7-2004

Realist Ennui and the Base Rate Fallacy

P.D. Magnus
University at Albany, State University of New York, pmagnus@albany.edu

Craig Callender
University of California - San Diego, ccallender@ucsd.edu

Realist Ennui and the Base Rate Fallacy*

P. D. Magnus and Craig Callender†‡

The no-miracles argument and the pessimistic induction are arguably the main considerations for and against scientific realism. Recently these arguments have been accused of embodying a familiar, seductive fallacy. In each case, we are tricked by a base rate fallacy, one much-discussed in the psychological literature. In this paper we consider this accusation and use it as an explanation for why the two most prominent 'wholesale' arguments in the literature seem irresolvable. Framed probabilistically, we can see very clearly why realists and anti-realists have been talking past one another. We then formulate a dilemma for advocates of either argument, answer potential objections to our criticism, discuss what remains (if anything) of these two major arguments, and then speculate about a future philosophy of science freed from these two arguments. In so doing, we connect the point about base rates to the wholesale/retail distinction; we believe it hints at an answer of how to distinguish profitable from unprofitable realism debates. In short, we offer a probabilistic analysis of the feeling of *ennui* afflicting contemporary philosophy of science.

1. Introduction. A recently fashionable claim in philosophy of science is that the realism/anti-realism debate ought to be dissolved rather than solved. There is a feeling that the debate is not entirely well-formed, that the disputants are speaking past one another. Motivating this view, Blackburn writes:

> The issue of realism…is apt to prompt a particularly acute gestalt-switch. On the one hand it seems absurd, a Berkeleian folly, to

*Received June 2003; revised November 2003.

†To contact the authors write to: P. D. Magnus, University at Albany, SUNY, Department of Philosophy, HU 257, Albany, NY 12222 e-mail: pmagnus@fecundity.com. Craig Callender, Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla CA 92093-0119; e-mail: ccallender@ucsd.edu.

‡This paper grew out of discussions in the scientific realism reading group at the University of California, San Diego in 2001–2; we are grateful for those discussions and to Mark Newman, Carl Hoefer, Peter Lewis, Elliott Sober, Stathis Psillos, Colin Howson, Jonathan Cohen, and two anonymous referees for comments on previous versions of this paper.

Philosophy of Science, 71 (July 2004) pp. 320–338. 0031-8248/2004/7103-0003$10.00

> question the reality of the objects of common-sense, of core scientific theory. On the other hand realism is often seen as demanding the mythical God's eye view, whereby we step out of our own skins, and comment on the extent to which our best scientific theory corresponds with an independent reality. . . . In the one view realism seems almost indisputably true, and in another equally obviously false or undiscussable. So there is every opening for debates in which each side talks past each other. (2002, 112)

He then argues that, when clearly stated, a "surprising 'quietism' or pessimism about a metatheoretical position begins to seem attractive" (2002, 111). In explaining the success of some piece of science, there is no "getting behind the explanation," he writes; the best explanation for the success of the hypothesis that the world is round is that the world is round. There is no "further set of data about science (its success) that required something like an independent, sideways explanation" (2002, 130). Maddy draws a similar conclusion, urging that we not "add extra-scientific standards of justification to our repertoire" (2001, 47–48). Adding these extra-scientific standards yields nothing but dead-end debates over realism. Though distinct, Maddy and Blackburn's positions are both descendants of Fine's well-known attempt to dissolve the realism debate (1984).

The issue may be clarified by distinguishing what we call *retail* arguments for realism (arguments about specific kinds of things such as neutrinos, for instance) from *wholesale* arguments (arguments about *all* or *most* of the entities posited in our best scientific theories). Wholesale arguments promise a conclusion that applies to all mature science. Wholesale realism seeks to explain the success of science in general; wholesale anti-realism seeks to explain the history of science in general. Dissolving the debate, we suggest, involves attending to the retail arguments without trying to settle the debate in an all-or-nothing, wholesale manner. Dissolvers want us to answer the question, 'Are there atoms?,' by referring to the same evidence scientists use to support the atomic hypothesis; e.g., Einstein and Smoluchowski's Brownian motion theory and the experiments by Perrin in 1908. But they do not want us to answer by appealing to this as an instance of a more general claim, namely, that the theory and experiment of Einstein, Smoluchowski, and Perrin are part of a mature science and the posits of mature science are generally true. We sympathize with potential dissolvers of the realism debate, inasmuch as it concerns wholesale arguments. However, the exact reasons for being a dissolver are often unclear. Why exactly are questions like 'What explains the reliability and success of science?' taboo? That it remains unsolved is not a sufficient reason to think that a philosophical problem is illegitimate or even insoluble. We ask and answer plenty of similar questions: 'Is this mathematical modeling technique reliable?' 'What explains the success of bright silver fishing lures?' 'Is the eye generally reliable?' One is left wondering when to solve and when to dissolve.

