European Journal for Philosophy of Science (2025) 15:63
https://doi.org/10.1007/s13194-025-00699-5

PAPER IN GENERAL PHILOSOPHY OF SCIENCE

How to estimate the success chance of a scientific theory? On the no miracles argument and the base rate fallacy

Johannes Nyström[1]

Received: 26 February 2025 / Accepted: 13 October 2025

Abstract
Colin Howson (2000) claims that the no miracles argument in favor of a realist interpretation of a scientific theory falls prey to the base rate fallacy and is therefore invalid on logical grounds. In response, Dawid and Hartmann (2018) claim that Howson only reconstructs a limited part of the argument. They argue that a more complete reconstruction of the no miracles argument takes into account the success frequency of a wider spectrum of scientific theory building, and therefore avoids the base rate fallacy. In a critical response to Dawid and Hartmann, Boge (2020) presents two challenges to their approach, both of which are designed to provide reasons for skepticism about treating observed success frequencies in science as connected to the relevant base rates. In this paper, I argue that Boge's challenges are not effective.

Keywords Scientific realism · No miracles argument · Base rate fallacy

1 Introduction

Scientific realism is, roughly, the view that well-confirmed scientific theories are typically approximately true. The view is defended on the basis on one general philosophical argument: the no miracles argument (NMA). The core premise of the NMA asserts that the only plausible explanation of the predictive success demonstrated by science is that predictively successful scientific theories are typically approximately true (Putnam, 1975; Boyd, 1984). In other words, that success appears miraculous, or a 'cosmic coincidence' (Smart, 1985), in absence of a realist explanation. Ever since the initial formulation of the NMA, the core premise of the argument has been criticized, and a rich debate on its plausibility has since ensued. This debate underwrites

✉ Johannes Nyström
johannes.nystrom@philosophy.su.se

[1] Department of Philosophy, Stockholm University, Stockholm, Sweden

Published online: 03 November 2025

one of the most long-standing general philosophical disagreements in the philosophy of science, and no general consensus on this issue has been established. Both realism and anti-realism remain common views among philosophers of science to this day.

However, one perspective on the NMA threatens to disqualify the argument before debate on its core premise can even get started. Colin Howson (2000) has put forward the strong claim that, *irrespective* of whether or not the core premise of the NMA is true, the argument cannot justify the realist position, because it falls prey to a logical fallacy. On the basis of a Bayesian reconstruction of the NMA, Howson argues that the argument ignores the impact of base rate information about approximate truth. Such information, however, is crucial in order to understand whether or not the realist implication of predictive success, even if highly compelling in isolation, is ultimately a sufficient justification of scientific realism. For this reason, Howson concludes that the NMA commits the so-called *base rate fallacy* and is therefore invalid on logical grounds. This conclusion is also endorsed by Magnus and Callender (2004).

In response, some authors (Worall, 2007; Psillos, 2009) have expressed concerns about the adequacy of a Bayesian perspective on the NMA and the realist position, and suggested that the argument cannot be properly reconstructed in this framework. Against this criticism, Howson (2013) himself points out that rejecting a probabilistic framework in this context leaves the proponent of scientific realism without crucial tools to express the epistemic merits of the realist position. Buying into Howson's reply, two more recent attempts to counter his original argument grant that there are no inherent issues with a Bayesian reconstruction of the NMA, but instead express a line of criticism which targets Howson's understanding of the NMA. Henderson (2017) and Dawid and Hartmann (2018) [Henceforth, 'DH'] each suggest that Howson only reconstructs a part of the full NMA structure. They then each claim that a more complete reconstruction of the argument shows that the argument in fact takes into account base rate information about approximate truth and conclude that the NMA therefore avoids the base rate fallacy.

Henderson's argument is based on the suggestion that the overall high success rate of science is evidence for the hypothesis that theory selection in science is biased in favor of approximately true theories, and that supplementing the NMA with this evidence leads to a valid argument. Against this claim, both Boge (2020) and Dyck (2023) submit that, in the absence of an established base rate, concluding that a bias of this kind exists in no way constrains the relevant prior probability of approximate truth for a given scientific theory. But Howson's core claim is exactly that the NMA is not valid because of the absence of a constraint of that kind. Hence, Henderson's argument fails to improve the realist's position with respect to this issue.

DH's line of reasoning, however, leads to a constraint on the prior probability of approximate truth and therefore avoids this problem, as both Boge (2020: 4344) and Dyck (2023: 766) also note. On the basis of a formal argument, DH demonstrate that this prior is constrained by the core premise of the NMA, given that the theory targeted in the NMA is assigned a sufficiently large probability of predictive success, prior to testing. Next, they adopt a statistical approach, and suggest that the frequency of predictive success in the relevant scientific research field can provide a well-founded estimation of this probability. In other words, if the research field of choice demonstrates a suitably high frequency of predictively successful theories, a NMA in favor of realism about a successful theory in that field is not a fallacious argument. DH call this the *frequency-based* NMA, and suggest that Howson's reconstruction is not sensitive to the frequency-based aspect of the NMA.

