

Optimism about the Pessimistic Induction

Sherri Roush

Abstract

I argue that pessimistic inductions over the history of science have not made the case that the failures of our predecessors give us reason to dial down our confidence in our scientific theories. I explain why an effective pessimistic induction must do more than meets the eye; it must show that 1) the supposed unreliability of our predecessors is relevant to what we should think about our own reliability, and 2) if we believe that we are lacking reliability – a second-order property – then we are rationally obligated to withdraw confidence in our first-order beliefs. I assist the pessimist by providing an account of fallibility that explains why the second is true. However, I show that the fact that we use different methods than our predecessors breaks the pessimistic induction from their unreliability to a conclusion that we are unreliable. To threaten our rational confidence in our particular theories, the pessimist needs much more than he has offered.

Why worry?

How confident does the history of science allow us to be about our current well-tested scientific theories, and why? The scientific realist thinks we are well within our rights to believe our best-tested theories, or some aspects of them, are approximately true.¹ Ambitious arguments have been made to this effect, such as that over historical time our scientific theories are converging to the truth, that the retention of concepts and claims is evidence for this, and that there can be no other serious explanation of the practical success of science than that its theories are approximately true. There is appeal in each of these ideas, but making such strong claims has tended to be hazardous, leaving us open to charges that many typical episodes in the history of science just do not fit the model. (See, e.g., Laudan 1981.) Arguing for a realist attitude via general claims -- properties ascribed to sets of theories, trends we see in progressions of theories, and claimed links between general properties like success and truth that apply or fail to apply to any theory regardless of its content – is like arguing for or via a theory of science, which brings with it the obligation to defend *that* theory. I think a realist attitude can be maintained rationally and without such a theory, as I will try to explain.

The starting point at which questions arise as to what in general we have a right to believe about our theories is one where we have theories and evidence for them, and we are involved in the natural or unreflective activity of apportioning our belief in each in

¹ Recent efforts of this sort can be found in Leplin 1997, who argues for a link between novel predictive success and truth, and Psillos (1999), who argues for a link between referential continuity and descriptive accuracy.

accord with the strength of the evidence.² The devil's advocate sees our innocence and proceeds to do his best to sow seeds of doubt as best he can. If our starting point is as I say, though, the innocent believer in particular theories does not have to play offense and propose sweeping views about science in general, but only to respond to the skeptic's challenges. The burden of initial argument is on the skeptic. The greatest strength of the realist attitude lies at this starting point, as Arthur Fine (1996) realized, and I will argue here that the innocent has so far not been given reason to give it up. In particular, no pessimistic induction over the history of science has done what it needs to do in order to undermine our right to apportion our beliefs in particular theories to the evidence we have for them. To show this I am going to reconstruct what the pessimistic inductivist could possibly be saying when he tries to undermine our confidence, and show that there is an argument form that is potentially threatening, and that the innocent realist has no right to ignore. However, this argument form requires much more than anyone has realized if it is to serve the pessimist, and the pessimist has not actually ever offered what is needed.

For an example of how to start from the beginning, consider the realist of the sort I described, whom I will also call the "optimist," who is typically confident that Quantum Theory has gotten some very big things right. Her confidence in Quantum Theory is based on trust in the work of the scientific community, her understanding of the good evidence we have for this theory, and her own thought and knowledge about the physical world and logical and mathematical matters. There may be circumstances in which it would be helpful to her to have a further argument at her disposal that explains why all of this speaks to her having a right to her belief about Quantum Theory. But the having of a right to a belief does not require her to have an argument available that she has a right to this belief. If it did then the justificational regress that follows would undermine our right to any of our beliefs.

If this is where we start then the anti-realist will have to provide reason to believe that our innocent does *not* have a right to such beliefs. One way to do that would be to cast doubt on the optimist's evidence for Quantum Theory, meaning arguing about particular experiments and so on, and if those doubts were compelling then she would be obligated to take them seriously. Of course, this has not been the strategy of the anti-realist, whose only other option, then, is to make a general argument for doubt that will apply to Quantum Theory qua theory, or qua theory of a particular sort. That is, the pessimist will have to find a fault or weakness that can be described without reference to the contents of particular theories. If he is to do this by an induction over history, then he must draw our realist into reflection about all (many, most, typical, etc.) theories and scientists of the past, and show why those reflections are both relevant and damning for us. The pessimist's tools have to be at a general level, but the confidence of the optimist that the pessimist would like to undermine is in particular theories, e.g. Quantum Theory, so the pessimist needs to give us something if he is to provide a connection. If instead of history he is going to invoke general concepts, such as underdetermination of theory by

² This stance is of course somewhat similar to Arthur Fine's Natural Ontological Attitude. (Fine 1996) However, besides my focusing more on epistemology rather than ontology, there is also the difference that here it is a starting point against which, I argue, there is a compelling form of argument for ascent from our object-level beliefs to reflection on them, and from there descent to withdrawal of confidence in particular theories. It's just that, as I show below, making that form of argument successful requires much more than the anti-realists have offered.

evidence, or general trends in the history of science, like getting it wrong by our lights, then the anti-realist will have to explain why these things have an upshot for belief in Quantum Theory. I don't say these things cannot be done – sometimes a theory exhibits a simple instance of one of these epistemologically relevant concepts, and surely there are some similarities between our predecessors and ourselves. I am making the point that this is what needs to be done in order to provide a challenge to optimism about the approximate truth of any particular theory. You cannot do a legitimate induction from properties of swans to properties of bananas.

