On the Pessimistic Induction and Two Fallacies

Juha T. Saatsi†‡

The Pessimistic Induction from falsity of past theories forms a perennial argument against scientific realism. This paper considers and rebuts two recent arguments—due to Lewis (2001) and Lange (2002)—to the conclusion that the Pessimistic Induction (in its best known form) is fallacious. It re-establishes the dignity of the Pessimistic Induction by calling to mind the basic objective of the argument, and hence restores the propriety of the realist program of responding to PMI by undermining one or another of its premises.

1. Introduction. Probably the best known argument against scientific realism is the argument from Pessimistic Induction (Poincaré 1952; Putnam 1978; Laudan 1981). This argument in some form or another has been part and parcel of the realism debate for quite some time now. It is therefore interesting to come across two recent papers which both claim that the argument in its best known form is actually fallacious (Lange 2002; Lewis 2001). Here I want to re-establish the dignity of the Pessimistic Induction by calling to mind the basic objective of the argument, and hence restore the propriety of the realist program of responding to this argument by undermining one or another of its premises.

I take the *Pessimistic (Meta-)Induction* (PMI) against scientific realism to be in essence the argument employed by Laudan (1981). Laudan appeals to an historical record of successful yet false theories to argue against the connection that realists like to draw between successfulness of a theory and its approximate truth—the connection that a successful theory is deemed probably approximately true. This connection is at the heart of the realist's *No Miracles Argument* (NMA), the intuition being that the best explanation of the success of science is the approximate truth of its

†To contact the author, please write to: School of Philosophy, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK; e-mail: mat8jts@leeds.ac.uk.

‡I want to thank Angelo Cei, Steven French, and Scott Shalkowski for helpful comments, and Philip Good for sparking the initial interest to the topic.

Philosophy of Science, 72 (December 2005) pp. 1088–1098. 0031-8248/2005/7205-0036$10.00
Copyright 2005 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/508959 Published online by Cambridge University Press

theories. PMI was devised—in the hands of Laudan, at least—to deliver a lethal blow to NMA. Hence the exact content of PMI is connected in a subtle way to our understanding of NMA, and the latter must be kept firmly in mind in considering the validity of the former.

Laudan's version of PMI can be succinctly reconstructed as the following reductio (Lewis 2001, 373; Psillos 1996), which we may call [PMI]:

(1) Assume that success of a theory is a reliable test for its truth.
(2) So most current successful scientific theories are true.
(3) Then most past scientific theories are false, since they differ from current successful theories in significant ways.
(4) Many of these past theories were also successful.
(5) So successfulness of a theory is not a reliable test for its truth (since this leads to contradiction in (3) and (4))

A typical realist response to this reductio can take issue with, for example, the implicit premise of step (3) by describing (usually via careful case studies) some theoretical elements solely responsible for the successfulness of past theories in a way that renders these theories continuous with otherwise incompatible current theories, and hence candidates of approximate truth in some suitable, restricted sense. (e.g., Psillos 1999) I am personally very optimistic about such a line of response, but the purpose of this paper is not to question the premises of Laudan's argument. Here my sole purpose is to stand up for the dignity of such premise defeating work against two lines of thought that claim to remove the antirealist threat of PMI by denying the validity of the argument to begin with.

2. Lange's Turnover Fallacy. Lange (2002) presents the *turnover fallacy* as a potential source of invalidity of pessimistic inductions in general (and not just of PMI against the realist). The basic idea of this fallacy can be conveyed by the following example:

Assume there is a board of directors comprising of ten members and that you are introduced as a new member to this board replacing someone else. You are told that the company is in turmoil: there has been a change in the assemblage of the board 240 times in the past ten years, but you don't know who's been sitting in the board for how long. You pessimistically infer, inductively, that someone is going to be replaced again very soon. It could be you or it could be someone else for all you know.

You might be tempted to pessimistically infer that the probability of most of you getting the boot within a year, say, is quite high. But this would be to commit the turnover fallacy! For it could be that nine out of ten members of the board have actually sat in throughout the past ten years and it is only your 'predecessors', as it were, who came and went.

https://doi.org/10.1086/508959 Published online by Cambridge University Press

Just by knowing *the number* of personnel changes in the board does not allow you to inductively infer anything about the probability for *any one particular* individual to get replaced—all you can infer is the high probability for *someone* to get replaced.