In a critical response to DH, Boge (2020) presents two challenges to the frequency-based NMA, designed to provide reasons to doubt the viability of DH's frequency-based approach to estimating the prior probability of predictive success of a scientific theory. Boge's two challenges are interesting and may be understood to represent intuitive concerns about the frequency-based NMA. Hence, they put into question whether the frequency-based NMA is an adequate response to the charge that the NMA falls prey to the base rate fallacy. In this paper, however, I argue that the challenges are not effective.

The structure of the paper is as follows. Section 2 sets up the relevant background; In section 2.1, I reconstruct Howson's criticism of the NMA, and in section 2.2, I present DH's response to Howson, and their frequency-based version of the NMA. In sections 3.1 and 3.2, I turn to Boge's two challenges. With the help of a Bayesian reconstruction of DH's frequency-based approach to estimating the prior probability of predictive success of a scientific theory, I explain why the challenges do not provide grounds for doubting the logical validity of the frequency-based NMA.

2 The no miracles argument and the base rate fallacy

2.1 Howson's reconstruction of the no miracles argument

Consider a diagnostic test for a rare disease that occurs only in one in a thousand people. The test is extremely sensitive (it always correctly diagnoses a positive case of disease) and very specific (the probability of a false positive is only 0.05). Next, consider a test that gave a positive result, conducted on a randomly picked test subject. What is the probability that the test subject has the disease? Given the high sensitivity and specificity of the test, one may be led to conclude that the patient most likely has the disease. But this would be neglecting the small base rate of the disease in the population from which the test subject was picked. Given the low base rate of disease in the population, only one of every thousand tests conducted in the specified way will correctly diagnose a positive case of the disease. And given that one in every twenty tests will be a false positive, the rate of true positive to false positive tests is just 1/51. In other words, the probability that any particular test is a true positive is approximately 0.02. Hence, the probability that the test subject actually has the disease, conditional on a positive test, is still very small.[1]

The *base rate fallacy* is typically understood as a reasoning error which disregards the significance of base rate information of this kind (e.g. Howson & Urbach, 2006: 24). Formally reconstructed in Bayesian epistemology, the base rate of disease is represented as the prior probability that any randomly picked test subject has the disease. If this prior is very small, i.e., if one's initial belief that the test subject does *not*

[1] Both Howson (2000) and Henderson (2017) use this analogy in their discussion of the NMA.

have the disease is very strong, even a very sensitive and specific positive test may be considered fairly insignificant with respect to establishing infection.

Howson's criticism against the NMA turns on the claim that this argument falls prey to a similar reasoning error. Let H be a scientific theory, and let *T* be a binary propositional variable with values T: H is approximately true, and its negation ¬ T. Let *S* be a binary propositional variable with values S: H is predictively successful, and its negation ¬ S. The core premise of the NMA can now be understood as the following two propositions, roughly analogous to the sensitivity and specificity of the medical test:

$\mathbf{A}_1$: *P*(S T) is large.

$\mathbf{A}_2$: *P*(S ¬ T) $< k << 1$.

Where A_1 formalizes the assumption that approximately true theories should be expected to be predictively successful, and A_2 formalizes the assumption that theories that are not at least approximately true should not be expected to be predictively successful. In other words, *S* can be treated as a fairly sensitive and specific test of *T*. From A_1 and A_2, the scientific realist infers that *P*(T S) is large. But, as Howson points out, this conclusion does not follow from the premises, because the realist's inference does not take into account the possibility of a very small prior probability of T. *P*(T S) can be calculated with Bayes' Theorem:

$$P(T \mid S) = \frac{P(S \mid T)}{P(S \mid T)P(T) + P(S \mid \neg T)P(\neg T)} P(T) \tag{1}$$

The product at the right-hand side of this equality implies that no (non-dogmatic) assignments of values to A_1 and A_2 guarantees the realist's conclusion, given an arbitrarily small prior probability *P*(T). Hence, concluding that *P*(T S) is large, perhaps larger than 0.5, the realist neglects the possibility of a very small base rate of approximate truth, and therefore falls prey to the base rate fallacy.

Formally, the argument may of course be validated by adding a premise which suitably constrains *P*(T). But the problem runs deeper. *P*(T) is difficult to constrain, given that the probability that the theory under scrutiny is approximately true is the core epistemic object of disagreement between scientific realists and anti-realists. Asserting *a priori* constraints on that probability hence begs the question. For this reason, Howson concludes that the NMA therefore is bound to fail already at the level of establishing its logical validity.

2.2 The frequency-based no miracles argument

DH (2018) offer a response to Howson that is based on claiming that his reconstruction of the NMA is not fully comprehensive. DH individuates two different versions of the NMA: *individual theory-based* NMA, and *frequency-based* NMA. The former starts out by selecting a scientific theory which is known to be successful, and then deploys the realist conjecture to explain that theory's success. The latter, on the other hand, deploys the realist conjecture to explain the observation that theories which are developed and tested in science *tend to be* predictively successful.[2] DH then claim that Howson only reconstructs the individual theory-based NMA, and that his conclusion that this argument falls prey to the base rate fallacy is correct: Individual theory-based NMA does not specify any constraints on P(T) and therefore does not convince anyone who is sufficiently initially skeptical about the theory's approximate truth, even if that skeptic accepts $\mathbf{A}_1$ and $\mathbf{A}_2$ for the sake of the argument. However, they show that, once a frequency-based NMA is adopted, the base rate fallacy can be avoided, because the frequency-based NMA *does* specify certain constraints on P(T).