In the pessimistic induction over the history of science the relevance of the failures of our predecessors is supposed to come by way of likeness. We are like them, and since they failed at the project of getting true theories, we have insufficient reason to believe that we are succeeding. I will argue that there is only one form of the pessimistic induction over the history of science that is a *prima facie* compelling challenge to the hunkered down realist position just described. However, as I explain below, it takes much more care than has been recognized to make that argument succeed.

In particular, I will argue that for the pessimist's historical evidence base of failure to be relevant to us at all, his argument must be formulated as a meta-induction, meaning an induction over our beliefs in scientific theories and the properties of our beliefs, where these take the position that swans and white respectively do in the generalization to "All swans are white." Similarity at the level of these meta-facts – they had beliefs ... , we have beliefs ... – could be the basis of the induction the pessimist desires, but simply because of the way such an induction must work, two more elements are also needed: to know why exactly facts about their beliefs are relevant to claims about our beliefs – the contents of the beliefs are different after all – and to know why even a meta-claim about ourselves, a claim about our having certain beliefs and dispositions to believe, should affect our confidence in the matters they are beliefs about. We will see that the pessimist must make some quite substantive claims to make all of this work, and that so far no pessimist has adequately made the case. It is not clear to me that they can.

The most obvious first pass at rendering the pessimistic induction is simply to say: our predecessor scientists believed theories and those theories were false. We believe theories, too. What is the reason to think they are not false? The realist might reply at this point that our evidence is better than our predecessors', but, the pessimist will say, though you have better evidence than your predecessors by your lights, they had better evidence than their predecessors too, by their lights. Still they were wrong. However, drawing the conclusion from this that we are probably wrong presumes either that objectively better evidence cannot advance us toward the truth no matter how much better it is, or that there is no way for us to judge better and worse evidence objectively, that our view about this is merely our view. To presume that even objectively better evidence cannot get us conclusions that are more likely to be true is to cite the general problem of induction, and if that is the pessimist's basis then he did not need to cite past scientific failure. The second issue of whether we have objectively better evidence, and how we could know that, is one I will address below.

But the optimist needn't have gone so far as to insist that our evidence is better. For though there may be reason to think our theories are false, the fact that our predecessors were wrong is not a reason to think our theories are false unless our case is

relevantly similar to theirs. If we cite further facts that relevantly distinguish us, then it is the pessimist's burden to say why we are for all relevant purposes the same. This is just an instance of a general fact about induction, namely that even if all F so far are G, the inference to the claim that the next case with F will be G is rendered illegitimate if the next case is known to have a further property which we know is relevant to whether G arises. There is such a further fact in this case: our evidence set is simply distinct in content from that of the evidence our predecessors', and the content of evidence claims is obviously relevant to whether they support a given theory. If the pessimist simply cites the fact that our predecessors had (good) evidence for their beliefs, then there is a counter: the content of a set of evidence propositions is relevant to whether the belief one might base on them is true. If our evidence set is different from our predecessors', then it remains to be shown why their failure is a reason to believe in ours. The optimist can stop here, for it is not now her burden to say why ours is better, but the pessimist's burden to say why it is not.

The pessimist might instead try to find the relevance of our predecessors' evidence to our situation by pointing to the specific evidential claims; our evidence *set* is different from theirs, but they overlap over some propositions. This will not work, because even particular evidence claims that we may share with our predecessors can be relevant to whether our theories are true only if their content is relevant to the content of our theories in the right way. Typically it is not. Whatever it was about the facts and the phlogiston theory that made scientists give up that theory is not evidence against our theory or the parts of that chemical picture that we might have retained. Our theory did not retain parts of that theory that we (and the pessimist) thought that evidence had falsified, so the evidence that falsified that theory does not falsify our chemical theory.³

The relevant evidence of our predecessors was supposed to be the evidence they had *for* their theories, one will say. But this fares even worse. Anything we did retain from their theory had not been falsified by any of our predecessors' evidence (or ours), so any positive evidence they had for their view is also positive evidence for that part of ours. Evidence, always being incomplete, may support more than one theory or falsify more than one theory. Thus the positive evidence for their theory may be relevant to our theory even if we did not retain their picture. But if that evidence supported our theory, how exactly could it give a reason for us to withdraw confidence in our theory? It is not, after all, our *only* evidence. And if it told against our theory we would not be espousing our theory, since this is evidence that we knew about in choosing our own theories.⁴

³ Obviously, I am using a highly idealized picture of how later theories deal with earlier evidence for earlier theories (broadly, denying Kuhn loss, and assuming what Feyerabend called the Consistency Condition), but this does not presuppose convergence of theories to the truth, because it does not presuppose that there is any amount of evidence we could get that is good enough to give us a right to say we are closer to the truth than our predecessors. Respect for the evidence of our predecessors of the kind I'm supposing here is compatible with our both being not only quite wrong, but *prima facie* justified in our belief. Stanford's pessimistic induction, in particular, does not depend on Kuhn loss. See Ch. 1, Stanford (2006).