Here we are concerned to provide a clear reason for believing that the wholesale realism debate should be dissolved, a reason that does not rely on contentious claims about the nature of truth, the status of normative epistemology, and what-have-you.[1] We hope that our analysis explains what dissolvers find objectionable—that some debates about scientific realism amount to adamant, futile table thumping—without the baggage of some of their epistemological analyses. We do this by considering the two most powerful wholesale arguments in the literature, the no-miracles argument for scientific realism and the pessimistic induction for anti-realism. The former has been dubbed the "ultimate argument" for realism, and anti-realists from Duhem through Laudan have relied primarily on the latter. Worrall (1989) recently billed these two arguments as the main considerations for and against realism. They pull in opposite directions with comparable, perhaps even balanced, force. Yet Colin Howson and Peter Lewis have independently suggested that these arguments embody a familiar, seductive fallacy. Howson (2000, 52–54) makes the point about the no-miracles argument;[2] Lewis (2001) about the pessimistic induction. Interestingly, neither applies the point to both arguments. With each argument we are tricked by a base rate fallacy. If this is correct and the intuitions marshaled by argument are phantoms of that fallacy, then there is much sound and fury in debates over realism that signifies nothing.

In this paper we appeal to the arguments of Howson and Lewis and use them as part of an explanation for why the two most prominent wholesale arguments in the literature seem irresolvable. Though the formal reconstruction of these arguments is artificial in various respects, reconstructing the arguments this way allows us to see a crucial flaw in both. In particular, framed as Howson and Lewis would have it, we can see very clearly why realists and anti-realists have been talking past one another. We then formulate a dilemma for advocates of either argument, answer potential objections to our criticism, discuss what remains (if anything) of these two major arguments, and then speculate about a future philosophy of science freed of these two arguments. In so doing, we connect the point about base rates to the wholesale/retail distinction; we believe it hints at an answer of how to distinguish profitable from unprofitable realism debates. In short, we offer a probabilistic analysis of the feeling of *ennui* afflicting contemporary philosophy of science.

1. See Psillos (1999) for criticisms of Fine on these points.

2. Howson's attention was drawn to the fallacy by Korb 1991.

2. The No-Miracles Argument. Some scientific theories are astonishingly successful. Classical thermodynamics has made correct predictions about all manner of substance for almost two centuries; meanwhile, quantum electrodynamics successfully predicts an electron's magnetic moment to more than one part in a billion. The no-miracles argument claims that if these theories did not latch on to the world in some way—that is, if they were not approximately true—then this success would be nothing short of miraculous. Some theories are too good not to be true.

The argument might be formalized in this way: For any theory *x,* let *Sx* stand for the expression '*x* is successful' and let *Tx* stand for the expression '*x* is true.' Let $\neg A$ be the negation of *A* and let $\Pr(A|B)$ be the probability of *A* conditional on *B*. We may now gloss the argument in this way for some current theory: [1] The theory *h* is very likely successful. [2] If *h* were true, it would be very likely to be successful. [3] If *h* were false, it would not be likely to be successful. [4] Therefore, there is a high probability that *h* is true. Formalizing this version of the argument yields:

$$\Pr(Sh) \gg 0 \tag{1}$$

$$\Pr(Sh|Th) \gg 0 \tag{2}$$

$$\Pr(Sh|\neg Th) \ll 1 \tag{3}$$

$$\therefore \Pr(Th|Sh) \gg 0 \tag{4}$$

For instance, assume the values .9, .95, and .05 for $\Pr(Sh)$, $\Pr(Sh|Th)$, and $\Pr(Sh|\neg Th)$ respectively; then, $\Pr(Th|Sh) = .997$. The reconstruction is crude, but it is sufficient to frame much of the action surrounding no-miracles reasoning. One can, as Howson (2000, 45) does, complain that it is silly to think that there is a well-defined outcome space, probability distribution, etc. We sympathize but do not want to decide the matter on a technicality. One can also, as Howson does, point out that 'gruesome' hypotheses short circuit the argument, for there are an infinity of these that will make the same predictions as our successful theories. But we don't think most realists see the no-miracles argument as solving the problem of induction; rightly or wrongly, that problem is being bracketed here (assumed 'solved' or 'unresolvable'). The crux of the debate then becomes assessing—often in a qualitative way—the relevant probabilities. Many anti-realists have directed their objections at forms of (3). Larry Laudan, for instance, notes that many past theories which were successful proved in the fullness of time not to be true; in effect, Laudan is chipping away at (3), recommending a higher value for $\Pr(Sh|\neg Th)$ by considering the history of science.

3. The Pessimistic Induction. Reflections on the history of science motivate the anti-realist argument. Considering past theories, we observe that many once successful theories are now believed to be false, e.g., caloric theory. We sample the successful theories of the past and find that many or most of them were false. We generalize and, by induction on these cases, evaluate $\Pr(\neg Tx|Sx)$ as being rather high for an arbitrary theory *x*. This holds for our present successful theories; hence we should think that they, too, will turn out to be false.[3] Contemporary discussions of this argument begin with Laudan (1981), who writes:

> I daresay that for every highly successful theory in the past of science which we now believe to be a genuinely referring theory, one could find half a dozen once successful theories which we now regard as substantially non-referring. (1981, 35)

If Laudan can show that $\Pr(\neg Th|Sh) = \frac{6}{7}$, then—on the further assumption that $\Pr(Sh)=1$—it follows that $\Pr(Th) = \frac{1}{7}$. This is the *pessimistic induction.* It invites us to conclude that the probability of some present, successful theory being true is rather low.[4]

4. The Fallacy. To see the point clearly, consider a case removed from the giddy heights of realism. Suppose that there is some disease that in the course of time inevitably produces unique and readily identifiable symptoms. Imagine that there is some reliable test for this disease that can identify people infected with it who have not yet developed symptoms. Let *Dx* stand for '*x* has the disease' and let *Px* stand for '*x* tests positive for the disease.' Now suppose that if someone has the disease, then they are sure to test positive; that is, assume $\Pr(Px|Dx)=1$. Suppose further that if someone is not infected there is some small chance that they will nonetheless test positive; that is, there is a chance that a positive test result will be a *false positive.* (In the language of significance tests, a false positive is a Type II error.) Let that chance of a false positive be five