In order to demonstrate that the frequency-based NMA avoids the base rate fallacy, DH construct a formal Bayesian model of the argument. They start by identifying the historical frequency of predictively successful theories n_S to the total number of tested theories n_E in some scientific research field D: $n_S/n_E := R$. Next, they assume that each new theory H in D can be treated as a random pick from D with respect to predictive success, and its probability of predictive success P(S) can therefore be estimated according to R.[3] Finally, DH prove that, according to the law of total probability, P(T) is bounded from below by the difference P(S) – k. If $k < P$(S), that bound is therefore positive. Since k is small by assumption $\mathbf{A}_2$, satisfying this inequality requires only a modestly large value of R, in which case the core premise of the NMA leaves no room for an arbitrarily small P(T). In other words, if the frequency of predictive success in a research field is sufficiently high, assuming $\mathbf{A}_1$ and $\mathbf{A}_2$ means assuming a prior probability of approximate truth for theories in that field that is positively bounded from below. On the observation that H was predictively successful, the argument can then proceed according to Howson's reconstruction of the individual theory-based NMA.

The frequency-based NMA consists of two parts. The first part is the conclusion that P(T) is bounded from below by the difference P(S) – k, and is obtained mathematically on the basis of the law of total probability. The second part is the claim that P(S) can be estimated according to a statistical inference based on the frequency of predictive success in a scientific research field. This part of the frequency-based NMA is less rigid, because an inference of this kind relies on making several nontrivial distinctions and observations. What counts as a predictive success? What counts as a theory? What are the boundaries on a scientific research field? Indeed, the difficulty in satisfactorily answering these questions have led some authors to reject the tractability of this inference. Considering a kind of prototype version of the frequency-based NMA, Magnus and Callender (2004: 325) hold that "it is impossible to count up or even fairly sample all the theories that were considered for our mature sciences, and so it is impossible to evaluate whether [R>>0] obtains". More recently, Dyck (2023: 775) points out that it may be especially difficult to produce a candidate set of theories that is satisfactory to both camps of the realism debate.

[2] This distinction is analogous to Henderson's (2017) distinction between *local* and *global* NMA, and to Dawid's (2008) distinction between *analytic* and *epistemic* NMA.

[3] DH adopt the frequentist estimate $P(\mathrm{S}) = R$ for conceptual simplicity, but note that a Bayesian estimate is more involved, since it takes into account prior beliefs about S. However, on DH's understanding of the argument, those beliefs are treated as subjective prior probabilities which are washed out in the large n_E limit, and the frequentist and Bayesian estimate therefore converge in that limit.

Nevertheless, Dyck also suggests that proceeding with this analysis may be "the best option available" for advancing the scientific realism debate. That something is difficult does not mean that it is impossible. Furthermore, what is at stake here are not exact probabilities, but the plausibility of the realist position. Justifying that position with respect to H only requires showing that $P(\mathrm{T}\quad\mathrm{S})$ is larger than some relevant threshold.[4] Concluding simply that R falls into a certain range or interval may well be enough to establish inequalities of this kind (given, of course, A_1 and A_2).[5]

However, even if one holds against Magnus and Callender that R can be satisfactorily established, there is another potential problem with DH's suggested statistical inference. In many contexts, the available data may not be substantial enough to take R as a reliable estimate of $P(\mathrm{S})$, because the sample size may just be too small. In this situation, the frequency-based NMA is exposed to a similar problem that befalls the individual theory-based NMA: the assessment of the significance of the relevant observations with respect to the realist conjecture will be strongly subjective, even if the core realist assumptions A_1 and A_2 are assumed for the sake of the argument.

While a situation of this kind does present a real problem for the scientific realist, it also reveals why DH claim that, unlike the individual theory-based NMA, the frequency-based NMA does not fall prey to the base rate fallacy. The individual theory-based NMA does not offer a framework within which the subjective element can be discharged and this element therefore remains structurally necessary in the justification of scientific realism. The frequency-based NMA, on the other hand, is based on an open series of experimental tests of new scientific theories in D. This series ensures that the influence of subjective priors can eventually be washed out in the large n_E limit. Hence, an anti-realist who accepts A_1 and A_2 for the sake of the argument but disagrees with the realist about the significance of the available data R with respect to the crucial probability $P(\mathrm{S})$ can just wait for new data to come in to settle the disagreement. In DH's words, "the deep reason why frequency-based NMA avoids the base rate fallacy ... lies in the fact that it provides a framework in which the convergence behaviour of posteriors under repeated updating can be exploited" (Dawid & Hartmann, 2018: 4071-4072).

