Studies in History and Philosophy of Science 45 (2014) 103–108

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Does the miracle argument embody a base rate fallacy?

Cornelis Menke

Department of Philosophy, University of Bielefeld, PO Box 100131, D-33501 Bielefeld, Germany

ARTICLE INFO

Keywords:
Miracle argument
Prediction
Novelty
Base rate fallacy

ABSTRACT

One way to reconstruct the miracle argument for scientific realism is to regard it as a statistical inference: since it is exceedingly unlikely that a false theory makes successful predictions, while it is rather likely that an approximately true theory is predictively successful, it is reasonable to infer that a predictively successful theory is at least approximately true. This reconstruction has led to the objection that the argument embodies a base rate fallacy: by focusing on successful theories one ignores the vast number of false theories some of which will be successful by mere chance.

In this paper, I shall argue that the cogency of this objection depends on the explanandum of the miracle argument. It is cogent if what is to be explained is the success of a particular theory. If, however, the explanandum of the argument is the distribution of successful predictions among competing theories, the situation is different. Since the distribution of *accidentally* successful predictions is independent of the base rate, it is possible to assess the base rate by comparing this distribution to the empirically found distribution of successful predictions among competing theories.

1. Introduction

Successful predictions of novel phenomena are often regarded as the triumphs of science. According to the miracle argument for scientific realism, the success of predictions of novel phenomena would be miraculous if the theories that made the predictions were not at least approximately true. A paradigm of a predictively successful theory is quantum electrodynamics:

> Quantum electrodynamics predicts that the magnetic moment of the electron (expressed in a well-defined unit which is unimportant for the present discussion) has the value
>
> $1.001159652201 \pm 0.000000000030$
>
> (where the "±" denotes the uncertainties in the theoretical computation, which involves several approximations), while a recent experiment gives the result
>
> $1.001159652188 \pm 0.000000000004$
>
> (where the "±" denotes the experimental uncertainties). This 11-decimal-place agreement between theory and experiment—particularly when combined with thousands of other similar though less spectacular ones—would be utterly miraculous if quantum electrodynamics were not saying something at least approximately true about the world. (Sokal & Bricmont, 2004, pp. 32–33)

One way to reconstruct the reasoning underlying this argument is to regard the miracle argument as a statistical inference: since it is exceedingly unlikely that a false theory makes successful predictions, while it is rather likely that an approximately true theory is predictively successful, it is reasonable to infer that a predictively successful theory is at least approximately true. This reconstruction has led to the objection that the argument embodies a base rate fallacy: by focusing on successful theories one ignores the vast number of false theories some of which will be successful by mere chance. Thus, the inference from a theory's predictive success to the approximate truth of the theory is not valid.

In this paper, I shall argue that the cogency of this objection depends on the explanandum of the miracle argument. The objection is cogent if we take the predictive success of a particular theory to be the explanandum of the miracle argument. The situation is different if the explanandum is the distribution of successful prediction among competing theories: if the predictions are only

E-mail address: cornelis.menke@philosophie.uni-bielefeld.de

0039-3681/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.shpsa.2013.10.006

accidentally successful, this distribution is independent of the base rate, and some distributions are hardly compatible with a base rate fallacy.

In the following, I shall first (following Colin Howson) give a statistical reconstruction of the miracle argument (2), and discuss the objection that the argument commits a base rate fallacy (3). The following two sections deal with alternative explananda of the argument: the overall success of science (4) and the distribution of successful predictions among theories (5). I shall illustrate the significance of the distribution of successful predictions by two case studies: nineteenth-century optics and theories of gravitation (6), and finally discuss the objection that the selection of successful predictions (and thus the distribution) might be severely biased.

2. A probabilistic reconstruction of the miracle argument

In *Hume's problem*, Colin Howson offers the following probabilistic reconstruction of the reasoning underlying the miracle argument:

> (i) If a theory T independently predicts some observational data E, and T is not approximately correct, then its agreement with the facts recorded in E must be accidental, a chance occurrence. (ii) The facts recorded in E are such that a chance agreement with them is exceedingly improbable. [...] (iii) We can regard such small chances as being so extraordinary unlikely that we can confidently reject the hypothesis that they are *just* chance occurrences, at any rate if there is an alternative explanation which accounts better for them. (iv) Hence we can confidently infer that T is approximately true, the smallness of the chance in (iii) being an index of the degree of confidence justified. (Howson, 2000, p. 36)

The first premise states that an independent prediction (a prediction of a novel phenomenon) ensures that if the theory T is not "approximately true" (whatever precise meaning, if any, can be given to this expression) then the agreement between prediction and empirical findings must be accidental. The second premise says that the probability of an accidental agreement should be low—so low that we can reject this possibility (third premise). The reasoning is obviously similar to tests of significance as described by Ronald Fisher, whom Howson includes "in the roll of No-Miracles workers" as the "first to see that the familiar chance model, on which the No-Miracles argument is based, could be turned into a methodological tool of apparently great power" (Howson, 2000, p. 37).