3. One referee suggests limiting the induction to *fundamental* theories, since all of which but the present ones are known to be false. However, it is not clear which theories are fundamental. Is quantum field theory fundamental? Looking at problems with renormalization, many think not. Then of course there is the question of its compatibility with general relativity. Also, which theories are considered fundamental may change with time. Thermodynamics, devised under a caloric interpretation, might have been fundamental; and later, under an energist interpretation like that Ostwald's might have been fundamental again. Yet today it is clearly not fundamental. The restriction to fundamental theories thus runs the danger of limiting the induction base to nil.

4. The argument is sometimes called a metainduction, since it generalizes over past inductive inferences, but of course the sample includes past scientific inferences whether inductive or otherwise.

percent, i.e. $\Pr(Px|\neg Dx) = .05$. Now suppose a patient *a* tests positive for the disease. What is the probability that she actually has it?

It is tempting to say that $\Pr(Da|Pa) = .95$ or at least to assign a high value to $\Pr(Da|Pa)$. We can construct the inference so as to be formally analogous to the no-miracles argument as we formulated it above: From $\Pr(Pa) = 1$, $\Pr(Pa|Da) = 1$, $Pr(Pa|\neg Da) = .05$, infer $\Pr(Da|Pa) \gg 0$. As any elementary statistics text will remind us, however, we must consider the sample from which this patient was drawn. Suppose, among the people tested, the disease is rare. If only 1 in 1000 people has the disease, then given the assumptions above we should expect about 51 in 1000 to test positive. Of those 51 who test positive, only 1 will actually have the disease. Thus, the chance that this patient who tests positive has the disease would be 1 in 51; $\Pr(Da|Pa) = .02$. Thinking that $\Pr(Da|Pa)$ *must* be rather high is the false positives fallacy, a form of *base rate neglect.*[5]

Now return to the no-miracles argument. Setting worries about (3) aside, there is an additional premise hidden in this formulation of the argument: *h* is, by stipulation, some current theory of a mature science. Let $\mathcal{H}$ be the set of present candidate theories. Now the no-miracles argument takes this form for all *x*:

$$\Pr(Sx|x\in\mathcal{H}) \gg 0 \tag{5}$$

$$\Pr(Sx|Tx \,\&\, x\in\mathcal{H}) \gg 0 \tag{6}$$

$$\Pr(Sx|\neg Tx \,\&\, x\in\mathcal{H}) \ll 1 \tag{7}$$

$$\therefore \Pr(Tx|Sx \,\&\, x\in\mathcal{H}) \gg 0 \tag{8}$$

The argument revised in this way is still valid, but its soundness should tug less at our intuitions. Premise (5) will hold only if any arbitrary member of the population is likely to be successful. On the assumption that success is a reliable indicator of truth, this is tantamount to assuming that any arbitrary member of the population is likely to be true. If $\Pr(Tx|x\in\mathcal{H})$ is low (and how can we know if it is not?), then (5) fails and the conclusion does not follow.

We might attempt to assess (5) by inspecting the pool of theories, $\mathcal{H}$. We defined $\mathcal{H}$ as the set of candidate theories, but what theories were candidates for our present mature sciences? 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 (5) obtains.

5. The argument for $\Pr(Da|Pa) \gg 0$ fails because the assumption $\Pr(Pa) = 1$ is not true given this population.

The realist may insist that $\mathcal{H}$ is the set of theories *actually professed* by our mature sciences and, thus, that we can assess (5) in a straightforward way: sample the overt declarations of mature sciences and check their success. Yet this would be a biased sample, since theories in mature sciences were chosen (in no small part) because they were successful. Any theory $x \in \mathcal{H}$ would probably be successful, just on account of its membership in $\mathcal{H}$. Thus, (6) would hold trivially and not on account of any connection between success and truth; (7) would simply be false, since almost all $x \in \mathcal{H}$ would be successful, whether true or not. The realist would thus avoid base rate neglect, but at the cost of sample selection bias.

Suppose realists win the point that success-to-truth inference is reliable, even that it is as reliable as our hypothetical diagnostic test. If true theories are rare enough—that is, if $\Pr(Tx|x \in \mathcal{H}) \ll 1$—then $\Pr(Tx|Sx)$ may be very low indeed. The no-miracles argument turns on neglecting this base rate.

The same goes for the pessimistic induction. The anti-realist hoped to show $\Pr(\neg Th|Sh) \gg 0$ on the basis of the historical record. Mindful of base rates, we should be careful to include the sample population explicitly in formulating the argument. Let the set of past scientific hypotheses be $\mathcal{H}_p$. Suppose that historical enquiry does show that, for an arbitrary member of $\mathcal{H}_p$, $\Pr(\neg Tx|Sx) = \frac{6}{7}$; that is, suppose that

$$\Pr(\neg Tx \,|\, Sx \,\&\, x \in \mathcal{H}_p) = \frac{6}{7}. \tag{9}$$

If (9) were true, it would not have a direct bearing on the reliability of the success-to-truth inference as expressed in $\Pr(Th|Sh)$. Past successful theories might typically be false because most *successful theories* are false ($\Pr(\neg Tx|Sx) \gg 0$), but they might typically be false instead because most *past theories* were false ($\Pr(Tx|x \in \mathcal{H}_p) \ll 1$).[6] If the latter, then the realist may insist that the population of past theories $\mathcal{H}_p$ was a different kettle of fish than the population of present theories $\mathcal{H}$. On that assumption, (9) is compatible with the realists' desire to infer truth from the success of *present* theories, expressed as (8).