3 How to estimate the success chance of a scientific theory?

The two concerns about DH's statistical approach to estimating $P(\mathrm{S})$ described above are based on the claim that the relevant statistical inference may in many research fields suffer from inconclusive or ambiguous data. These are clearly serious concerns for the scientific realist who relies on the NMA to justify a high credence in

[4] For example, DH show that, for $P(\mathrm{S} \mid \mathrm{T}) = 1$, $P(\mathrm{S})$ needs to be larger than $2 \times P(\mathrm{S} \mid \neg \mathrm{T})$ in order to obtain $P(\mathrm{T} \mid \mathrm{S}) > 0.5$ (Dawid & Hartmann, 2018: 4073).

[5] DH's own perspective aligns with this understanding. They claim that "[the realist] normally is not in the situation to provid[e] a complete count of the successful and failed theories in the field based on precise criteria for what counts as a theory. The presented formalisation [of the frequency-based NMA] does not suggest that scientific realists must provide an actual count of theories any more than Bayesian confirmation theory suggests that scientists must carry out Bayesian updating from explicitly specified priors of their theories" (Dawid & Hartmann, 2018: 4069).

the approximate truth of her scientific theory of choice. However, they are practical problems that can in principle be addressed on the basis of a careful assessment of the relevant data sets. The current most critical response to DH's frequency-based NMA, however, is based on a more fundamental issue. On the basis of two independent arguments, Boge (2020) presents two challenges for DH's suggested statistical approach to estimating *P*(S). In the next sections, I evaluate these challenges.

3.1 Why transfer expectations of success to a new scientific theory?

Boge's first challenge concerns the inductive justifications for DH's core assumption that *P*(S) can be estimated according to *R*. Boge begins with delimiting his challenge. To simplify their formal model, DH adopt the strict frequentist equality $P(S) = R$. Boge notes that a Bayesian will reject this strict equality, since she would take into account her prior beliefs about S when updating *P*(S) on *R*. However, he accepts DH's practical choice to stick with the frequentist estimate. Since DH treat these beliefs as subjective priors which are eventually washed out by testing ever more theories in *D*, the Bayesian and frequentist estimates of *P*(S) are taken to converge in the large n_E limit. Hence, according to Boge, "the 'roughness' of the estimate n_S/n_E is really not what's problematic here" (Boge, 2020: 4345). What he is rather concerned with is "why the success of the *other* theories in *D* should confer *any* probability to the success of an entirely new theory, H, just because H also falls into *D*, i.e., concerns the same subject matter" (Boge, 2020: 4345).

Boge illustrates his concern with an analogy, which captures its intuitive pull. Consider a hitherto untested weather forecaster, picked from a group of somewhat successful weather forecasters "in ancient Greece with their predictions informed by past experience and their a priori ideas on how to conceptualize the weather" (Boge, 2020: 4345). Now Boge asks;

> would we ... assume that the forecasts of a *new* forecaster, so far unknown to us, would also be somewhat successful, *just* because she tells us something about the weather, i.e., in the absence of any further information, as required by DH? Clearly one's willingness to transfer the expectation of success to the new member may depend on all sorts of *other* factors: personal connections between the new forecaster and the known ones, similarities and differences it their basic beliefs, similarities in method, and so forth. Hence in the absence of further knowledge, one should be at liberty to remain far more skeptical about the new forecaster's success (Boge, 2020: 4345).

Boge supplements his analogy with a concrete argument. He first notes that predictive success in scientific realism is typically understood as successful prediction of novel phenomena, rather than accommodation of known phenomena. Then, he asserts that it is unclear why one would assume that the success rate of earlier theories in *D* would be relevant when estimating the new theory's chance of predictive success in the domains where the theory's prediction diverges from those of earlier theories, which is required for novel predictive success. To illustrate this concern, he takes an example from theoretical physics: "why should the *mere* success of Newtonian mechanics (NM) ... and that of special relativity (SR) ... have committed scientists to the belief in general relativity (GR) ... ? Could GR not have easily made *false* predictions where it diverges from Newton's theory?" (Boge, 2020: 4346). Indeed, Boge points at the numerous potential alternatives to GR which were considered at the time before GR was empirically confirmed, and states:

> If it were correct that "in the absence of further knowledge, the success chances of a new theory should be estimated according to [the] rate of predictively successful theories in *D*" (DH: 5), should the success of NM and SR not have led us to the false expectation of these alternative theories' successes, before they were studied in detail or tested empirically? (Boge, 2020: 4346).

Boge's challenge identifies two apparent issues for DH's statistical approach. The first issue conceptual: in the absence of specific knowledge about the connections between a new theory H in a research field and earlier theories in that field, why should the probability of H's novel predictive success be estimated according to the success rate of those earlier theories? The second issue is set up like a reductio. A consequence of assuming $P(\mathrm{S}) = R$ in the context of physics is that one's expectation about the success of the alternatives to GR would have been much higher than what was ultimately proven to be the case, which looks like a problem for that assumption. A successful response to Boge must address both issues associated with his challenge. Let me begin by addressing the first issue on the basis of a more elaborate Bayesian analysis of DH's suggested estimation, before then going on to address the second issue.

