematical nature* of KL divergence doesn't require *f* to be true. Indeed, for treating any version of RKL that is weaker than or equivalent to ROKL, this move won't change anything. This is because within the range of observable phenomena, the two ways of understanding *f* are equivalent, since "a theory is empirically adequate exactly if what it says about the observable things and events in this world is true" (van Fraassen 1980, 12). Only when it comes to stronger forms of RKL (like RKL^+) does it matter which of the two understandings of *f* one adopts. The problem with taking *f* to be only empirically adequate is that in that case one cannot even *formulate* any stronger form of RKL than ROKL (regardless of whether one accepts or rejects it). But since my goal is to determine how strong a realist thesis NMA can establish, I'd want to be able to formulate stronger forms of realism (even if I end up not defending them). Moreover, recall that, as I mentioned in the introduction, my main thesis is that *even if one grants the realist that IBE can in principle tell us about unobservables*, NMA cannot give us anything stronger than some form of RKL (where the form depends on how we understand P(X)) the strongest of which is RKL^+. Thus, the nature of my argument demands that I allow every realist-friendly assumption I plausibly can, lest I set the debate up in ways that disfavor

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

realism. Taking f as the true function is one of those assumptions. Notice, however, that I don't assume we *know* anything special about f.

4.b Other divergences

Consider an example in which the error in g is additive (i.e., g: $y = d(x) + \mathcal{E}$) and the distribution of error is the same for f and g (e.g., E and $\mathcal{E}$ both have normal distributions with mean 0 and variance 1). Then RKL is true about g just in case the curve that represents $d(x)$ (the deterministic part of g) is close to the curve that represents $D(x)$ (the deterministic part of f) for those values of X for which P(X) is not extremely small. Similarly, RKL$^+$ is true just in case this holds for (almost) all values of X. Thus, an intuitive way of understanding RKL is this: the curve representing our theory is close (in numerical value) to the curve representing the truth. This way of thinking about KL-closeness helps illustrate how weak a notion of "closeness" it is, because it requires nothing about the mathematical forms of g. The mathematical form of a theory—the quantities it posits and the mathematical relation(s) among them—affects, sometimes significantly, the plausibility of various claims about the causal or explanatory structure of the system under study. Thus, if the realist takes science to tell us about *this structure*, RKL falls far short of her ideal.

Alfonso García-Lapeña (Forthcoming) argues that a theory's truthlikeness requires not only "accuracy" (getting the numbers nearly right) but also "nomicity," getting the shape of the curve roughly right. RKL doesn't require anything about nomicity. This is a potential weakness of RKL that I won't discuss here. Instead, I will mention three other ways in which RKL might be insufficiently strong, which go beyond the combined demands of accuracy and nomicity but which are required for the truthlikeness of fundamental/ontological commitments. (Clarification: I'm not arguing that the realist *must* insist that our theories are similar to truth in any of these forms. She can opt for weaker versions of the Realist Thesis like RKL. I'm only *describing* how strong RKL is.)

Nested Models. Suppose f is $y = 2x + 1 + N(0, 1)$ and we're studying it using polynomial models. Here g might be of the form $y = \delta x^2 + ax + b + N(0, 1)$, where δ is small but nonzero. The realist might think (or hope) that by increasing the data size or by setting more stringent empirical standards on the success of g (for example, by insisting that g must have a track record of successful predictions), one can reasonably expect δ to become zero after a certain point. However, this is unlikely on the assumption of RKL. It is a well-known fact in model selection that if one maximizes AIC, one's hypothesis converges to truth (δ approaches zero) but one's *model* need not be "correct" no matter how large the data is (δ need not be zero for any finite body of data).[18] In other words, no matter how stringent one's standards of empirical success are, the model to which g belongs (the family of quadratic functions) can be different from the smallest model that contains f—the family one typically associates f with (the family of linear functions). This is relevant to RKL because to maximize AIC is to minimize KL divergence. Thus, one can have a theory that is very KL-close to truth, while one is "wrong" about the family to which it belongs. This isn't surprising. If δ is small enough, $y = \delta x^2 + 2x + 1 + N(0, 1)$ can be arbitrarily KL-close to $y = 2x + 1 + N(0, 1)$, thus it can be predictively successful to an arbitrarily high standard.