⁴ I am making crude assumptions about falsification. But the pessimist must also be assuming the past theories are falsified or his negative evidence from the history of science disappears. And he must be assuming that our theories are not falsified or he would not need to make his argument in the first place. One can be more subtle about how much negative evidence it takes to be falsified. One counterexample is never taken to be enough in real science. But the structure I have described remains, and follows from the symmetries and asymmetries the pessimist needs as a basis.

This point is even stronger when we consider that the pessimist is trying to make an argument over all theories from the past, and his argument sounds better for it. Somehow the track record of all theories together speaks against our particular confidence in Quantum Theory. However, that general track record over all such theories is not relevant to our confidence in each of ours via our predecessors' particular evidence claims, since there is no reason at all to think that the particular evidence against or for the theory of bodily humors, or spontaneous generation, etc., that eventually came to light, casts doubt on Quantum Theory. None of these cases support the idea that the particular evidence that falsified or supported old theories says something negative or positive about any particular theory of ours.⁵ When put this way the point is obvious, and it may be for that reason that it has not tended to be remarked upon; we need to draw it out here in order to scan all of the pessimist's possible options.

The evidence for or against false theories of the past is not relevant to whether we have a right to believe our theories insofar as the contents of those evidence-claims are not related to the contents of our theories in an appropriate way. I drew this out in detail to show that the pessimist is forced to use a meta-induction. For what does strongly appear to be relevant to our case is the fact that our predecessors had evidence that supported their theories, and that evidence was inadequate to their presumed goal of having true theories. The similarity for a potential induction must be between the point of view of many previous scientists and the point of view of our own, between how their evidence (or beliefs) related to their theories' truth values and how our evidence (or beliefs) relate to our own theories' truth values. This general feature is the only thing we share with our predecessors that is potentially relevant in a pessimistic way, to our right to believe in Quantum Theory. The similarity in cases that makes for the basis of a pessimistic induction is indeed the property of their being in some sense justified by their evidence, and our being in some sense justified by ours.

This is where the power of the pessimistic induction over the history of science lies. As noted, I call it a "meta-induction" because its claims refer not to things in the world that scientific theories are about but to our *beliefs* and evidence about things, and our tendencies to believe. In the pessimistic argument are not simply acting in our epistemic situation – believing theories on the basis of evidence – but reflecting on that activity. The pessimist bases his conclusion on the observation that people like us have been rational with their evidence yet come to false theoretical beliefs, and categories with no referents, time and again. Thus, what links us to our predecessors is not that what falsified or supported their theories T_1, T_2, \dots, T_n might falsify our theories $T_{n+1}, T_{n+2}, \dots, T_m$. The similarity that makes the basis for an induction is the more general claim that they had good evidence, in some sense of good, just as we do. The property that came along with this was being wrong. This is right as far as it goes, and standard pessimist have even sounded like this.

It is not often noticed that this is not enough to make the pessimistic argument go. One reason is that our evidence is in many ways different from that of our predecessors, beyond the difference in content just discussed, now in general properties. We have more.

⁵ It would be pointless for the anti-realist to object to the assertion that our theories are not falsified by appealing to the Duhemian point that this is so only relative to auxiliary assumptions we must be assuming we know. The claim that our predecessors' theories were false is also relative to auxiliary assumptions, and, as noted in the previous footnote, the pessimist needs this claim to get his induction going.

It covers more territory. It comes from different experiments. This difference is there even if both sets of evidence are good, in some sense, for their times. To do an induction we need a similarity there. What is it? Likewise, the property of our predecessors that “they were wrong” must be a property we can have too, if the pessimist is to infer that we have it. The attribution to our predecessors can’t simply be the statement “-T₁ and -T₂ -T_n” denying their particular theories, because that can’t be a property of our theories T_{n+1}, T_{n+2}, . . . , T_m, since the contents are not relevantly the same (see above) and the negated contents of our predecessors’ theories is the only content of the property described. If an inductive projection is to occur, there must be a property of our predecessors that we or our situation can be substituted into as argument, keeping the property itself the same. An example would be something general like:

For all n and all theories T_n that were believed by those human beings doing science, and unfalsified at the time, PR(-T_n) > .90,

that is, the probability that T_n is false is greater than 90%. If the induction goes through, then our property will be

For all n and all theories T_n believed by these human beings doing science, and unfalsified at the time, PR(-T_n) > .90.

The number .90 is not special; if there is an important challenge here at all, then any number approaching .5 will do. The point is that the pessimist’s argument needs a lot more structure, including higher-order structure, if it is to make the past record connect to us at all.

The property in question is a relation between our beliefs and evidence on the one hand and the truth of the theory believed on the other, namely the general property of fallibility. From there the argument must include at least two steps. The first is to show the *relevance* of the fallibility of our predecessors to our own level of fallibility. Most of us will readily admit that we, and even scientists, are fallible. The first claim in the pessimist’s induction though, must be that our predecessors’ fallibility is a *good reason* to believe in our own fallibility at a level as significant as theirs. This, I will argue, is questionable, and at least has not yet been shown by the pessimists. The second claim the pessimist needs is that *our* fallibility would be a reason to withdraw confidence in our particular theories. This claim is widely assumed but in serious need of explanation. I will argue that it is true and non-trivial that discoveries about our own fallibility can give us reason to withdraw confidence in our object-level beliefs, and I will supply the explanation of this below. Of course, as just said, the fact that this type of inference is legitimate does not get the pessimist his conclusion without an argument that our predecessors’ unreliability is a reason to believe in ours, which is where he has failed.