In order to obtain a more detailed Bayesian perspective on DH's statistical approach to estimating $P(\mathrm{S})$, we first need to identify a (statistical) population of theories in some research field D. This population consists of the set H^D of past, present and future theories in D. Next, we are interested in the *relative frequency* of predictively successful theories in H^D. Since we are operating in a Bayesian framework, we begin by formalising our prior beliefs about that frequency. We assign a prior probability distribution $\mathcal{P}^{\mathcal{F}}$ over a propositional variable $\mathcal{F}$, such that $\{\mathcal{F} = j\} := \mathcal{F}_{\mathrm{j}}$, and j is a real number in the [0, 1] interval:[6]

$\mathcal{F}_{\mathrm{j}}$:= the relative frequency of predictive success in H^D is j.

Now consider a theory H from H^D. By the law of total probability, H's probability of success $P(\mathrm{S})$ is expressed as:

$$\mathrm{P}(S) = \sum_{j=0}^{1} \mathrm{P}(S \mid \mathcal{F}_j)\mathrm{P}(\mathcal{F}_{\mathrm{j}}) \tag{2}$$

[6] We must assume that $\mathcal{P}^{\mathcal{F}}$ is non-extreme, i.e., that we are not already in advance fully certain about the relative frequency of success in H^D. If we would be, it would be trivially true that we could not update $\mathcal{P}^{\mathcal{F}}$. Moreover, if we already know the relative frequency of success in H^D, we get a base rate of approximate truth from $\mathbf{A}_1$ and $\mathbf{A}_2$. Hence, the validity of the NMA could be guaranteed by a sufficiently small choice of $\mathbf{A}_2$ (of course, if exactly *no* theory in H^D is successful, no particular choice of $\mathbf{A}_2$ generates a valid NMA).

Now, in the absence of further knowledge about H other than that it is from D, it will be treated as a random pick with respect to predictive success from D. Hence, for all j, we get $P(\mathrm{S} \mid \mathcal{F}_j) = j$, because if the relative frequency of predictive success in H^D is j, the probability that a random pick from H^D is predictively successful is by definition exactly j. Hence, we obtain:

$$\mathrm{P}(S) = \sum_{j=0}^{1} j\mathrm{P}(\mathcal{F}_{\mathrm{j}}) = \bar{\mathcal{F}} \tag{3}$$

where $\bar{\mathcal{F}}$ is the $\mathcal{P}^{\mathcal{F}}$– weighted mean of $\mathcal{F}$.

The equality expressed in (3) is a formalisation of the fact that if one has no knowledge about H except $\mathrm{H} \in \mathrm{H}^D$, one's opinion on the relative frequency of success in H^D directly determines P(S), on pain of logical inconsistency. In other words, if H is treated as a random pick with respect to predictive success from H^D, its probability of predictive success is equal to the relative frequency of predictive success in H^D.

In order to understand how observing earlier instances of predictive success in D impacts H's probability of success, we assume that some earlier theory H was predictively successful when exposed to empirical testing, and assess any (dis)confirmatory effect of this observation on each $\mathcal{F}_{\mathrm{j}}$. Let S stand for the observation that H was successful. We can calculate the posteriors $P(\mathcal{F}_j \mid S')$ with the help of Bayes' theorem:

$$P(\mathcal{F}_j \mid S') = \frac{P(S' \mid \mathcal{F}_j)}{P(S')} P(\mathcal{F}_j) \tag{4}$$

In absence of further information about H , it is treated as a random pick with respect to predictive success, and we obtain (S $\mid \mathcal{F}_j) = j$ and P(S) $= \bar{\mathcal{F}}$. Hence, we can rewrite Eq. 4 as:

$$P(\mathcal{F}_j \mid S') = \frac{j}{\bar{\mathcal{F}}} P(\mathcal{F}_j) \tag{5}$$

Hence, we see that, given $0 < P(\mathcal{F}_j) < 1$, if $j > \bar{\mathcal{F}}$, then we obtain $P(\mathcal{F}_j \mid S') > P(\mathcal{F}_j)$, that is, $\mathcal{F}_j$ is confirmed by S . If $j < \bar{\mathcal{F}}$, then we obtain $P(\mathcal{F}_j \mid S') < \mathrm{P}(\mathcal{F}_j)$, that is, $\mathcal{F}_j$ is disconfirmed by S . The (dis)confirmation value depends on the difference of j and $\bar{\mathcal{F}}$: The larger (smaller) j is compared to $\bar{\mathcal{F}}$, the more it is (dis)confirmed, and the probability distribution across $\mathcal{F}$ will thereby be updated to favour higher (lower) $P(\mathcal{F}_{\mathrm{j}})$. Given (3), we see that H's probability of success thereby increases or decreases by updating on the success or failure of earlier theories in D. That is, observing the predictive success of the earlier theory H updates one's opinion on the relative frequency of success in H^D, and, therefore, one's credence in the success of an entirely new theory H.[7]

[7] DH's frequentist equality $P(\mathrm{S}) = R$ avoids the arduous task of updating the probability distribution across $\mathcal{F}$ for each instance of predictive success and failure in D. As long as $\mathcal{P}^{\mathcal{F}}$ is treated as a strict prior distribution (i.e., as a fixed initial condition of the probabilistic model), the two frameworks converge in the large n_E limit, and the frequentist estimate therefore involves only quantitative distortions at small n_E.