Now for small enough δ, g is *nearly linear*. So one can argue that g's mathematical form approximates that of f's. However, fundamental/ontological commitments of a theory can be *highly sensitive* to its mathematical form in ways that there might be a big difference between a quadratic but nearly linear function and a linear one. Here is an example. Newton rejected vortex theories of planetary motion in part because he believed vortices can only act on the surfaces of planets. Thus, if gravity was due to a vortex, it should have been proportional to the surface areas of the planets not their volumes (masses). He sometimes favored another explanation according to which gravity was

[18]In the statistical jargon, AIC is said to be an inconsistent estimator *of the number of parameters* in the model (which shouldn't be mistaken for a logical inconsistency in AIC). See section 7 of Forster (2002) for helpful discussion.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

due to the elastic force of an all-pervading ether that interacted with the *entire* body of planets. Now suppose our law for gravity suggested that (instead of being proportional to R^3) gravitational force is proportional to either $R^{3+\delta}$ or $R^3 + R^\delta$, where R is the radius of the planet. For small enough δ, both forms can be arbitrarily KL-close to R^3. However, neither particularly justifies the ether explanation of gravity. $R^{3+\delta}$ seems inconsistent with the ether theory. And $R^3 + R^\delta$ suggests that another mechanism must be producing the R^δ-part. Here we have three functions that can be arbitrarily KL-close but can suggest massively different causal structures.

Spurious Contributions. A special case of the above phenomenon deserves attention. This is when f is a function of one variable, say, $y = ax + b$, but g is a function of two (or more) variables, say, $y = \delta z + cx + d$—where δ remaining nonzero no matter how stringent one's empirical standards are. For small δ, one could argue that the contribution of Z is small or negligible relative to X, but it's never zero and therefore, the ontological/fundamental picture suggested by g, (based on two or more contributing factors) is different from that of f.

Partial Causes. Suppose f has the form $f = f_1 + f_2$. One might hope that if g is "close" to f, it must be explicitly of the form $g = g_1 + g_2$, where g_1 is close to f_1 and g_2 is close to f_2. RKL doesn't satisfy this requirement. For example, one can think of f as the sum of two ordinary trigonometric functions and g as a very good approximation of it by polynomial functions. In that case, one cannot infer from looking at g that f is the sum of two functions. Important corollary: if f_1 and f_2 represent the contributions of two partial causes that conjointly determine a full cause, g gives no clue about the existence of two partial causes, since it isn't explicitly of the form $g = g_1 + g_2$. Thus, we could have an arbitrarily KL-close theory to truth, while we are quite oblivious even to the most rudimentary facts about the metaphysical structure.

5. The realist's rejoinder

Recall that one of NMA's premises was Uniqueness: that the Realist Thesis offers the *only* plausible explanation for Remarkable Success. Having RKL as a potential explanation for Remarkable Success, the problem for the realist is that she cannot easily appeal to NMA to establish a stronger version of the Realist Thesis. She has two options. She can argue that a stronger version is a better explanation of Remarkable Success than RKL. Or, she can argue that RKL itself calls for an explanation and *that* explanation must appeal to a stronger thesis. I won't discuss the first option in detail. As I argued in section 3, if the proposed explanation doesn't entail RKL, then Remarkable Success becomes an improbable coincidence. This makes the first option unlikely to work. But no more could be said about this without knowing the exact content of such a rival explanation.