In what follows I will proceed in roughly three steps. First, I will formulate more explicitly what the steps of the second-order challenge must be which claims, on the basis of the history of science, that we have significant fallibility. Then I will explain the relevance of such a challenge to our first-order beliefs in our particular theories. Finally, I will use these points to show why the pessimist has not yet succeeded in making an argument of this form. Since the relevant property in all steps of the argument is the

general property of fallibility, formulating the challenge requires first defining this concept carefully.

Fallibility

To define the property of fallibility adequately, we must take into account at least two things. One is that the property isn't trivial; it both has bite, and, as I will argue, is something distinct from merely lacking confidence in particular propositions. Two, we must be able coherently to attribute it to ourselves and see what its proper relation to our other beliefs might be. In this section I will argue that thinking of fallibility as a second-order property, that refers not just to things but also to our beliefs about things, helps enormously with these goals.

There are already many areas where it would be helpful to have a finer formulation of the property of fallibility than we have. A classic example occurs in the Preface Paradox. I write a book, and thus assert each of its sentences. It is also natural for me to write a preface acknowledging that people are not perfect and that some of my conclusions may be, might be, or probably are, wrong. Both of these claims seem natural, and they are apparently intuitively coherent. Yet standard ways of expressing these two thoughts lead to contradiction. Let us use the following informal expression to see why. On the one hand, I assert each of the sentences of my book with high confidence:

$p_1, p_2, p_3, \dots, p_{10,000}$

On the other hand I admit my confidence that at least one of those thousands of claims is false:

not- p_1 or not- p_2 or not- p_3 or ... or not- $p_{10,000}$

Immediately we see that we have portrayed me as having high confidence in each of an inconsistent set of propositions.

Some have offered solutions that involve denying that conjoining conjuncts preserves justification (closure under conjunction). Others propose that the lesson of the preface paradox is that it is possible to rationally believe two contradictory claims – the conjunction and the disjunction – as long as we avoid conjoining them and believing a contradiction. These options have high price tags, and fortunately I do not think we need them. The significance of the difference between the list of conjuncts and their conjunction is that probabilistically speaking, a high confidence in p_1 , and high confidence in p_2 , and, ..., high confidence in $p_{10,000}$ do not rationally require me to have a high confidence in the conjunction $p_1 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_{10,000}$. In fact, as long as the confidence in each of the conjuncts is only high rather than maximal, I am rationally required to have quite the opposite, a low confidence in the conjunction. For consider, if my confidence in each of the p_n is 95%, I only need to get to 59 such claims in order for the required confidence in the conjunction to drop to 5%. This is because the probability of a conjunction is the product of the probabilities of the conjuncts. 95% raised all the

way to the 10,000th power, as in this case, makes the probability of the conjunction practically nil. It immediately follows that the probability of the negation of this conjunction, that is, of the disjunction of the negations of the conjuncts, must be high. My degrees of belief in the conjuncts, conjunction, and disjunction must be related according to these probabilities if I am to be coherent. Thus, as long as my confidence is not maximal, a high degree of confidence in the disjunction of the negation of these conjuncts is consistent with a high degree of confidence in each of the conjuncts. Indeed the former is required by the latter.

Resolution of the problem via non-extreme degrees of belief may not have seemed promising because on pain of incoherence one might not have expected the proportionality to work out right. Since her beliefs in each of the conjuncts collectively *contradict* the disjunction of the negations of the conjuncts the possibility that all have high credence may not have leapt to mind. But looking at degrees of belief showed that it is obligatory, and this could be seen as the silver lining in the phenomena that make us worry about the closure of justification over conjunction.

But on this picture the probability or rational credence in the conjunction is still low, one might object, whereas the preface should surely be confident of the conjunction. There is no reason to suppose the person writing the preface is confident in that conjunction. She came to have reasons to believe any of the conjuncts only by local arguments for each, we are imagining. She surely recognizes the logical relation between the conjuncts and conjunction, and the logical relation may make her feel she should assert the conjunction because of asserting the conjuncts, but her instincts about fallibility and the disjunction tell her not to. Discomfort comes from not recognizing why low confidence in the conjunction is rational given the rest of her commitments, but via non-maximal degrees of belief we have just seen the reason for this. For confirmation of this intuitive side of the situation notice that if we asked the one who came to confidence in each conjunct relatively separately whether she is confident that *all* of the claims in her book are correct (the conjunction), she will surely say: Well, no, probably at least one of them is wrong (the disjunction of negations).