Now consider the case of scientific PMI. Looking at the set of current, well confirmed, successful theories we may want to ask: "How likely is it that most of these theories will turn out to be false and will be replaced by new theories incompatible with them?" Given a very bad *numerical* historical record of successful yet false theories we may be tempted—vaguely remembering the intuition behind the PMI argument—to answer "Very likely." But this would be to commit the turnover fallacy! For it could be that most of the current theories have been stable throughout the historical record tracking period, and all the numerous theory changes involve the 'predecessors', as it were, of only one current theory.

Although this is a point about a type of induction in general, Lange takes it to be telling against Laudan's argument in particular. The alleged lesson is that to validly infer the wanted conclusion—that most current theories are probably false—one needs to use a premise much stronger than (3) above in an argument of slightly different form.

> . . . a pessimistic induction of a somewhat different and less familiar form is made impervious to the turnover fallacy by employing a historical premise that is not cumulative: at most past moments, most of the theories receiving wide acceptance at that moment are false (by current lights). (Lange 2002, 284)

This is significant since the usual premise "that most of the theories that have ever been accepted were false is inevitably more plausible than the needed premise: that at most past moments, most of the theories then accepted were false" (2002, 285). A fallacy is committed, Lange proposes, since a typical statement of PMI (such as Laudan's) only refers to the *number* of past false theories as an inductive basis, and yet draws a conclusion about the high likelihood of any one of our present theories to be found false and replaced in the course of future science.

It must be admitted that Lange makes a fine point about pessimistic inductions in general, but nevertheless it seems that this potential fallacy *cannot* be incorporated against the scientific PMI of interest (that is, Laudan's PMI). Here we need to be more careful about the real objective of the PMI argument—what is the conclusion being inferred exactly? To begin with, note that the conclusion (5) above makes no reference to future times: what will be found false or whether any theory shifts will take place. This argument [PMI] is therefore *not* an argument to the *time-dependent* conclusion that most of our current theories will be most likely found false and will be replaced. Rather, it is an argument to the *timeless*

https://doi.org/10.1086/508959 Published online by Cambridge University Press

conclusion that "successfulness of a theory is not a reliable test for its truth." As a matter of fact, in this conclusion no reference is made even to the probable falsity of any one theory of the current successful science; this conclusion would indeed hold even if the current theories were all likely to be true! And nonetheless the force of the argument is considerable given the key role of the claimed explanatory connection between success and approximate truth in the realist's game plan. It is interesting to notice that in the literature the term 'Pessimistic Induction', originally coined by Putnam in 1978, is invariably tagged on to the antirealist line of thought the canonical formulation of which is taken as [PMI] (i.e., Laudan's reductio argument as presented above). The failure to properly distinguish Laudan's argument from Putnam's rhetoric is behind Lange's undue optimism to be able to sidestep the antirealist worry about the historical facts of science as they are typically told.

This reading of PMI—viz. merely as something to counter NMA—may feel unintuitively *neutral* to some.[1] One may feel that PMI should have some pessimistic force on its own and not just as a reactive opposition to NMA, and we can indeed discern different levels of pessimism which PMI is sometimes taken to be an argument for. For example, witness Psillos' informal summary of Laudan's argument:

> Therefore, by a simple (meta-)induction on scientific theories, our current successful theories are likely to be false (or, at any rate, are more likely to be false than true), and many or most of the theoretical terms featuring in them will turn out to be nonreferential. (1999, 101)

This sentence perhaps typifies a more customary reading of PMI as entailing the probable falsity of any one of our current theories, and indeed this is the reading that Lange explicitly adopts. Is this reading of the argument, referring to the probable falsity of our current theories, now subject to the *turnover fallacy* as Lange suggests?