However, the second option appears promising. Indeed, this is a rather typical argumentative strategy for the realist. For example, Arthur Fine argues that Remarkable Success can be explained by the instrumental reliability of our theories, where 'instrumental reliability' is understood as a kind of "capacity" to produce good predictions.[19] In response, Stathis Psillos writes,

> Although certainly in the right direction, this account is incomplete. Not because there are no dispositions, or powers, in nature, but rather because one would expect also an explanation of why and how theories have such a disposition to be instrumentally reliable … Is it a brute fact of nature that theories—being paradigmatic human constructions—have the disposition to be instrumentally reliable? This seems hardly credible. If dispositions of this sort need grounding, then there is an obvious candidate: the property of being approximately true would ground the power of scientific theories to be instrumentally reliable. Since Fine would certainly deny this account, he owes us an alternative story of how this disposition is grounded. (1999, 90–91)

[19]Fine 1991, 83.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

A virtually identical argument can be offered against RKL. What explains the fact that our theories are remarkably KL-close to truth other than its fundamental commitments being approximately true? RKL *itself* cries for explanation.

Fair enough. I owe the realist an explanation for RKL and I think I have one. Our theories are KL-close to truth because the methodology of science involves (among other things) replacing theories that are less simple[20] and less close to the extant data by those that are more so. Akaike's results tell us that if we follow this methodology, we will continually get KL-closer to truth. Thus, if there are theories in our radar that are meaningfully KL-close to truth, we will sooner or later choose them. Nothing about the truth (or near truth, whatever that means) of fundamental/ontological aspects of our theories is required.[21]

Here the main idea is that scientific methodology involves (among other things) balancing considerations of goodness-of-fit with those of simplicity. To the extent that this is the case, scientific methodology jibes with the Akaikean recommendation. Given Akaike's results, it follows that subsequent theories are typically KL-closer to truth than their predecessors.

Two points must be mentioned here. First, since RKL concerns only the predictive elements of the theory, the above explanation is satisfactory only if considerations of fit and simplicity *of the predictive elements* of the theory play an important role in theory choice. Fortunately, the fit of the predictive elements and the fit of the theory are the same thing—the theory is connected to data via its predictive elements. However, simplicity considerations are not exhausted by those of predictive elements. For example, although *other things being equal,* theories with more parsimonious ontologies have simpler predictive elements (a linear function of one variable is simpler than a linear function of two variables), the ontological parsimony of the theory need not be reflected in the simplicity of its predictive elements. Second, considerations of fit and simplicity ground many other considerations that are important in theory choice. Earlier I claimed that the evidential significance of predictions is grounded in simplicity-favoring considerations. Forster and Sober (1994) argue that unification is similarly grounded.[22] This explanation of RKL works if the totality of such considerations plays a major role in theory choice. I think they do.

Objection: the Akaikean framework recommends optimizing a *very particular* balance between goodness-of-fit and simplicity. Thus, to show that the methodology is responsive to the same considerations doesn't establish that scientific methodology and the Akaikean recommendation are in line. They must assign these considerations the same relative weights. This objection, although strictly speaking true, overlooks an important fact. Scientific inferences, at least insofar as our best theories are concerned, are different from typical curve-fitting or model selection problems in an important respect: they are based on very large data. When data size increases, the first term in AIC ($\log L(L(M))$) is proportional to n, the number of data points. (Because with large n, the average *per datum* log-likelihood is almost constant for different values of n.) However, the penalty term for complexity remains constant. Therefore, considerations of goodness-of-fit dominate those of simplicity. In other words, when data size is large, there is a lexical ordering between the two

[20]Here and throughout the paper, my use of 'simplicity' and 'simplicity-favoring considerations' is concerned only with simplicity as paucity of adjustable parameters—the sense of simplicity that, as Akaike shows, bears on KL-closeness. I adopt this terminology for the sake of convenience. However, I do not intend to suggest that this is the only conception of simplicity that matters to theory choice. For example, Forster et al. (2018) introduce the principle of 'frugality' in connection with causal inference. This is an important sense of simplicity, which is different from paucity of adjustable parameters. I am thankful to an anonymous referee for suggesting that I clarify this point.