Pressing the point against the realist directly would require showing that the proportion of true to false theories is now about the same as it has always been; i.e. that

$$\Pr(Tx \,|\, x \in \mathcal{H}) \approx \Pr(Tx \,|\, x \in \mathcal{H}_p). \tag{10}$$

6. Of course, (9) could also be explained by some interaction of these two factors.

The realist will contest (10), of course, and insist that the $\Pr(Tx|x \in \mathcal{H}_i)$ increases through history. The anti-realist will probably deny that we have any way of evaluating (10), since evaluating it would presume we could determine matters of truth. Lewis suggests how a new induction might address this problem:

> Given that past theories are not automatically successful, the only way to ascertain whether the history of science supports convergent realism or undermines it would be to conduct a thorough survey of past theories, true and false, successful and unsuccessful. A moment's reflection on the difficulties of such a survey perhaps indicates why nothing like it has been attempted. (2001, 379)

Indeed, it is unlikely that there is any neutral way to count up past theories. Mindful of change over time and inter-personal variation, how many theories of Newtonian mechanics were there?

We can now offer a diagnosis of the feeling of futility in the realism debates. By pointing to apparently successful but false theories, anti-realists responding to the no-miracles argument seek to increase $Pr(\neg Th|Sh)$. By restricting the pool of theories under consideration to those that meet various strict conditions like maturity, novelty, and so on, realists responding to the pessimistic induction seek to lower the value for $Pr(Sx|\neg Th)$. (That is, they are essentially replacing S with new, more stringent conceptions of success, S^*.) One is working on the likelihood, the other on the probability. Cases might shift the above probabilities/likelihoods, but it won't matter all that much. It won't matter that much because to connect the likelihood with the probability there is a third crucial ingredient, the base rates of S and T in the population, and these aren't talked about—for good reason.

5. A Dilemma. Perhaps the most natural counter-argument is to suppose that the base rates or priors are determined in some way. Let's consider whether this might be so. In the case of our imagined disease, the problem of false positives need not undermine the usefulness of the test. If some group has been exposed to the disease, we might have good reason to think that many of them have it—for instance if the disease is highly contagious and all members of the group were exposed to it. Even without such knowledge, we might quarantine anyone who tests positive. If only 1 in 1000 has the disease, then we should expect quarantine to inconvenience 50 healthy people. However, this would contain the disease, and it would certainly be better than inconveniencing 999 healthy people as we would if we quarantined potential victims without testing. So, the test may still be a useful thing. It may also be that we have superior but more expensive tests. We could then save money by testing only 51 people rather than all 1000 with the expensive test. Also, we know that infected people will ultimately eventually break out with symptoms, giving us an independent way of checking to see if a positive result was a false positive.

The realist will be hard-pressed to save the success-to-truth inference in these ways. Selecting successful theories may increase the probability that we will have a true one, just as quarantining people who test positive increases the proportion of people under consideration who have the disease from 1 in 1000 to 1 in 51. Success of course is a good thing. Yet the true theories will not break out with any other symptoms—they will merely continue to be true. Nor are there further tests. This is a crucial point. The realist gambit is in a way posing further tests and looking for symptoms of a theory's truth. For example, they look for empirically successful theories that 'break out' with novel predictions, maturity, unification, etc. Granting realist claims, whittling down the pool like this will increase the positive probability shift. Settle on a definition of success, though, and we still need to know base rates; unless we know them, these shifts are useless for helping advance the no-miracles argument. The realist is faced with a dilemma: Either there is a way of knowing the approximate base rate of truth among our current theories or there is not. If there is, then we must have some independent grounds for thinking that a theory is very likely true; yet if we had such grounds, the no-miracles argument would be superfluous. If there is not, then the no-miracles argument requires an assumption that some significant proportion of our current theories are true; yet that would beg the question against the anti-realist.

Anti-realists, meanwhile, can keep finding successful false theories. By finding more and more of them, they can eventually drive the realist to make a choice between giving up the reliability of the success-to-truth inference and making the base rate of true theories very low. If Laudan found *hundreds and hundreds* of successful but false theories, this ought to make the realist squirm a bit. Even so, as reflection on the above numbers shows, realists can always find a value of the base rate that suits their purposes. More importantly, given the tremendous controversy over merely a handful of cases (aether, phlogiston, the wave theory of light, caloric, etc.), it is unlikely that a sufficient number of uncontroversial successful-but-false theories would ever be found to even slightly nudge a realist cognizant of the base rate fallacy.

Of course, if the base rates are interpreted as prior probabilities in a subjective Bayesian framework, the problem can be avoided. If some interval of values can be deemed the subjectively 'reasonable' priors, then—as Dorling (1992) shows—the Bayesian can be a realist or an anti-realist. Given certain priors and evidence, Bayesians will be committed to realism about particular entities; given other priors, Bayesians will be committed to anti-realism. But notice that these are retail arguments about particular entities. In Dorling's cases, there may well be some plausible set of priors available, priors that realists and anti-realists could have agreed on before all the evidence came in. In the present wholesale case, however, where the entire fate of realism or anti-realism seems bound up with the priors, we can't imagine how one could find a reasonable set of priors.