I believe not.[2] First of all, we need to notice that this new argument is

1. Laudan (1981) does not use the term PMI, but I believe this 'weak' reading of PMI is closest to the use Laudan makes of his pessimistic historical record. This version of the anti-realist's argument is obviously already damaging against the realist, given the respective objectives of the two positions: even if the PMI does not conclude that most current successful theories are probably false, the anti-realist has shown (by undermining NMA) that there is no rationale for taking these theories to be true either, and agnosticism follows. The anti-realist, of course, can be quite happy with this (cf. van Fraassen 1980).

2. This is not to say, of course, that there are no other weaknesses or fallacies that the proposed simplistic classical statistical reasoning may succumb to.

https://doi.org/10.1086/508959 Published online by Cambridge University Press

no longer just the reductio presented above.[3] Rather, we now add to the above reductio a statistical argument along the following lines, call it [PMI*]:

(1*) Of all the successful theories, current and past, most are taken to be false by the current lights.
(2*) The current theories are essentially no different from the past successful theories with respect to their "observable" properties. (Viz. properties potentially figuring in the realist's explanatory argument.)
(3*) Success of a current theory is not a reliable indicator of its truth (by the reductio argument above), and there is no other reliable indicator of truth for the current theories.
(4*) Therefore any current successful theory is probably false by statistical reasoning.

This argument concludes that any one current successful theory, *ceteris paribus*, is probably false for all we know. The ceteris paribus clause effectively amounts to the premises (2*) and (3*) above: NMA is taken to be indiscriminating so that the current observer is taken to have no advantage over the past observers in evaluating the truthlikeness of a successful theory. Furthermore, this clause should be also taken to rule out all kinds of 'relativisations' of NMA to specific scientific domains: scientific methodologies and mechanisms are taken to be homogeneous across the domains and the competing realist and antirealist arguments apply across the board. I take the content of these premises to be implicit in the standard construal of PMI.

The argument [PMI*] does not fall foul of the turnover fallacy. However, one may be tempted to *further* infer from such probable falsity the high probability of finding a theory false and it getting replaced, but such an inference would go beyond the confines of—and indeed beyond the validity of—this version of pessimistic induction. Hence a *timeless* conclusion (4*) is inferred from timeless premises and no fallacy of turnover is being committed; this fallacy requires a reference to a time dependent property (e.g., getting the boot within the next two weeks) in the conclusion but 'being false' is not such property.[4] And a further argument to the conclusion that false theories *will be* replaced in the course of future

3. The argument is usually presented as a reductio as I have presented it (cf. Laudan 1981; Lewis 2001; Psillos 1999). Lange also refers to Laudan and Psillos in his discussion of the scientific pessimistic induction.

4. Notice that there is a time-dependent part in the above quote from Psillos (1999) invalidly going beyond the confines of PMI. Curiously enough there is no such explicit mistake to be found in Lange's exposition of PMI.

https://doi.org/10.1086/508959 Published online by Cambridge University Press

science, whilst perhaps not unthinkable, is surely not part and parcel of the contemporary NMA vs. PMI debate.

Moreover, the conclusion of [PMI*] is clearly compatible with the kind of possible (asymmetric) state of affairs that Lange puts forward as problematic. Assume that all theory changes have taken place within just one domain of scientific enquiry, say. It seems, *pace* Lange, that we nonetheless have reason, ceteris paribus, to believe that all domains of enquiry are currently ridden with false theories. This is because the only feature of theories appealed to in NMA is their successfulness and not, say, the duration of their reign. Once the connection between success and truth has been demolished by [PMI], all the current successful theories (including those which we inductively have *no* reason to expect to get replaced) are on a par with all the past successful theories in one big domain of theories most of which are false, and the conclusion (4*) can be drawn. Furthermore, whilst the assumed asymmetric state of affairs undoubtedly begs for *some* explanation, it is not clear that we have any reason to think that *the best* explanation is achieved by hypothesising the stable theories to be true. What the realist needs is an argument to the conclusion that the combination of successfulness *and* long lifespan of a theory is best explained via truthlikeness, or something like that. As far as I know, no such version of NMA has yet been developed. On the other hand, our degree of confidence to realism as a *possible* explanation of the asymmetric state of affairs is significantly lowered by Laudan's PMI and the availability of numerous other explanations, together with the ceteris paribus clause.