[21]A clarification (almost entirely borrowed from White [2003]) about how RKL is understood as the explanandum here. I report the results of White's analysis and refer the interested reader to his essay for the argument. Suppose T is a theory with a remarkable track record of predictive success. That T is KL-close to f is a necessary fact having to do with their contents and doesn't need any explanation. RKL requires an explanation when it is understood as the idea that "*our theory* is KL-close to truth," where "our theory" *nonrigidly* refers to whatever theory we hold. The Akaikean explanation I discuss is an explanation of RKL in this sense.

[22]Notice that the relevant sense of simplicity for all of these cases is paucity of adjustable parameters. See also footnote 20.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

considerations: simplicity-favoring considerations are relevant only when rival theories have (or are supposed to have) equally good fit with the data.

Although it goes well beyond the scope of the present essay to show this conclusively, I think scientific methodology does give simplicity a secondary status. The most famous historical example in which simplicity-favoring considerations supposedly played a major role is the case of Ptolemaic versus Copernican astronomy. In the seventeenth century, when Copernicanism became dominant in scientific circles, there existed a relatively massive body of observational data on the motions of the planets. (Notice that every single precise-enough observation is a datum here.) The two astronomical theories were about equally fit with those observations. Thus, the relative simplicity of the Copernican system could play a decisive role in theory choice.[23]

Examples like this are suggestive but are no proof for my claim. The best argument I can offer (still not a proof) is this. Scientists often adopt a theory despite its inability to accommodate part of observational data in part because it is simple or has other epistemic merits. But they do so believing (justifiably or not) that the discrepancy will be sorted out. That is, they don't *accept* the "errors" in the deliverances of the theory in a permanent fashion. They temporarily allow for those "anomalies" because they are confident in the theory's ability eventually to accommodate them or for the recalcitrant data to turn out to be misleading evidence. If there was an in-principle competition between simplicity and goodness-of-fit, it would have been an acceptable argument that the anomaly exists and *is not going to go away*, but this epistemic vice is counterbalanced by the theory's other merits, like simplicity. This doesn't appear to be a typically acceptable argumentative strategy in science, especially when it comes to our best theories.

The above argument might lead to a misunderstanding. The idea that scientific methodology is roughly in accord with the Akaikean recommendation might appear to suggest a clear defense of methodological instrumentalism (that science aims at instrumental success). The goal in the Akaikean framework is maximizing predictive accuracy—an undoubtedly instrumentalist goal. Thus, to suggest that the practice of science is in accord with the Akaikean recommendation might seem an endorsement of methodological instrumentalism. However, I have argued that scientific methodology jibes with the Akaikean recommendation only insofar as it gives a secondary status to simplicity when data size is large. Interestingly, we have another framework based on Bayesian Information Criterion (BIC) (defined by $\mathrm{BIC}(F) \stackrel{\mathrm{df}}{=} \log \mathrm{L}(L(F)) - (k/2)\log(n)$, where n is the number of data points) that can be interpreted in ways that are harmonious with methodological realism,[24] but similarly assigns a secondary status to simplicity for large data. (When n is large, the first term, which is proportional to n, dominates the second term, which is proportional to $\log(n)$).[25] Thus, my

[23]The case of Copernican versus Ptolemaic astronomies is also one of the best historical examples in which one can see how paucity of adjustable parameters can lead to higher predictive accuracy. The Ptolemaic theory could in principle be made compatible with any astronomical data by the help of more and more epicycles. However, once you select the best fitted Ptolemaic hypothesis given previously acquired data, the fitted hypothesis would do poorly in predicting new data, whereas the Copernican theory was able to predict new data with much higher accuracy. This is exactly what Akaike's theory predicts because for any fixed body of extant data, the best fitted versions of the two theories have about equal fit with that data; but the Ptolemaic theory has far more adjustable parameters. See also Forster and Sober (1994, 14–15). I am thankful to an anonymous referee for inviting me to say more about this case.