The above analysis assumes, with DH, that "each new theory that comes up in *D* can be treated as a random pick with respect to predictive success" (Dawid & Hartmann, 2018: 4067). However, there may be situations where this assumption should be rejected. For example, one may consider a situation where information is available which leads to a substantially higher credence in the success of earlier theories in the research field *D* than in the success of the new theory. Boge's weather forecasting analogy could be interpreted as spelling out such reasons. If one adds to his analogy the assumptions that the methodology and conceptualizations employed by the ancient forecasters can be associated a priori with a higher probability of forecasting success (and further that those tools are not known to be common to all forecasters) while the new forecaster is simply treated as a random pick from what is considered an overall fairly unsuccessful bunch of forecasters, then one encounters a situation where the power of the frequency-based estimation of the new forecasters success may be significantly reduced. In the case of the NMA, this information controls the P(S) which, as demonstrated by equation (4), is crucial to understand the significance of instances of predictive success in the field of interest with respect to beliefs in the relative frequency of success in that field. If this probability is large, the updating in credence in the relative frequency of success by observing S will not be very significant, and one will be left with a fairly unchanged opinion on that relative frequency. Alternatively, or in addition, one may have information about specifically H which disqualifies treating H as a random pick, and which thereby impacts one's credence in H's success.

These situations are by no means ruled out by DH's suggested framework in isolation, as they operate under the assumption that theories in *D* are treated as random picks with respect to predictive success. Hence, they underwrite a line of criticism that is in principle available to the anti-realist. However, insofar as a situation of this kind actually obtains, the relevant information must be identified for each case and each theory. In order to put forward this kind of argument against the frequency-based NMA, one would have to identify reasons to believe that the earlier theories in *D* differ in some specified way from the new theory that is related to their respective probabilities of predictive success. Simply postulating that there *can be* a difference of this kind does not undermine the validity of DH's suggested statistical analysis. In the absence of information of this kind, theories will be treated as random picks with respect to predictive success. Moreover, in order to reject the validity of a frequency-based NMA, the anti-realist would further have to demonstrate that this information speaks against, rather than for, the theory H that is the target of that argument. Assuming that one has special information about the theories in *D* which disqualifies treating them as random picks with respect to predictive success may just as well be taken to imply a improvement of H's success chance. For example, as a reviewer for this journal suggests, one may be tempted to assume that, *ceteris paribus*, a new theory may be considered more likely to succeed than theories that preceded it, simply due to the natural progression of science. In any case, anti-realists who seeks to ascribe to H a small success chance due to the availability of information of this kind carries the evidential burden of presenting that information.

Before moving to the second issue associated with Boge's challenge, let me briefly address a final point made by Boge that is connected to his skeptical line of reasoning about the justification for assuming $P(\mathrm{S}) = R$ described above, and which is related to the way the Bayesian updating described above proceeds:

> ... given the above arguments [described above], we must ask whether the fact that the new theory *is from D* suffices to sanction the induction in question. A detailed Bayesian analysis would allow a more precise assessment of this inductive justification. But it would presumably appeal to the theorem "of fundamental importance to the interpretation of the probability calculus as a logic of inductive inference" that "[i]f h entails e and p(h) >0 and p(e) < 1, then p(h e) >p(h)." (Howson & Urbach, 2006, p. 20; notation adapted) In the present case the hypothesis, h, at any stage of the updating procedure must be that the next theory in D will succeed and the evidence, e, that a number of theories in D have succeeded so far. ... A core problem with the justification for using R or a similar value that comes from conditioning on the empirical success of other theories is hence that they may be largely unrelated, or not related in the appropriate ways, whence the sort of entailment relation required for the updating may not even hold: The success of the older theories will not be entailed by any *genuine* success of a new one, in the sense specified above, and one could hence reasonably maintain that older successes have no say in the estimation of the new theory's success (Boge, 2020: 4347 [This quote is based on the corrected version (2021) of Boge's text].

While this is an interesting suggestion on how to reconstruct the relevant updating procedure in a Bayesian framework, this point is based on a misunderstanding of how the relevant updating plays out that overly constrains the confirmation relation between R and S. In particular, Boge's claims that the detailed Bayesian analysis he has in mind would appeal to the described theorem, and that an entailment relation of the kind Boge describes is *required* for updating h on e, are both false. The theorem Boge references describes a paradigmatic case of confirmation in Bayesian epistemology, but does not spell out necessary conditions for confirmation.[8] The formal analysis of updating $P(\mathrm{S})$ on R carried out above is a good illustration of this fact, and also serves to illustrate how the relevant updating procedure plays out in absence of an entailment relation between hypothesis and evidence.