Nor can we imagine how one could find the objective ratio of true theories among all past theories. We might admit that *in principle* there is such a ratio, but there is in practice no way to ascertain it. And if there were, there would then be no *need* to make either argument in the first place.

6. Doing Justice to Intuitions. The fallaciousness of the no-miracles argument and the pessimistic induction may come as no surprise. Worrall suggests that they might better be called "considerations" for and against realism. Although not valid arguments, they reveal a deep realist intuition and a deep anti-realist intuition (Worrall 1989, 101).[7] What do the considerations amount to? Well, any story about science and its history had better allow for the fact that sometimes scientists have thought highly of theories that later turned out false; and any story about science had better reflect the fact that science sometimes makes very detailed predictions that are confirmed with beautiful precision.

Yet one may feel an urge to say more than this. Even after the exposure of the fallacy, the two 'arguments' may still prime us to feel realist or anti-realist impulses. Whether we should acknowledge these impulses as probative intuitions is another matter.

One may object: 'There must be something wrong with the foregoing discussion! Debates that so many smart people have taken seriously can't be predicated on such an elementary fallacy.' Yet there is evidence that people are prone to commit the false positives fallacy. Psychologists have been aware of the phenomenon at least since Meehl and Rosen's classic paper (1955), and it has been demonstrated experimentally by Tversky and Kahneman (1982) and others following after them. It is not confined to the psychologist's subject of choice, the college undergraduate; Casscells, et al. (1978) demonstrate its prevalence among physicians. There is no reason to expect that philosophers would be immune.

The propensity to commit this fallacy explains why these two 'considerations' have rhetorical force even after their logical force has been shown lacking. Just as we are susceptible to optical illusions after we understand them for what they are, we may continue to be susceptible to these logical illusions. We should not try to *do justice* to the intuitions,

7. In response to the tension between these two, he suggests *structural realism* as a position that might sit well with both.

except in the sense that the court may do justice to a killer. This diagnosis, if it is correct, means that the major considerations for and against realism come to naught.

We have argued so far that the no-miracles argument and the pessimistic induction are instances of the same fallacy, that research suggests that educated people are apt to commit this fallacy, and thus that the intuitive appeal of the no-miracles argument and the pessimistic induction should not be taken as a sign that they have any probative force. This does not entail that there is no issue between realism and anti-realism. Nevertheless, one wonders what debates about realism would be like without appeals to either of these attractive but fallacious arguments. We return to this train of thought in Section 8.

7. Reformulations. Following Howson and Lewis, we've formulated both arguments as statistical inferences of a particular kind. Perhaps the point is merely an artifact of a deficient formulation of the argument?

The no-miracles argument might be reconstructed as a likelihood inference (cf. Sober 1990). As we've already noted, this could provide *some* incremental confirmation of a theory's truth. That is, learning that a theory is successful gives us a reason to revise upwards the probability that it is true. Nevertheless, nothing guarantees that the resulting probability will be high enough to merit belief. Recall the hypothetical disease: The probability that an arbitrary person has the disease is .001; testing positive shifts this upward to .02. Yet the realist seems to think that the probability of the truth of theories is close to 1 or at least that it's greater than .5. An argument purely in terms of likelihoods cannot secure this conclusion.

The realist might object that the no-miracles argument is not fallacious when supplemented with an appropriate auxiliary assumption. As argued above, however, suitable assumptions (e.g., that there is a significant proportion of true theories in $\mathcal{H}$) either beg the question or render the no-miracles argument redundant. The realist is still free to suggest that the supplemented no-miracles argument is valid and that he the realist thinks it is sound. It then joins inductive defenses of induction and abductive defenses of inference to the best explanation. Here we enter murky waters and cannot do justice to the subtleties of the discussion in the literature (see Howson 2000, chap. 10; Lipton 1994; Psillos 1999).

It might be better to reconstruct the argument such that the chances and miracles are not expressed as probabilities.[8] It is common to say that the no-miracles argument is not a probabilistic argument at all, but instead an *inference to the best explanation* (IBE). Critics of IBE complain that the

8. As Howson points out (and we agree), it is doubtful that there are really well-defined sample spaces and probability distributions over the space of all theories anyway.

argument is viciously circular—anti-realists doubt that IBE is in fact a reliable inferential method, so it can't be used to defend realism. Defenders of the argument respond that there is 'good' circular and 'bad' circular, that if the argument is only as bad as deductive defenses of deductions it is not so bad. The upshot of these debates seems to be that defenders of IBE view the IBE no-miracles argument as having the rhetorical force to make the hearer believe what the hearer believes. Perhaps, as Lipton (1994) argues, these 'sermons to the choir' have a legitimate purpose, but they cannot settle debates over realism.[9] As such, we do not see how the no-miracles argument can be meaningfully saved by another formulation.

What about the pessimistic induction? One might defend it as a classical statistical inference: If the theories are drawn at random from the history of science, then we may infer from the theories sampled to the population of all scientific theories. The question, of course, is whether it's appropriate to think of scientific theories as balls in an urn. For a classical statistical inference, the draws from the urn must be independent and identically-distributed. Cases offered for the pessimistic induction are not drawn in this way, since they all come from the past. This is analogous to pouring half of the balls from the original urn into a second urn and taking random draws only from this second urn. It will only be legitimate to generalize from the second urn to the original urn if the distribution of balls in the second urn is the same as the distribution in the original. The analogous assumption in the historical case is that the distribution of true theories in the past is the same is the distribution of true theories in the present and future—that is, it is just the contested assumption about base rates. So we don't see how even a completely random sampling of *past* theories could resolve the problem.