[24]BIC is an approximation of the (average) log-likelihood of the model. Assign prior probabilities to your models and multiply them with their BIC scores and you have an approximation of their posterior probabilities (modulo a normalizing constant). The goal in this framework is maximizing model probability, which *can be* interpreted in line with methodological realism. (This is not to say that a proponent of BIC-based model selection must be a realist. I am grateful to an anonymous referee for suggesting that I clarify this point.)

[25]BIC and AIC (and some other model selection criteria, like $\mathrm{AIC_c}$) measure the goodness-of-fit of a model by log-likelihood of its MLE. However, they differ on the relative weights they assign to simplicity (in balancing it against goodness-of-fit). The penalty term for complexity is $(k/2)\log(n)$ for BIC, while it is k for AIC. Thus, assuming the number of data points isn't very low (that is, assuming $\log(n) > 2$), for a fixed n, BIC penalizes complex models more than AIC does (while the weight of simplicity relative to goodness-of-fit diminishes with larger data sets for both AIC and BIC). (See Forster [2002] for a more thorough comparison between AIC and BIC.) I am thankful to an anonymous referee for suggesting that I clarify this.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

observation that simplicity has a secondary status in scientific methodology is compatible with both realist and instrumentalist frameworks in model selection.

This ends my argument for the claim that successive scientific theories are typically KL-closer to truth than their predecessors. This is the main idea in my explanation for RKL, but two more ideas are needed. (A) That there aren't too many theories in our theory pool, so that we could always get KL-closer but never close enough in any meaningful way. And more importantly, (B) that some of our candidate theories for explaining the entirety (or a big part) of the data are meaningfully KL-close to truth to begin with.[26] Neither fact is trivial, but we can be fairly confident about both. We have direct access to the number of theories taken seriously by scientists. Although the number of candidate theories (both at any given time and collectively over time) isn't meager, it's not unmanageable either. So A is true. B is more interesting and more difficult to establish. Indeed, I think B is neither true of all sciences, nor always about any particular science. Sometimes we don't have any good candidate theory that can satisfactorily fit the existing data without unacceptable complexity. In fact, the hitherto falsity of B for certain scientific disciplines, like psychology, might be the reason why they arguably haven't attained the status of mature science. However, when it comes to scientific areas to which our best scientific theories belong, we have direct evidence that some theories in our radar are KL-close to truth, because we have direct evidence that our theory itself is such.

The realist might justifiably ask for an explanation for A and B themselves. Why is it that we are able, at least in certain areas of inquiry, to formulate theories that are KL-close to truth? I think this is a fair question, but one that is unlikely, even prima facie, to help the realist's position. Consider the set of all the relatively simple models describing functional dependencies between the "natural" properties of a system (properties that are describable by predicates like green and blue, as opposed to grue and bleen) and call it M(S), where S is the system under study. We can think of M(S) as the set of our candidate models for S prior to consulting the data at any given point in the development of a science. (It's a mistake to think of M(S) as a fixed set determined a priori. More about this momentarily.) This is because we hardly ever consider hypotheses that are either formulated in terms of functions of great complexity (like the family of polynomials of degree 1000) or in terms of other, "nonnatural" properties like grue. M(S) is huge but it's not unmanageably so, especially given the fact that data typically weeds out an absolute majority of its members. Thus, the fact that M(S) is the set of our candidate models explains A. B can be explained by the fact that there are members of M(S) that are KL-close to truth (given an error distribution). But what explains this fact?