One might object that how confident one is allowed to be in the disjunction of negations depends, on this picture, on the number of conjuncts asserted, and must be small if there are few conjuncts. Our preface-writer, on the other hand, is surely highly confident she has made at least one mistake. The allowed confidence in the disjunction does indeed depend on the number of conjuncts on this picture. However, intuitively, the preface's level of confidence that she's made at least one error also depends on how many claims she has made. If you had confidently made only three claims, how confident would you be that at least one of them is wrong? Surely not very. In that case it would be quite strange for the preface to express high confidence of error. She may have some credence that she has erred somewhere in the three, but the resolution of the problem via degrees of belief allows this. The reason that the preface paradox deals with a book isn't just that books have prefaces, but that books are long.

The other possibility is that my degrees of belief concerning these matters are maximal (extreme), 1 or 0. If my confidence in each of the conjuncts is full then probabilistic coherence requires my confidence in the conjunction to be perfect ($1^n = 1$), and also (thereby) requires zero confidence in the disjunction of negations of the

conjuncts.⁶ This is surely the picture we have when we sense that the person asserting her conjuncts and expressing her fallibility is incoherent. Coherence expects her to have zero confidence in the disjunction of negations, and she appears to have more than zero, apparently significantly more. However, this case is not paradoxical either. A person with maximal confidence would have no reason to give credence to those denials, since she does not leave open any possibility of being wrong. (The formal expression of this is the fact that the maximal degree of belief (1) that is the way we must express her confidence if we are to generate paradox, makes impossible any revision of that belief on the basis of change in other beliefs.) The person we imagine having perfect confidence that requires zero confidence in the negations is not plausibly a person who would have written confidently of the possibility of her being wrong in the first place.

The paradox of the preface does not arise merely from the fact that $p_1, p_2, p_3, \dots, p_{10,000}$ are together incompatible with $\text{not-}p_1$ or $\text{not-}p_2$ or $\text{not-}p_3$ or ... or $\text{not-}p_{10,000}$. We also need to have reason to suppose that those propositions (or other appropriately contradictory relatives of them) are ones that the writer of the imagined, apparently reasonable, preface is asserting. We just saw that we can suppose the writer of the preface has very high confidence in each of the conjuncts – as high as you like though not 1 – with significant credence for the disjunction of their negations, and no incoherence. And though there is a type of subject who should not give any credence to the negations, for coherence reasons this can only be a subject who is maximally confident in the conjuncts. Such a person has no doubt about those conjuncts, from any source, so we would not expect her to have written the self-doubting preface anyway.

The application of this discussion to the issue of realism is that we have just shown it is possible to acknowledge a high probability of error somewhere in your beliefs, maintain high confidence in particular beliefs, and remain perfectly rational. Concerning ourselves with maximal degrees of belief had led us astray here.⁷ So, we can have high confidence that some of our theoretical claims are wrong, while coherently maintaining high confidence in each. Acknowledging that we err, does not require pessimism. This is another expression of the fact that the pessimist cannot win at the first order but must deal with claims *about* beliefs and theories.

There is a richer way of representing the claim of fallibility which gives the pessimist another chance. It is of course not impossible for a maximally confident person to write the self-doubting preface. If we express fallibility in a more adequate way than we have so far, we will see that it is not even first-order probabilistically incoherent for the maximally confident person to write that preface. There is a broader irrationality involved in it though, that corresponds to our expectation that a person would not feel inclined to do that. The formulation of fallibility above implicitly took my expression of my fallibility about the topic of my book as itself in direct competition with my confidence in each of the propositions in question, equating my recognition of my

⁶ Thus, a perfect level of justification is closed under conjunction. What is not closed under conjunction is rational *degree* of belief. The failure to distinguish these two facts and the intuitions that correspond to them, causes much confusion about these kinds of cases.

⁷ Philip Kitcher (2001a, 2001b) first invoked the preface paradox and the fact that the book-author's assertions are intuitively coherent to argue that fallibility of science is not a reason to doubt our theories. I am effectively fleshing out that point by explaining why the book-author's assertions are coherent. However, my arguments also show why this coherence does not mean the scientific realist is automatically off the hook, since we can formulate fallibility in another way.

fallibility with a lack of perfect confidence that all of those propositions were true. This is one way to understand the problem, and it generates the most direct version of it that we just dealt with. However, what I am asserting when I say with high (not maximal) confidence:

not- p_1 or not- p_2 or not- p_3 or ... or not- $p_{10,000}$

is not giving me or my audience any more information over my lack of perfect confidence that:

p_1 and p_2 and p_3 and, ..., and $p_{10,000}$

(except the fact that with regard to these degrees of belief I am coherent and responsive to the relation between the contents of these propositions). My high, or at least non-zero, confidence in the disjunction of the negations of the conjuncts is not a statement of my fallibility. It is simply an expression of my confidence about particular matters.⁸

When one discovers that he is sometimes wrong about such things, what he discovers is unfortunate, his discovering it is fortunate, but neither is just the same thing as his lacking confidence in the particular propositions. Rather, it says he has a general property, a tendency among his beliefs; it is a statement *about* his beliefs, a statement which would include among its terms not only “ p ” but also “ $B(p)$,” meaning the subject believes b ; as such, a statement of my own fallibility is second-order, expressing a belief about my beliefs. This way of thinking about fallibility avoids the preface paradox too, in that a second-order claim holds no threat of failing to cohere in the strict sense with first-order claims like the p ’s above (not without further assumptions).⁹ However, intuitively, believing ourselves fallible surely places some kind of rational demands on us in our management of our first-order beliefs. It is a general issue exactly what those demands are, and it is a question the pessimistic inducer needs an answer to if his arguments that we are not generally reliable are to have any punch for our beliefs in Quantum Theory. Here I will do the pessimist the favor of filling in this blank in his argument by defining fallibility as a second-order property, and explaining why and how it places rationality constraints on first-order beliefs.