We can imagine IBE formulations of the pessimistic induction: the absence of any robust connection between success and truth might be offered as the best explanation for the eventual failure of past successful theories. This is a tantalizing suggestion. We leave it undeveloped, since we doubt that any anti-realists would be charmed by such a schizophrenic defense of anti-realism in the first place.

Perhaps there is some formulation of one or the other argument that escapes these worries. We note simply that the obvious formulations do not. Consider, then, a realist who finds realism appealing because of a pretheoretic intuition that the success of science could not be a miracle. Is it

9. Psillos (1999) defends a version of the no-miracles argument that *presumes* realism and then aims to show that IBE is reliable. We have no immediate quarrel with such an approach, since it grants what we hope to demonstrate—viz. that the no-miracles argument fails as a probative argument for realism.

plausible to think that her intuition harkens to some elusive formulation of the argument? Isn't it more plausible to think that she has the obvious version of the argument in mind, especially since the fallacy in the obvious version is one that even educated people are prone to commit? The same holds for an anti-realist who finds anti-realism appealing on account of the history of science. Since he has a psychological propensity to neglect base rates in a way that would make the obvious version of the argument convincing for him, then why think his intuitions anticipate some as-yet-undiscovered argument?

8. Imagining a Future. Imagine, if you will, what the literature on scientific realism would be like if we set aside no-miracles arguments and pessimistic inductions. As we mentioned at the outset, the no-miracles arguments and pessimistic induction are wholesale arguments, in that the conclusions are supposed to hold for all (or most) of the theories of our present, mature sciences. If we eschew these two arguments, we might look for some other wholesale motivation to settle realism debates *en masse* or instead for retail arguments that resolve questions only about particular kinds or individuals.

8.1. Other Wholesale Arguments. The *underdetermination of theory* by evidence, as it is often construed, is a wholesale argument for anti-realism about unobservable entities. Every theory has empirically equivalent rivals, we are told, and there is no epistemic distinction to be made between empirically equivalent theories. This is not a statistical argument, since we are promised that not just most but *all* theories have empirically equivalent rivals. Thus, there is no population under consideration and there are no base rates to neglect.

Yet, anyone hoping to exploit underdetermination in this way faces a dilemma. First horn: The underdetermination should not be too sweeping. Since the anti-realist has no better answer to the problem of induction than the realist has got, underdetermination threatens to sweep away predictions about observables just as it sweeps away claims about unobservables. That way lies scepticism. Second horn: The underdeterminaton must be sweeping enough that the details of particular cases will not obviate it. In contemporary quantum physics, for instance, there are several theories which are (plausibly) underdetermined with respect to one another. Nevertheless, as Cordero argues, the common ground between these several theories "does manage to tell us a great deal about what actually exists and what it is like" (2001, S310). Some modest realism would be justified *in this case,* so putatively wholesale underdetermination would give way to retail evaluation of particular scientific episodes. The dilemma for wholesale underdetermination arguments, then, is how to find underdetermination strong enough to apply without reference to particulars but not so strong as to yield scepticism.[10]

The *conjunction argument* is another wholesale consideration that does not rely on neglecting base rates. When chemists tell us that an object is 'negatively charged,' we expect that object to act according to electrostatic laws governing things that are 'negatively charged.' If we were to consider only the empirical parts of each model, such prediction would be illegitimate. We thus derive a predication from the chemical and physical theories taken together, considering not only unobservable parts of the theories but also assuming that a shared phrase ('negatively charged') has the same meaning in each theory.[11]

Yet, it's not *always* appropriate to apply the conjunction of our best scientific theories. This is most obvious when our best accounts in different domains are logically inconsistent; for instance, the spacetime metric being dynamical in general relativity and non-dynamical in quantum theory (see Callender and Huggett 2001). Some conjunctions are appropriate and others aren't. Attending to when conjunction is good practice and when it is not demands attending to the details of specific cases. It may deliver realism about *some* entities, but it would mean abandoning hopes for a wholesale answer to the question of scientific realism.[12]

There may yet be other wholesale arguments waiting in the wings, and our arguments against underdetermination and conjunction *qua* wholesale arguments have been necessarily brief. Regardless, its worth imagining what sort of realism or anti-realism would survive if it were secured without the no-miracles argument or pessimistic induction. Without these, we lose the rationale for both *entity realism* and *structural realism,* two accounts that struggle to sharply divide theoretical structures from entities posited by the theory.

8.2. Retail Arguments. Nothing said here dooms retail arguments for or against realism. There may be good reasons to be a realist about neutrinos, an anti-realist about top quarks, and so on. Indeed, it would have been a

10. One of us has dealt with the question of underdetermination in considerable detail elsewhere (Magnus 2003a).

11. For one version of the argument, see Boyd 1982. There are objections to the conjunction argument, as in van Fraassen 1980, but it interests us here merely as a putatively wholesale argument.