The realist might argue that only the (near) truth of the fundamental/ontological aspects of our theories can explain this. Richard Boyd has offered such an argument. Accordingly, all aspects of scientific method—including which properties of a system are deemed "natural" or the kind of functional dependencies that are considered "simple"are theory-dependent. Therefore, the judgment to include a given function in M(S) is itself theory dependent. Boyd adds,

> The only scientifically plausible explanation for the reliability of a scientific methodology that is so theory dependent is a thoroughgoingly realistic explanation: scientific methodology, dictated by currently accepted theories, is reliable at producing further knowledge *precisely because, and to the extent that, currently accepted theories are relevantly approximately true.*" (1990, 362)

I agree that the determination of the members of M(S) (among other aspects of scientific methodology) is a theory-dependent judgement. However, I think the move from this idea to "a

[26]It is important to bear in mind that (B) concerns only the KL-closeness of those theories that are sophisticated enough to be candidates for explaining (most of) the phenomena in a given scientific discipline. Otherwise, (B) becomes trivial, since there are many trivial theories we can come up with that are very KL-close to truth or outright true. I am grateful to an anonymous referee for pointing this out to me.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

thoroughgoingly realistic explanation" isn't warranted. Take the concept of 'inertial mass' as it was introduced to natural philosophy by Newton. Why were he and other natural philosophers who came after him convinced that mass is a "natural" property of a physical system? Because they observed that simple hypotheses formulated in terms of mass can fit the extant data nicely. That is, the fact that scientific methodology jibes with the Akaikean recommendation explains the fact that the development of M(S) over time leads to the inclusion of new hypotheses that are KL-closer to truth than their predecessors.

This is not (yet) a full explanation of the fact that some members of M(S) are meaningfully KL-close to truth. We could be getting closer (in adding new members to M(S)) but not be meaningfully close. Thus, the remaining question is: Why is it that we are able to formulate theories that are meaningfully KL-close to truth at all? I don't know the answer, but I can think of only two initially promising ones. (I) Nothing. It is just a brute fact that there are members of M(S) that are KL-close to truth. This might not be as bad an idea as it initially appears. As I argued earlier, being KL-close to truth is not a particularly exacting requirement. Therefore, that our most fundamental epistemic tendencies (e.g., the concepts in terms of which we tend to think) are such that in *some* areas of inquiry, there are members of M(S) that are KL-close to truth might not be a particularly impressive epistemic condition to need an explanation. (II) It is evolutionarily beneficial to think in terms of concepts that are such that simple hypotheses formulated in their terms are KL-close to truth. Notice, however, that although constructing theories that are KL-close to truth (or instrumentally reliable) might have evolutionary benefits, forming ontologically/fundamentally true theories has no *further* evolutionary benefit. (That is, having ontologically/fundamentally true beliefs has no evolutionary benefit conditional on the belief being instrumentally reliable.) Thus, if this explanation is indeed true, it doesn't favor the realist's position.[27]

Is there any realist-friendly explanation for the fact that some of the members of M(S) are KL-close to truth? The only thing that comes to my mind is something like Descartes's appeal to God's benevolence in creating us (which justifies his clear-and-distinct-idea criterion). According to that explanation, we tend to think about concepts that enable us to discover deep truths, which then entail predictive elements that are KL-close to truth. Regardless of what one thinks about this epistemological view, I think even Descartes would have probably agreed that it is particularly unconvincing to argue for it *by way of explaining the Remarkable Success of science*. I cannot think of any plausible realist rejoinders at this point.

Therefore, I conclude that the strongest version of the Realist Thesis that NMA can plausibly establish is RKL and, in some instances, its extensions like RKL^{+}.