If I am asserting a claim of my fallibility in the preface, I am not expressing a low confidence in the collective set of claims – although it seems I should have that too – but expressing something *about* my beliefs. I am not withdrawing confidence in them but

⁸ This is also one problem that afflicts Andy Egan and Adam Elga’s (2007) argument that one can’t coherently believe one is unreliable without withdrawing the beliefs in question. There fallibility of a given degree has simply been formulated as with the original Preface Paradox here, as the disjunction of the negations of the propositions I believe, which automatically cancels that same degree of confidence in the claims in question. The 2nd-order formulation of fallibility I give and the solution I just provide two independent resolutions of their problem.

⁹ In particular, with the second-order way of expressing fallibility it will turn out to be conceptually possible to have both maximal confidence in p and recognition of one’s fallibility. This may seem paradoxical given my claim above that we wouldn’t expect someone to have these two states together. But though on my view it turns out to be conceptually possible, and first-order coherent, this state is not rational – you should revise your confidence when you recognize your reliability does not match it – so the preface writer we regard as making sense is not best imagined as in this mixed state.

making an observation about the probability that creatures like me, using methods like this, get it right about matters like this. The claim that p and the claim that I or those like me have a general tendency to imperfect beliefs on p-like matters, are distinct kinds of assertion.

It is reasonable to think that fallibility is a second-order property, and it seems clear to me that it is an inverse of reliability, which is also a second-order property. We can say that a person's belief-forming process is *x% reliable* when:

$$PR(q/B(q)) = x^{10}$$

That is, the (objective) probability that it is true, given that the subject has full belief in it is *x*, where for simplicity we are taking "B(p)" to mean perfect confidence.¹¹ For the moment we may think of this probability as a rate of success, representing how *often* p is true when the subject believes it. Obviously, the probability that a person is right about q when she has perfect confidence in it is not necessarily 1; confidence does not imply reliability.

From this definition of reliability, fallibility can then be expressed as the size of the gap between her reliability and perfect reliability. Perfect reliability means the rate, PR(q/B(q)), equals 1. Always, when she is sure of it, it is true. Thus, she is fallible (in her full beliefs) to degree *y* when:

$$1 - PR(q/B(q)) = y$$

To acknowledge one's fallibility is to acknowledge that $y > 0$, and that $PR(q/B(q)) = 1 - y$. That is, it requires a degree of belief in a statement about the objective conditional probability PR(q/B(q)). But without further assumptions this belief about this conditional probability does not constrain one's belief about q in any way. This is the most precise way of saying why there is no preface paradox. When we say we might be wrong somewhere, we mean this second-order claim which does not immediately conflict with our first-order beliefs. To make a connection we would need to make an assumption about how a degree of belief about a property of our degree of belief in q should relate to degrees of belief in q. Probabilistic coherence alone – conforming one's degrees of belief to the probability calculus, does not by itself give a constraint concerning that connection.

To admit that you are fallible about q is not *eo ipso* to withdraw belief in q. Similarly for partial belief, to admit that your reliability level about q when believing it to some specific degree is *x* is not *eo ipso* to reduce or increase your degree of belief in q. Admitting your fallibility may be a reason to reduce your degree of belief in q, but it is not the same thing, and if it is a reason we will have to say why.

Non Sequitur

¹⁰ This is not the only way to express reliability, and hence fallibility, though I think that any adequate way of expressing it must be second-order, not first-order.

¹¹ One might be worried about the lower end of the spectrum of degrees between 0 and 1. What could it mean to say that someone was 20% reliable? It means, simply, that they get it right 20% of the time. They wouldn't be a very useful source (except in reverse, in a yes-no query), but that isn't the problem here.

We now see how to express admission of fallibility in a way that is *consistent* with first-order assertion of particular claims, so that the admission is not *eo ipso* a withdrawal of confidence. But is it not clear that the admission still undermines our justification for the first-order belief by an obvious and immediate inference? Wouldn't our serious fallibility as scientists, if shown, automatically give reason for an equally serious reduction of our confidence in Quantum Theory?

Consider one consequence if we did think that the admission of fallibility immediately undermines our justifications for first-order beliefs, even domain-specifically. Most of us will have been frustrated at least once in our lives by a typical Creationist (or Intelligent-Designist) line of rhetoric one hears, which points out that the theory of evolution has not been *proved*. Therefore, it is concluded, opposing views (especially one in particular) have a right to be taught in public schools. We should have an "open mind," and not be so confident in the views of the scientific mainstream that we shove them down students' throats. It certainly looks like the Creationist is appealing to the fallibility of science in this argument. One obvious reply to this is that he is mistaking a small quantity of fallibility for a problem so large that it puts all views and frameworks on an equal footing of plausibility. "Proof" is for mathematics and logic, not for empirical research where fallibility is just the way it is. But if the level of fallibility is significant, then what? The resemblance of the form of the pessimistic induction over the history of science to this Creationist rhetoric is somewhat alarming. In charity we must suppose there is a difference, but it would be nice to know what it is. We will see below that when we understand what has potential in the pessimist's argument – by understanding the rational place of fallibility claims – we will also see what is fallacious about the Creationist argument.