12. In recent work, Philip Kitcher deploys an argument form he calls the *Galilean Strategy* to underwrite realism (2001a; 2001b, chap. 2). As one of us has argued elsewhere, however, the Galilean Strategy leads to realism about some particular kinds but fails to underwrite the inference from success to truth (Magnus 2003b); as we now put the point, the Galilean Strategy succeeds as a retail argument form but fails as a wholesale argument for realism.

surprise if base rates were neglected in all the arguments marshaled for and against the reality of particular entities or structures. Return to the case of our imagined disease, and ask if we should be realists about the disease even when it is incubating but yet to show symptoms. As noted above, we may sometimes know the base rate of the disease in a population by independent means. Moreover, we might be able to identify the disease-causing pathogen under a microscope. This would provide an argument for realism that is not statistical and, hence, does not implicate base rates in any obvious way.

We therefore have at least one clear way of distinguishing profitable realism debates from unprofitable ones. Profitable ones will either not be sweeping statistical arguments or, if they are statistical arguments, will be such that the base rates can be determined through independent means. Consider some classic examples. The debate over the reality of atoms at the turn of the twentieth century was one of huge significance in science. This was not a general statistical argument. No one used (say) 'most atomistic theories turn out to be true' as an argument for atomism. With argument and experiment it was thus possible for the most forceful opponent of atomism, Planck, to change his mind and be committed to atomism. With the debate among caloric theory, the wave theory of heat, and the dynamical theory it was the same. The arguments were non-statistical. Here again it was possible for a notable defender of caloric, Thompson, to completely change his mind due to theoretical and experimental arguments.[13]

Of course, there are also statistical arguments made in realism debates, especially in debates in high energy particle physics. Base rates may well be used. In any retail case, however, there will in general be independent handles on determining these base rates. Consider for instance the discovery of the $W\pm$ particle, the charged intermediate vector boson needed by the Weinberg-Salam theory of electroweak interactions. Observed events matched what was predicted by theory: the mass, type of momentum, and lack of particle jets. But how do we know that these events were not background effects pretending to look like these candidate effects? It is possible that this happened, and one might worry that if the base rate of such mimicking events is high then we can't state with confidence that we have likely found the $W\pm$ particle. The key here is to understand that there are independent theories of many other particles, collisions, etc., each of these independently tested, that in fact provide

13. Within some theories of confirmation (e.g. Bayesianism), a retail inference from confirming evidence to theory is treated in terms of probabilities—and so problems of base rates *might* arise. Wholesale arguments need to address populations of theories, however, so they invite statistical considerations—and so problems with base rates *do* arise.

independent predictions of the probabilities of such background effects (see Franklin 1986). Here is a case where we do need to know various base rates to corroborate our theory, but these base rates are obtainable. In general, science proceeds with a variety of methods at independently estimating crucial base rates, e.g., using a variety of instruments, experimental checks and calibration, eliminating alternative explanations, etc.

One may object that these independent theories themselves require base rates, the theories corroborating these base rates require base rates, and so on. In other words, maybe there is no way of breaking out of the circle involving base rates. We see no particular reason to believe that this is the case. If it is and the circle is wide enough, however, then this claim begins to sound like merely a restatement of the problem of general scepticism—not something we are concerned to tackle here.

One may also object:[14] 'Retail arguments rely on general assumptions that can be used to construct a wholesale argument. You agree that in the early twentieth century it was appropriate to believe in the existence of atoms for all the usual reasons. Let *F* be whatever features of the situation at the time made it reasonable to believe in atoms. There is then a wholesale argument that, whenever *F* obtains, one should believe in (say) the principle theoretical posits and entities of the relevant science.'

We believe this objection would have us reprise a standard debate in epistemology, namely, the problem of the criterion (Chisholm 2003). How do we know which beliefs are justified? Do we have a criterion, like Descartes' clarity and distinctness, and then see which particular beliefs satisfy this criterion? Or do we begin with particular beliefs we know to be justified, like Moore, and then generalize from them? Retail arguments turn on particularism in the present context. We acknowledge that it may be possible to get a kind of wholesale argument by discovering something in common among all good retail arguments for realism. Without trying to settle the larger epistemological issue, we offer a note of caution. Reflecting on the vast complexities of various historical episodes in science, there is no reason to think that the general assumptions one finds will be at all simple, natural, or even non-disjunctive; in short, there is no guarantee that the criterion one finds will be either interesting or useful. So although it is logically possible to turn a retail argument into a kind of wholesale argument, the resulting wholesale argument may appeal to 'general assumptions' that are long, gruesome, and can do none of the heavy lifting that wholesale arguments are usually meant to do.

Perhaps philosophers inclined to dissolve debates over realism will be no happier with retail arguments than they were with wholesale arguments. Fine, for instance, rejects what he calls *piecemeal realism*. He decries

14. We owe this interesting objection to an anonymous referee.

realists' attempts "to relocate the school to where conditions seem optimal for its defense, and then to insinuate that the case for such a 'piecemeal realism' could be made elsewhere too, were there but world enough and time" (1991, 79). Fine describes a plausible stratagem for a realist who sees the value of retail arguments but has a hidden yen for wholesale conclusions. Yet licensing the move from a collection of realist cases to conclusions about all or most of science requires an implicit statistical argument that the cases are representative. Such statistical arguments are doomed.