Perhaps there is a deeper lesson we can learn from NMA's inability to establish a strong version of realism, independently of the particular nature of RKL and its explanation. In many academic disciplines, it is part of common wisdom that there ain't no such thing as a free lunch. Philosophers are by no means strangers to this idea. It is a notorious fact about deductive reasoning that it doesn't discover anything that's not already contained in the premises. That is why there's been more attention in recent decades to various nondeductive forms of reasoning, both as objects of study and as forms of reasoning within philosophy. Granted that such nondeductive forms of reasoning are ampliative—they add *something* to the premises—we should still be wary of arguments that purport to give us something *big* for nothing. But that is exactly what NMA purports to give us. It begins with the instrumental success of science and purports to establish the (near) truth of ontological or fundamental commitments of scientific theories. Even if we didn't have a well worked out answer to the realist's challenge for explaining Remarkable Success, this big leap in NMA would have been sufficient ground for concern. Indeed, I find it particularly problematic that NMA is

[27]The idea here is quite different from van Fraassen's Darwinist explanation of the *Remarkable Success* of scientific theories. (The argument is presented in van Fraassen [1980, 40]. See Lipton [1991, 192–95] and Psillos [1999, 93–94] for compelling critiques.)

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

insensitive to the particular *content* and the prima facie metaphysical plausibility of the theory's ontology—instrumental success is supposed to be enough.

A useful concept in studying a physical system is that of its 'degrees of freedom.' When the number of constraints on the system are less than its degrees of freedom, its state is indeterminable. Perhaps a similar notion is useful in epistemology. When all you know about a theory is that it is predictively successful, there is no way you can decide the truth of particular parts of it. More information is needed to settle the issue one way or the other. And in all likelihood, the issue is going to be settled differently for different theories, even if they are all predictively successful. This is compatible with the plausible idea that the predictive success of a theory provides *evidence* for its fundamental commitments. But having evidence for something is different from having a (fallible) proof for it.

Acknowledgments. I am grateful to Adam Elga, Snow Xueyin Zhang, Gideon Rosen, Mehdi Hatef, Curtis Haaga, two anonymous referees, and audiences at Bilkent University and Institute for Research in Fundamental Sciences in Tehran for their helpful feedback.

References

Akaike, Hirotugu. 1973. "Information Theory as an Extension of the Maximum Likelihood Principle." In *Second International Symposium on Information Theory*, edited by B. Petrov and F. Csaki, 267–81. Budapest, Hungary: Akademiai Kiado.

Boyd, Richard. 1990. "Realism, Approximate Truth and Philosophical Method." In *Scientific Theories*. Minnesota Studies in the Philosophy of Science, vol. 14, edited by C. Wade Savage, 355–91. Minneapolis: University of Minnesota Press.

Chang, Hasok. 2003. "Preservative Realism and Its Discontents: Revisiting Caloric." *Philosophy of Science* 70 (5): 902–12. https://www.jstor.org/stable/10.1086/377376.

Fatollahi, Alireza. 2023. "Predictivism and Model Selection." *European Journal for Philosophy of Science* 13 (1). https://doi.org/10.1007/s13194-023-00512-1.

Fine, Arthur. 1991. "Piecemeal Realism." *Philosophical Studies* 61 (February): 79–96. https://doi.org/10.1007/BF00385834.

Forster, Malcolm. 2002. "The New Science of Simplicity." In *Simplicity, Inference and Modeling*, edited by Arnold Zellner, Hugo A. Keuzenkamp, and Michael McAleer, 83–119. Cambridge: Cambridge University Press.

Forster, Malcolm, and Elliott Sober. 1994. "How to Tell When Simpler, More United, or Less Ad Hoc Theories Will Provide More Accurate Predictions." *British Journal for the Philosophy of Science* 45 (1): 1–36. https://www.journals.uchicago.edu/doi/10.1093/bjps/45.1.1.

Forster, Malcolm, Garvesh Raskutti, Reuben Stern, and Naftali Weinberger. 2018. "The Frugal Inference of Causal Relations." *The British Journal for the Philosophy of Science* 69 (3): 821–48. https://doi.org/10.1093/bjps/axw033.

Frankel, Eugene. 1976. "Corpuscular Optics and the Wave Theory of Light: The Science and Politics of a Revolution in Physics." *Social Studies of Science* 6 (2): 141–84. https://doi.org/10.1177/030631277600600201.