The first problem is a group of quantitative issues. Judging intuitively, the admission that we are not perfectly reliable should not produce a leveling in the credibility of all theories, though we have yet to say exactly why. The pessimistic argument is most compelling if we convince ourselves that our predecessors were often or always wrong, meaning they had a low, or very low, reliability. But how should we count? What if there are more theories in the last 100 years than in the previous 5,000, and whereas all of the previous ones were false (or at least abandoned) a high percentage of those in the last 100 years are still not falsified (and are retained)? All of these folks are our predecessors and the overall numbers aren't yet shown to be bad. Nevertheless, suppose we can argue for a high unreliability of our predecessors. It's clear we share with them a lack of perfect reliability, but why think we share with them a quantitatively large impairment?

I will return to this question below. But suppose we can find a reason to think that large impairment is shared. Another pressing problem, just described above, is why even if we successfully attribute a surprising, significant, general level of unreliability to ourselves, this should issue in a withdrawal of our confidence in specific first-order matters, such as particular claims of the Special Theory of Relativity. Suppose that $q =$ The speed of light is not different in different frames of reference, where frame of reference is defined by velocity. The claims that are probabilistically relevant to that claim are the results of the Michelson-Morley experiment, the claim that the interferometer arms were equal in length, and so on. The first is relevant because if the

experimental apparatus works as it should then the fact that no interference fringes show up in the experiment raises the probability that the speed of light is constant over reference frames. The second is relevant because length equality in the interferometer arms increases the probability that the apparatus works as it should, i.e., that fringes you may see are due to variation in the speed of light. Length equality is relevant to the hypothesis because it raises the probability of that hypothesis given the experimental outcome. Our being fallible is definitely not relevant to the Special Theory of Relativity in the way that typical evidence like this is.¹²

As pointed out above, it is a straightforward fact that the general level of unreliability we might attribute to ourselves,

$$PR(q/B(q)) = x,$$

is not probabilistically relevant to

q

and thus not to how much confidence we should have in q, without further assumptions. $PR(q/B(q))$ is related to $PR(B(q))$ via:

$$PR(q/B(q)) = PR(q \cdot B(q)) / PR(B(q))$$

But the question is what confidence we should have in q, given our credence for $PR(q/B(q)) = x$, not what confidence we should have in B(q) in that circumstance. How confident should I be in q if I believe that the probability that q is true given that I believe it is x? How my degrees of belief about these matters should affect my degree of belief in q is so far an open question.

Failure to respond to a change in my confidence in my reliability level $PR(q/B(q))$ with a change in confidence in q does not make me (1st-order) probabilistically incoherent. It is the burden of the argument the pessimist needs to be making to say what the extra assumptions or rationality constraints are that make the fallibility claim relevant to our first-order beliefs. Otherwise the track record of our predecessors has not given us a reason to withdraw confidence in our theories, even if it gives us reason to doubt our reliability, even significantly.

Descent

I will call this the problem of “descent” because we have to say why and how a second-order discovery about our tendency to believe should affect our confidence in specific

¹² One might notice that the second claim is the denial of a Type II defeater—a claim not that the purported evidence claims are false but that the purported support relation between the evidence and the hypothesis is weaker than thought. And one may wonder whether we don’t need something greater than first-order probabilistic relevance to model that relevance. But first-order relevance accounts for Type II defeaters handily. C is a Type II defeater of the claim that A supports B iff $Pr(B/A) > Pr(B/A.C)$. A Type II defeater *screens off* the relevance of A to B. The relevance of all the specific claims to the hypothesis that the speed of light is constant over reference frames can be represented as first-order probabilistic coherence. The relevance of claims about fallibility cannot.

first-order claims. Surely a discovery that, say, we fall into a class of people who have a 20% error rate on theoretical claims doesn't mean we need to drop our confidence in the relativity of simultaneity to 50-50, meaning that all bets are off. But it is reasonable to think it means *something*.

A notion that would give the pessimist's intentions their due is that of calibration: we are calibrated when we match our confidence to our accuracy. We can see an example of the concept of calibration in the way that an eyewitness's confidence on the witness stand relates or should relate to his other properties. We onlookers may have no information about the person's track record in making face recognition judgments, but there is a track record and it appears to be epistemically good, in some sense, if the witness is no more confident in his judgment now than the percentage of times he's gotten such judgments right in the past. His track record is a clue to how reliable he is. If he's been right 99% of the time it would be epistemically good if he were very highly confident. (Otherwise in trusting him only to his level of confidence we are deprived of the information he likely has.) At the other end of the spectrum, if he were right 50% of the time in past such cases, we hope that he would say "I don't know." He may have no inkling of his bad track record—he may have acquired amnesia since his last such judgment—and if we don't know about the amnesia or the track record we are out of luck. It is clear what would be most epistemically sound and helpful in these situations, from imagining it present and imagining it absent.