The analysis carried out hitherto has addressed the first issue associated with Boge's first challenge. Let me now move on to the second issue. In the statistical framework described above, there is nothing inherently problematic about the fact that one may ascribe a high probability of success to a scientific theory that ultimately fails. If the success rate of physical theory before the advent of GR was considered high, that framework indeed suggests that one might have a fairly high credence in the success of the theory which was up for testing next, since $\bar{\bar{\mathcal{F}}}$ might by then be fairly large. To be clear, it does not at all imply that GR, or its alternatives, *could*

[8] In the paragraph just after the quote from Howson and Urbach (2000: 20) that Boge references in support of his claim, the former write 'but [this theorem] is just one of the results that exhibit the truly inductive nature of probabilistic reasoning. It is not the only one, and more celebrated are those that go under the name of *Bayes's Theorems*".

not have failed. The observation that theories then started to fail in rapid succession would update that credence, and lower one's confidence in the success of future theories. The conclusion would then be that the nascent success of the research field was a statistical fluke, and that the relative success frequency of the field in reality is much lower than what was expected on the basis of that initial success. One would not, however, take the failure of that expectation to imply that it was not well-founded to begin with.

3.2 The argument from incompatibility

In the next section of his paper, Boge puts forward a second challenge against DH's core assumption that $P(\mathrm{S}) = R$, which is set up as another reductio, and which cannot be directly resolved by the formal analysis carried out above. The challenge is based on an argument that can be reconstructed as a three-step line of reasoning. First, Boge suggests that theories which have been developed in a research field can be partitioned into classes, such that theories within a class are compatible (i.e., can be approximately true simultaneously) but theories which belong to different classes are not compatible. Second, he asserts that $P(\mathrm{T})$ must be constrained by the number of classes c. Since only one of the classes can contain approximately true theories, $P(\mathrm{T})$ can *at most* be $1/c$. Finally, given a fairly large c, plausible values of $\mathbf{A}_1$ and $\mathbf{A}_2$ in conjunction with a high frequency of success in the relevant research field lead to a lower bound on $P(\mathrm{T})$ that exceeds $1/c$. Put simply, one is not rationally allowed to expect H to have an R>>0 probability of being successful, given that one accepts $\mathbf{A}_1$, $\mathbf{A}_2$ and $1/c$ as an upper bound on $P(\mathrm{T})$. The formal reason is that, given $\mathbf{A}_1$ and $\mathbf{A}_2$, the law of total probability strongly constrains $P(\mathrm{S})$ given a small upper bound on $P(\mathrm{T})$. According to Boge, this conclusion generates "a conflict with DH's reasoning", because "is easily conceivable that ... values for $\mathbf{A}_1$, $\mathbf{A}_2$ and c could apply but that more than $1/c$ of the theories in the given [research field] can claim success" (Boge, 2020: 4348).

Although Boge's claim that it is 'easily conceiveable' that large values for c could apply in some given scientific research field is questionable, his formal analysis is entirely correct. The lower bound on $P(\mathrm{T})$, established by $\mathbf{A}_1$, $\mathbf{A}_2$ and assuming $P(\mathrm{S}) = R$, can indeed become incompatible with $P(\mathrm{T}) \leq 1/c$. However, this incompatibility it is not *in conflict* with DH's statistical approach to estimating $P(\mathrm{S})$. DH's core claim is that given $\mathbf{A}_1$, $\mathbf{A}_2$ and a sufficiently high frequency of predictive success, $P(\mathrm{T})$ is positively bounded from below. DH do not claim that $\mathbf{A}_1$ and $\mathbf{A}_2$ are in fact justified.[9] The frequency-based 'part' of the NMA, advocated by DH, is required for establishing that the argument is valid, i.e., that the conclusion follows from the premises. It does not in itself make $\mathbf{A}_1$ and $\mathbf{A}_2$ any more plausible than they were to begin with. And, as Boge himself notes (2020: 4349), the observation that predictive success is demonstrated by a large number of incompatible theories can be taken to suggest that H's predictive success is quite probable, even if H is not approximately true (i.e., to a rejection of A_2), rather than that the demonstrated pattern of success will probably

[9] " ... a supporter of the frequency-based NMA must justify assumptions $\mathbf{A}_1$ and $\mathbf{A}_2$ ... whether or not that can be achieved lies beyond the scope of this paper" (Dawid & Hartmann, 2018: 4077).

not continue in the future (i.e., to a rejection of the assumption that $P(\mathrm{S}) = R$). This conclusion is a fully consistent response to the described observation. Hence, there is no general problem with taking the historical success frequency to deliver a fairly reliable prediction about the future success frequency, even in research fields where one finds a large c.

Indeed, this understanding would be consistent with a classic anti-realist argument, the pessimistic meta-induction (Laudan, 1981), against $\mathbf{A}_2$. Proponents of this argument claim that the predictive success observed in science should be expected to continue in the future, even if realism is false. According to them, the historical record of science shows that there is a fairly substantial probability that scientific theories which are absolutely false nevertheless turn out to be quite successful. The success rate of science may therefore be expected to continue in the future, even if one assumes a low rate of approximately true theories.