One may, of course, have grave doubts about the ceteris paribus clause in the above portrayal of [PMI*], and many realists indeed argue that at least some current successful theories are not on a par with the past theories which are employed as the basis of the statistical inference above. But while this may offer a way to encounter this version of PMI, it does so by undermining one significant premise of the argument and not by virtue of showing it to harbour the turnover fallacy.

I prefer to follow Laudan and read the argument as the reductio [PMI]. We should notice that Laudan's argument is a somewhat atypical case of induction. Usually induction is described as an inference from the particular to the general, and it typically concerns states of affairs at future times being inferred from states of affairs at past times. But we have seen that [PMI] is not best characterised in such terms. Rather, [PMI] should be viewed as a reductio of an indiscriminating realist image—a challenge to the realist's beloved connection between success and truth. Even if none of our current theories eventually succumbed to some incompatible successors, the antirealist could nonetheless appeal to [PMI] as an anti-NMA. To do this, all that is required is a pool of theories all of which are

https://doi.org/10.1086/508959 Published online by Cambridge University Press

successful at some time or another, yet most of which have turned out to be false.

So perhaps it is better to regard this meta-induction as a *statistical argument* against the realist claim that one 'observable' feature of our theories—successfulness—is a *reliable statistical indicator* of another, 'unobservable' feature of our theories: their truth(likeness). This is exactly what Peter Lewis (2001) does and claims that it falls victim to another kind of fallacy.

3. Lewis's False Positives Fallacy. Lewis presents an altogether different rationale for regarding PMI thus understood as harbouring a fallacy. For Lewis the problem is that "the premise that many false past theories were successful does not warrant the assertion that success is not a reliable test for truth" (2001, 374). More specifically: the *fallacy of false positives* that Lewis has in mind concerns the reliability of successfulness as an indicator of (approximate) truth. The notion of statistical reliability is usually characterised in statistics literature in terms of the rates of false positives and false negatives: a reliable indicator is one for which "the false positive rate and false negative rate are both sufficiently small, where what counts as sufficiently small is determined by the context" (2001, 374–375). An instance of false positive (negative) indication is, of course, one in which the existence (absence) of an indication fails to reflect the existence (absence) of the indicated. The rate of false positives (negatives) is then calculated as the number of such cases per all negative (positive) cases.

With statistical reliability characterised in these terms Lewis then takes successfulness to be a reliable indicator of the (approximate) truth of a theory T (picked at random out of *all* theories at time t) if and only if the rate of false-yet-successful theories is small and the rate of true-but-unsuccessful theories is small. With this notion of statistical reliability at hand Lewis explains why Laudan's reductio formulation of PMI is a non sequitur:

> At a given time in the past, it may well be that false theories vastly outnumber true theories. In that case, even if only a small proportion of false theories are successful, and even if a large proportion of true theories are successful, the successful false theories may outnumber successful true theories. So the fact that successful false theories outnumber successful true theories at some time does nothing to undermine the reliability of success as a test for truth at *that* time, let alone other times. In other words, the realist can interpret Laudan's historical cases, not as evidence against the reliability of success as a test for truth, but merely as evidence of the scarcity of true theories in the past. (2001, 377)

https://doi.org/10.1086/508959 Published online by Cambridge University Press

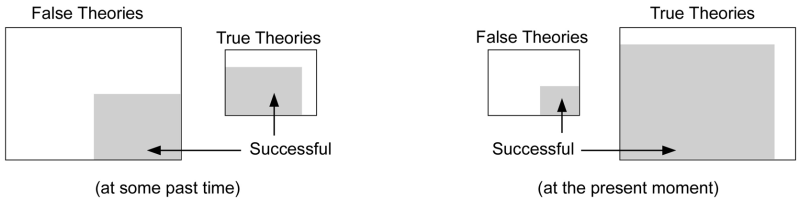


Figure 1. Domains compatible with both statistical reliability and 'bad' historical record. The ratio of successful (grey) and unsuccessful (white) is constant for each domain over time.

And to do otherwise is, Lewis proposes, to commit the fallacy of false positives.