What Fine calls piecemeal realism is thus only an ersatz retail argument; the particular case is offered as a proxy for all of science. As Fine sees it, the piecemeal strategy is motivated by the opposition between the no-miracles argument and the pessimistic induction (1991, 81–84). We have argued that these opposing wholesale arguments should be set aside. We should pay attention to particular cases for their own sake and not as proxies for something else.[15]

9. Conclusion. We can now answer the question with which we began. The question of why scientific methods succeed is, in a sense, just as legitimate as the question of why silver fishing lures succeed in catching fish. There is nothing wrong with the philosopher wanting an answer to the broader question. Because one can imagine a full continuum of cases where base rates are knowable to a varying degree, one can see that the distinction between wholesale and retail is a matter of degree and not kind. We concede that there is no sharp distinction between the wholesale and the retail. Nevertheless, good retail arguments for realism are on one side of the spectrum and wholesale arguments occupy an extreme position on the other side. The lunge for totality in wholesale arguments suggests that they will need statistical considerations about all or most theories and also that there will not be any independent methods for estimating the relevant base rate. Without independent methods for estimating crucial base rates, there is little to do but make arguments that beg the question. Wholesale realism debates persist not due to mere stubbornness, but because there is *no reason* for opponents to agree. The more modest reach of the narrower retail question allows for arguments that are non-statistical or for broad agreement in estimating base rates. These debates are profitable because there is reason to agree.

We suggest that the great hope for realism and anti-realism lies in retail arguments that attend to the details of particular cases. It is unlikely that

15. As a matter of terminology, there is some small reason to prefer 'retail' to 'piecemeal.' Retail arguments naturally contrast with wholesale arguments, but piecemeal realism has no such natural contrast. (Whole hog realism?)

either side will win *every* argument; it seems more likely that realism and anti-realism are options to be exercised sometimes here and sometimes there.[16] This equivocal victory for each might be uncomfortable for realists and anti-realists alike, but so be it.

REFERENCES

Blackburn, Simon (2002), "Realism: Deconstructing the Debate", *Ratio* 15: 111–133.

Boyd, Richard [N.] (1982), "Scientific Realism and Naturalistic Epistemology", in Peter D. Asquith and Ronald Giere (eds.), *PSA 1980,* vol. 2: 613–662.

Callender, Craig, and Nick Huggett, (eds.) (2001), *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity.* Cambridge: Cambridge University Press.

Casscells, Ward, Arno Schoenberger, and Thomas B. Graboys (1978), "Interpretation by Physicians of Clinical Laboratory Results", *The New England Journal of Medicine* 299: 999–1001.

Chisholm, Roderick (2003), "The Problem of the Criterion", in Louis Pojman (ed.), *The Theory of Knowledge: Classical and Contemporary Readings,* 3d ed. Belmont: Wadsworth, 2003, 9–17.

Cordero, Alberto (2001), "Realism and Underdetermination: Some Clues From the Practices-Up", *Philosophy of Science* 68 (Proceedings): S301–S312.

Dorling, Jon (1992), "Bayesian Conditionalization Resolves Positivist/Realist Disputes", *The Journal of Philosophy* 89: 362–382.

Fine, Arthur (1984), "The Natural Ontological Attitude", in Jarrett Leplin (ed.), *Scientific Realism.* Berkeley: University of California Press, 83–107.

——— (1991), "Piecemeal Realism", *Philosophical Studies* 61: 79–96.

Franklin, Allen (1986), *The Neglect of Experiment.* Cambridge: Cambridge University Press.

Howson, Colin (2000), *Hume's Problem: Induction and the Justification of Belief.* Oxford: Clarendon Press.

Kitcher, Philip (2001a), "Real Realism: The Galilean Strategy", *The Philosophical Review* 110: 151–197.

——— (2001b), *Science, Truth, and Democracy,* Oxford: Oxford University Press.

Korb, Kevin B. (1991), "Discussion: Explaining Science", *The British Journal for the Philosophy of Science* 42: 239–253.

Kukla, André (1998), *Studies in Scientific Realism.* Oxford: Oxford University Press.

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.

Lipton, Peter (1994), "Truth, Existence, and the Best Explanation", in Anthony A. Derksen (ed.), *The Scientific Realism of Rom Harré.* Tilburg: Tilburg University Press, 89–111.

Maddy, Penelope (2001), "Naturalism: Friends and Foes", in James E. Tomberlin (ed.), *Philosophical Perspectives* 15: 37–67.

Magnus P. D. (2003a), *Underdetermination and the Claims of Science.* Ph.D. Disseration. La Jolla CA: University of California, San Diego.

——— (2003b), "Success, Truth, and the Galilean Strategy", *The British Journal for the Philosophy of Science* 54: 465–474.

Meehl, Paul E., and Albert Rosen (1955), "Antecedent Probability and the Efficiency of Psychometric Signs, Patterns, or Cutting Scores", *Psychological Bulletin* 52 (1955): 194–216.

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

16. Kukla (1998, 28) suggests that science which is sometimes realist, sometimes anti-realist would be irrational, but we do not see why this should be the case. Tailoring one's opinions to the particular evidence at hand seems to us instead the paradigm of rationality.

Sober, Elliott (1990), "Contrastive Empiricism" in C. Wade Savage (ed.), *Scientific Theories.* Minnesota Studies in the Philosophy of Science, vol. 14. Minneapolis: University of Minnesota Press, 392–410.

Tversky, Amos, and Daniel Kahneman (1982), "Evidential Impact of Base Rates", in Daniel Kahneman, Paul Slovic, and Amos Tversky (eds.), *Judgement under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press, 153–160.

van Fraassen, Bas C. (1980), *The Scientific Image.* Oxford: Clarendon Press.

Worrall, John (1989), "Structural Realism: The Best of Both Worlds?", *Dialectica* 43: 99–124.