I do not take a stand on whether or not it is plausible to assume that c is large for a significantly large set of scientific fields. This is a difficult question that is related to the large and complicated debate on the pessimistic meta-induction and selective versions of scientific realism, and to the realist's interpretation of the concept of approximate truth. Whether or not Boge's argument for a large c is ultimately convincing is a question that lies beyond the scope of this paper, which is concerned with the logical validity of the NMA. The point made here is that Boge's argument, even if one grants for the sake of the argument that many scientific fields instantiate a large number c of incompatibility classes, does not identify any special issue for DH's statistical approach to estimating $P(\mathrm{S})$. Contrary to what he claims, accepting $1/c$ as an upper bound on $P(\mathrm{T})$ would not be in conflict with that estimation.

4 Conclusions

The following picture of the logical validity of the frequency-based NMA has emerged in this paper. Estimating the probability that a scientific theory will be predictively successful on the basis of the historical success frequency in the relevant discipline, as suggested by DH, is a fairly standard case of statistical inference. As such, it has all the benefits and drawbacks of such inferences in general. In some contexts, it may be difficult to establish that these inferences are reliable, while in others, they may appear fairly straightforward. In the end, whether or not those inferences turn out to support realism or anti-realism about some given theory or research field is an empirical question that must be addressed by the stakeholders of the scientific realism debate.

Questions about the soundness of the frequency-based NMA lie beyond the scope of the analysis provided in both this paper and by DH. Adopting a frequency-based perspective on the NMA does not amount to the claim that high success frequencies in science imply scientific realism on their own. The core assumptions $\mathbf{A}_1$ and $\mathbf{A}_2$ must still be justified in order to reach that conclusion. In situations where a high frequency of success is coupled with an understanding that an abundance of absolutely false theories have contributed to that success, either by way of a pessimistic meta-induction (Laudan, 1981) or, as Boge discusses, by way of considerations about incompatibility, those assumptions may be difficult to defend. However, the point of understanding the NMA as a frequency-based argument is not to improve the plausibility of the NMA's core premise, but rather to analyse the strongest (logically valid) version of the argument. Hence, whether or not the realist position is justified lies beyond what a frequency-based perspective on the NMA can offer in isolation.

While this picture have expanded and offered clarifications of DH's own presentation of the frequency-based NMA, it remains fully consistent with their core message that the frequency-based NMA does not fall prey to the base rate fallacy. The results of the analysis carried out in this paper therefore strengthen the scientific realist's position in light of concerns about the logical validity of their main general argument.

Author Contributions N/A

Funding Open access funding provided by Stockholm University. No special funding

Data Availability N/A

Declarations

Ethical Approval N/A

Informed consent N/A

Conflicts of Interest No conflicts of interest

References

Boge, F. J. (2020). An argument against global no miracles arguments. *Synthese, 197*(10), 4341–4363. https://doi.org/10.1007/s11229-018-01925-9

Boyd, R. (1984). The current status of scientific realism. In J. Leplin (ed.), *Scientific realism* (pp. 41–82). Berkeley: University of California Press. https://doi.org/10.1525/9780520337442-004

Dawid, R. (2008). Scientific prediction and the underdetermination of theory building. *Phil Sci Archive,* 4008. https://philsci-archive.pitt.edu/id/eprint/4008

Dawid, R., & Hartmann, S. (2018). The No Miracles Argument without the Base Rate Fallacy. *Synthese, 195*(9), 4063–4079. https://doi.org/10.1007/s11229-017-1408-x

Dyck, K. (2023). On No-Miracles and the Base-Rate Fallacy. *Philosophy of Science, 90*(4), 761–776. https://doi.org/10.1017/psa.2023.80

Magnus, P. D., & Callender, C. (2004). Realist ennui and the base rate fallacy. *Philosophy of Science, 71*, 320–338.

Henderson, L. (2017). The no miracles argument and the base rate fallacy. *Synthese, 194*, 1295–1302. https://doi.org/10.1007/s11229-015-0995-7

Howson, C. (2000). *Hume's problem: induction and the justification of belief*. New York: Oxford University Press. https://doi.org/10.1093/0198250371.001.0001

Howson, C. (2013). Exhuming the No-Miracles Argument. *Analysis, 73*(2), 205–211.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach* (3rd ed.). Chicago, IL: Open Court.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science, 48*, 19–49. https://doi.org/10.1525/9780520337442-012

Putnam, H. (1975). What is mathematical truth? In: Putnam, H.: *Mathematics, matter and method*, Collected Papers Vol. 2. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511625268.006

Psillos, S. (2009). *Knowing the structure of nature*. New York: Palgrave McMillan. https://doi.org/10.1057/9780230234666

Smart, J. J. C. (1985). Laws of Nature and Cosmic Coincidences. *The Philosophical Quarterly, 35*(140), 272–280. https://doi.org/10.2307/2218906

Worall, J. (2007). Miracles and Models: Why reports of the death of structural realism may be exaggerated. *Royal Institute of Philosophy Supplement., 61*, 125–154. https://doi.org/10.1017/S1358246100009772