García-Lapeña, Alfonso. Forthcoming. "Truthlikeness for Quantitative Deterministic Laws." *British Journal for the Philosophy of Science*. https://www.journals.uchicago.edu/doi/10.1086/714984.

Harker, David. (2013). "How to Split a Theory: Defending Selective Realism and Convergence without Proximity." *British Journal for the Philosophy of Science.* 64(1): 79–106. https://www.journals.uchicago.edu/doi/10.1093/bjps/axr059

Kitcher, Philip. 2001. "Real Realism: The Galilean Strategy." *The Philosophical Review* 110 (2): 151–97. https://doi.org/10.2307/2693674.

Kullback, Solomon, and Richard A. Leibler. 1951. "On Information and Sufficiency." *Annals of Mathematical Statistics* 22 (1): 79–86.

Laudan, Larry. 1981. "A Confutation of Convergent Realism." *Philosophy of Science* 48 (1): 19–49. https://doi.org/10.1086/288975.

Lipton, Peter. 1991. *Inference to the Best Explanation*, London: Routledge.

Lyons, Timothy. 2006. "Scientific Realism and the Stratagema de Divide et Impera." *British Journal for the Philosophy of Science* 57 (3): 537–60. https://www.journals.uchicago.edu/doi/10.1093/bjps/axl021.

Maxwell, Grover. 1970. "Theories, Perception and Structural Realism." In *The Nature and Function of Scientific Theories*, edited by in Robert G. Colodny. Pittsburgh, PA: University of Pittsburgh Press.

Niiniluoto, Ilka. 1987. *Truthlikeness*. Dordrecht, Nether.: Reidel.

Niiniluoto, Ilka. 2021. "Approaching Probabilistic Laws." *Synthese* 199 (3): 10499–519. https://doi.org/10.1007/s11229-021-03256-8.

Niiniluoto, Ilka, Gustavo Cevolani, and Theo Kuipers. 2022. "Approaching Probabilistic Truths: Introduction to the Topical Collection." *Synthese* 200 (2): 1–8. https://doi.org/10.1007/s11229-022-03516-1.

Popper, Karl. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge and Kegan Paul.

Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. London: Routledge.

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press

Putnam, Hillary. 1975. *Philosophical Papers: Mathematics*, *Matter and Method*, vol. 1. Cambridge: Cambridge University Press.
Rosenkrantz, R. D. 1980. "Measuring Truthlikeness." *Synthese* 45 (3): 463–88. https://www.jstor.org/stable/20115570.
Sober, Elliott. 1999. "Instrumentalism Revisited." *Critica* 31(April): 3–39. http://www.jstor.org/stable/40104484.
Sober, Elliott. 2002. "Instrumentalism, Parsimony, and the Akaike Framework." *Philosophy of Science* 69 (3): 112–23. https://doi.org/10.1086/341839.
Stanford, Kyle. 2006. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford: Oxford University Press.
van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Clarendon Press.
Vickers, Peter. 2019. "Towards a Realistic Success-to-Truth Inference for Scientific Realism." *Synthese* 196 (2): 571–85. https://www.jstor.org/stable/45096367.
White, Roger. 2003. "The Epistemic Advantage of Prediction over Accommodation." *Mind* 112 (448): 653–83. https://doi.org/10.1093/mind/112.448.653.

Alireza Fatollahi is an assistant professor of philosophy at Bilkent University. He received his PhD from Princeton University in 2020. His interests are in philosophy of science and history of early modern philosophy.

Cite this article: Fatollahi, A. 2023. Akaike and the No Miracle Argument for Scientific Realism. *Canadian Journal of Philosophy* 53: 21–37, doi:10.1017/can.2023.21

https://doi.org/10.1017/can.2023.21 Published online by Cambridge University Press