The basic intuition behind this argument is made most clear in pictorial terms in Figure 1. We can see immediately that by having a big enough domain of *false and unsuccessful* theories we can satisfy the requirement of statistical reliability even in cases in which, somewhat unintuitively perhaps, the probability of a randomly drawn successful theory to be true is small (less than 0.5, say). At both times pictured the requirement of statistical reliability is satisfied. Furthermore, given that we take *most of our current theories to be successful,* it follows "deductively that most current theories are true, as required by the realist" (2001, 375). This Lewis takes to be a reasonable justification for regarding statistical reliability to be a notion that adequately captures the realist's appeal to the success-versus-truth connection.

So the notion of statistical reliability works for Lewis on the assumption that the statistical reference classes (relative to which statistical reliability is determined) are of the right kind and vary radically as we move from past to current theories: the domain of all theories at some past time t_p must contain a much higher proportion of false and unsuccessful theories than the domain of all current theories. This immediately raises a couple of worries regarding the overall framework in which Lewis casts his realism and allegedly sidesteps the challenge of PMI: (1) How are the crucial reference classes defined in the first place? (2) Has there really been a change in the reference classes such as to enable Lewis's response to PMI to get off the ground?

(1) First of all, it is not at all clear that the notion of reference class of statistical reliability is well-defined in the context of scientific theories. It seems that the relevant domains of *all* true theories and *all* false theories (at some time t) with respect to which the rates of false negatives and false positives are calculated are not straightforwardly definable in the way a pool of people, say, is readily given in a typical case of medical

https://doi.org/10.1086/508959 Published online by Cambridge University Press

statistics, for example. Not much has been said in the discussion so far about the putative identity conditions of theories—it just has been surmised that they could in principle be given. But whereas this assumption may be reasonable with respect to both the set of successful theories and the set of true theories, I can make no sense of the idea of delineating a non-arbitrary, well-defined collection of *both false and unsuccessful* theories.

Lewis's realism-friendly scenario which makes Laudan's historical record compatible with success being a reliable statistical indicator depends on there having been a large domain of such false and unsuccessful theories relative to which the rate of false positives is small.[5] But what exactly are the theories which are neither successful nor true? Should we count in only the theory-proposals made by eminent scientists, or perhaps all the proposals actually published in scientific journals, or what? It is easy to imagine a variety of sociological factors, say, yielding scores of unsuccessful and false theories, directly affecting the notion of reliability at stake. But why should we care about *those* theories? It just seems that the debate between NMA and PMI does not involve unsuccessful and false theories (or true yet unsuccessful, for that matter) in anything like the way Lewis projects.

But perhaps a case could be made that the realist should really give us some rough idea of how many false and unsuccessful theories there are per each successful one, given that NMA—being a form of inference to the best explanation—seems to hang on the assumption that this ratio is not high enough to explain away the 'miracle' of successful science by the mere number of trials. But however we decided to delineate the domain of all theories it should not be the case that the realist explanation is held hostage to contingent matters regarding the number of false and unsuccessful theories in the strict manner implied by Lewis's strategy; realism simply cannot depend on the alleged (contingent) fact that most current theories are successful! Rather, it is implicit in the No-Miracles intuition that any feasible fluctuation in the number of false and unsuccessful theories—feasible to science as we know it—is not large enough to overthrow NMA as the *best* explanation around.[6]

5. Lewis's proposal for testing the history of science for the pessimistic conclusion of PMI in a *valid* way consists of taking "a random sample of theories which are known to be false, and show[ing] that a significant proportion of them are nevertheless successful" (2001, 378). The worry now is that this testing cannot be done since the domain in question is ill-defined.

6. Whether or not this intuition holds is another matter, of course. The point is that Lewis has not only provided a response to PMI but also a particular understanding of realism to go with it. The problems with the former really spring from the inadequacy of the latter.