This is a reasonable reconstruction of the miracle argument. Note the similarities to the way William Whewell characterized tests of a hypotheses one and a half centuries earlier:

> The hypotheses which we accept ought to [...] *foretell* phenomena which have not yet been observed;—at least all of the same kind as those which the hypotheses was invented to explain. [...] But the evidence in favour of our induction is of a much higher and more forcible character when it enables us to explain and determine cases of a *kind different* from those which were contemplated in the formation of our hypotheses [...] No accident could give rise to such an extraordinary coincidence. No false supposition could after being adjusted to one class of phenomena, exactly represent a different class, when the agreement was unforeseen and uncontemplated. (Whewell, 1847, pp. 65–66; italics in original)

Whewell demands that an acceptable hypothesis should not only explain phenomena already observed but also predict yet unobserved ones; this ensures that the agreement between the prediction of the theory and the empirical findings is accidental if the theory is false. But that the agreement is accidental does not mean that it is unlikely. To ensure that the probability of an accidental agreement is vanishingly small, Whewell draws a line between predictions of phenomena of the *same kind* as those an hypothesis was devised to explain ("predictions") and predictions of phenomena of a *different kind*; the latter he termed "Consiliences of Inductions." Consiliences he regarded as so extraordinarily unlikely (if the hypothesis is false) that it is reasonable to dismiss the very possibility of an accidental agreement.

Howson's version of the argument is an improvement on Whewell's. At first sight, Whewell seems to be more specific as to the criteria we can use to assess how unlikely an accidental agreement is. To speak of "different kinds of phenomena," though, only suggests that there are clear criteria of telling one kind from another. In any case, we cannot rely on the theory that makes the prediction of a novel phenomenon to distinguish between different kinds of phenomena: if the theory is correct, then the phenomena it explains are all of the same kind. What changes are boundary conditions and circumstances. When we say that a theory predicts phenomena of a class different from that it was designed to account for, we generally mean that it predicts phenomena of a class we *thought* to be different—a judgment that depends on the state of our knowledge. There are, to be sure, more or less remarkable predictions; however, the difference between a small variation of an experiment and a successful prediction of a new phenomenon is a matter of degree and may depend on contingencies. Novelty is a gradual (or vague) concept, and this characteristic is better reflected in Howson's version of the argument than in Whewell's.

3. Neglecting the base rate

Howson argues that the miracle argument thus reconstructed is fallacious: the miracle argument is a base rate fallacy (Howson, 2000, pp. 52–54).

The base rate fallacy is well-known from diagnostic testing; it became famous as the "Harvard Medical School Test." A test for a disease is characterized by two properties, its sensitivity and its specificity. The sensitivity of a test refers to its ability to detect the disease; it is the probability of a positive test result given that the patient is ill. A test with a high sensitivity has a small false negative rate: Most of the patients who have the disease will test positive. The specificity of a test is the probability that the test result will be negative if the patient does not have the disease; the higher the specificity of a test, the lower is its false positive rate.

Consider a diagnostic test with following characteristics: the false negative rate is nil, thus, whoever has the disease will test positive. The false positive rate is 5 per cent; so there is a small probability that a patient who does not have the disease will nevertheless have a positive test result. What is the probability of a patient having the disease given the test result is positive? The answer is—we cannot tell. The probability depends on how common the disease is (the base rate). If only one in a thousand patients tested have the disease, the one person who actually has it will test positive, but so will approximately 50 of the remaining 999 patients who do not; thus, the probability will be less than 0.02. If, however, the disease is quite common, the probability that a patient has the disease given a positive test result will be much higher: if every second patient tested has the disease, the probability is about 0.91.

The base rate fallacy is the fallacy of simply neglecting the base rate and assuming that a positive test result is a good indicator of the patient having the disease regardless of how common the disease is. A positive test result surely increases the probability that the patient has the disease; but even given a high sensitivity and