https://doi.org/10.1086/508959 Published online by Cambridge University Press

(2) So has there been a change in the reference class of the kind that Lewisian realism requires? The idea is that realism only requires that most of our current theories are true which deductively follows, given good statistical reliability of success as an indicator of truth, from the premise that most of our current theories are successful. That is, given any one successful theory—current or past—the best explanation for a Lewisian realist of its successfulness is that *either* it is (approximately) true *or* it is a member of a huge domain of false theories a small portion of which are successful. Regarding past successful science, at least, this is fully amenable to an antirealist reading. To an antirealist like Bas van Fraassen—who persistently denies the force of the No-Miracles Argument—an explanation such as the above is good enough and fully consonant with his Darwinian selectionist image of science. For van Fraassen, of course, this picture fits the bill with respect to current science just as well; that is, he denies the initial premise of Lewis's that most of our current theories are successful. But the soundness of that premise is neither necessary nor sufficient for the realist to make a case against van Fraassen; what is required is NMA as typically understood and the intuition that (approximate) truth is thus connected to successfulness—and for that intuition to have bite is for it to have bite *at all times*, regardless of the number of false and unsuccessful theories present at the time in question.[7]

As a matter of fact, Lewis's unorthodox formulation of the realist position seems to beg the question against this point to begin with. According to Lewis "convergent realism usually includes the thesis that most of our current theories are true" (2001, 371). But this is certainly an unreasonably strong thesis for any realist to aspire after: contingent matters regarding the number of false and unsuccessful theories produced by the scientific community depends on factors quite independent from realism and NMA—or so the realist argues—which is why convergence is typically characterised in terms of increasing level of 'truthlikeness' in a sequence of *successful* theories of cumulative empirical adequacy. Lewis's convergent realist is committed "to the empirical claim that successful theories were rare in the past and are common today" (2001, 377). Such commitment is not generally acknowledged to be part of any contemporary realist position. And it better not be! Keeping in mind how strict a qualification 'successful' can be for the realist and casually glancing

7. Unless, of course, that 'number' is *so high* as to undermine the credibility of NMA as the best explanation altogether as explained in (1) above. Lewis stresses "the inference that the realist wishes to draw from the success of most *current* theories to their truth" (2001, 378, my italics) but this requires that the realist accounts for some principled difference between the current and the past. And Lewis does not provide such an account.

https://doi.org/10.1086/508959 Published online by Cambridge University Press

through the *Journal of Mathematical Physics*, for example, one is bound to be convinced of the sheer incredibility of this premise upon which realism à la Lewis is erected.

4. Conclusion. Despite Lange's and Lewis's respective attempts to short circuit the Pessimistic Meta-Induction it remains a powerful force to be reckoned with. There is no easy way out for the realist; one or another of the premises must be defeated. To get properly started with this task the realist ought to recognise the variety of forms this intuitively straightforward argument can take when looked at in closer detail. This paper has focused only on how PMI should *not* be understood and much work remains to be done to understand the subtle interplay between PMI and NMA vis-à-vis the notion of success as an indicator of (approximate) truth. To achieve an adequate account of this notion we need to appreciate the timeless character of PMI as a reductio of NMA and not construe the latter in terms of mere statistical reliability.[8]

REFERENCES

Lange, Marc (2002), "Baseball, Pessimistic Inductions and the Turnover Fallacy", *Analysis* 62: 281–285.
Laudan, Larry (1981), "A Confutation of Convergent Realism", *Philosophy of Science* 48: 19–49.
Lewis, Peter (2001), "Why the Pessimistic Induction Is a Fallacy", *Synthese* 129: 371–380.
Magnus, P. D., and Craig Callender (2004), "Realist Ennui and the Base Rate Fallacy", *Philosophy of Science* 71: 320–338.
Poincaré, Henry ([1902] 1952), *Science and Hypothesis*. New York: Dover. Originally published as *La science et l'hypothèse*. Paris: Flammarion.
Psillos, Stathis (1996), "Scientific Realism and the 'Pessimistic Induction'", *Philosophy of Science* 63: S306–S314.
——— (1999), *Scientific Realism: How Science Tracks Truth*. London: Routledge.
Putnam, Hilary (1978), *Meaning and the Moral Sciences*. London: Routledge.
van Fraassen, Bas (1980), *The Scientific Image*. Oxford: Oxford University Press.

8. If we insist on casting the debate in such purely statistical terms, then the whole debate threatens to crumble as ultimately nonsensical twiddling of probabilities with undefined and indefinable base rates. (Cf. Magnus and Callender (2004).)

https://doi.org/10.1086/508959 Published online by Cambridge University Press