There is psychological evidence that jurors use confidence of a witness as an indication of accuracy by default, effectively assuming calibration without any information pertinent to the property. That is, on first encounter with a witness, a juror tends to believe what a confident witness says in direct proportion to the witness's confidence. Since confidence implies nothing without calibration, this appears to be an epistemic disaster for the system of trial by jury, and was greeted as such on its discovery. This illustrates both that calibration is a non-trivial property, i.e., that confidence and accuracy are distinct properties that do not necessarily line up in naturally occurring beliefs, and that it matters to us; it would be good for us if a witness's testimony were calibrated before we form any particular degree of confidence in what he says, even if, as a matter of fact, we tend to be poor judges of calibration. Incidentally, and in keeping with the idea that we do demand calibration, it has recently been discovered that (mock) jurors do make use of any further evidence of error by that witness that comes available, to reduce their judgment of his calibration and thereby of his credibility when he is confident. (Tenney et al. 2007) The jurors' default assumption of calibration is defeasible.

So, suppose calibration – having your confidence match your accuracy – is a good thing, and that by "accuracy" we mean reliability, probability of being right, the property defined above that we can get information about by looking at track record. This would imply that our confidence in our particular theories should match our reliability in making such judgments, and that if we aspire to that then our confidence should match what we *believe* our reliability is.¹³ We have hereby uncovered the basis of an intuition

¹³ This obligation can be formulated as a rule for updating our beliefs on discovery of information about our reliability level, so: $Pr_i(q) = Pr_i(q/[Pr_i(q) = x \cdot PR(q/Pr_i(q) = x) = y]) = y$. That is, the final degree of belief you ought to have in q given that your current degree of belief in q is x and that you think the objective probability of q when you believe q to degree x is y , is y . So, for example, if you know that in the past

that makes us feel a pull from the pessimistic induction: if the pessimist can show us that the (supposed) low reliability of our predecessors is good evidence that *we* have a low reliability, then we should reduce our confidence in our particular theories to the level of reliability we have now come to believe we have. If we believe our reliability level as scientists is 50%, then we should be no more confident than not that our theories are true. Our beliefs about our own reliability/fallibility in p-like matters *are* after all relevant to what our confidence in p should be, in both a coherent and a significant way. Thus, we can see how the pessimist can finish his argument on us, if he manages to convince us we are unreliable. But this is the only good news for the pessimist. In what follows I will discuss how the historical evidence the pessimist gives us *could* be a reason to believe we are unreliable, but argue that it is not.

Note that this way of viewing the relevance of the second-order property of fallibility to what our first-order confidence in particular theories should be does not license the Creationist move. This is because the rule I have described does not imply that fallibility implies equal plausibility of all theories. The fallibility will, speaking in idealization, be a given number, the reliability being 1 minus that number. Suppose, for illustration, that the fallibility level of contemporary science is 20%; then the reliability level is 80%. The rule says match your confidence to your reliability, so in this case our confidence should be 80%. Discoveries about your reliability level in q-like matters may obligate you to push your confidence in q up or down by a small or large amount, but they do not automatically mandate, or even license, a leveling in which every view is as good as every other. This view also shows why the amount of fallibility matters – our allowable confidence varies with it in a way that conforms to the fact that .02 fallibility is quite a bit better than .3 and almost as good as 0 (perfect).

What have they to do with us?

Suppose our predecessors had low reliability, meaning that they had a pretty high rate of believing things that were false. If our reliability level is the same as theirs then, because we should be calibrated, we should have low confidence in our theories. The question is whether our reliability level is the same or similar, and whether theirs being low is a *reason* to think ours is so. If I have been right above, then this is the best way of seeing the issue at stake in the pessimistic induction. The relevant question is not whether their theories were false and whether that should matter to us – as we saw above the pessimist can't get where he wants to go that way. Rather the question is whether their ways of getting to their beliefs were reliable and whether that should matter to us. In what remains I will provide a few substantive arguments, and indications of further similar investigations one might do, and I will provide some thoughts about how the optimist-

when you responded to the weatherman saying there was a 30% chance of rain on a given day by having a 30% confidence in rain that day, it actually rained 60% of the days, then you should update today's weatherman-induced 30% confidence in rain to 60%. This scheme is an extension of the Bayesian constraints on rationality, that relaxes certain idealizations in order to accommodate the fact that while it is unfortunate to discover you are less reliable than you expected, the only irrationality would be to fail to respond to that discovery. I am currently developing the details of this rationality constraint. Defense of a formal constraint is a very complicated matter, but my ground for endorsing calibration here is the empirical evidence that it is beneficial.

pessimist dispute ends up, depending on the outcomes of such investigations. I will conclude that generally speaking the optimist is not currently threatened.

One might wonder how we could investigate our reliability level without already knowing whether our theories are true. To see our reliability rate we would need to do a count among the theories we believe, and see how many are true. But theories that we know are false we no longer believe, and asserting the truth of any or all of the ones we believe would seem to be begging the question here. If there exists any discernible track record for us, must it not be in the broad sense of “us,” the set of human beings doing science, in which case, assuming the quantities above work out in favor of the pessimist, it is dominated by that very (supposed) faulty record of our predecessors? It seems the pessimist wins after all.

What this misses is that the thing we need an estimate of is not track record per se but reliability. Tallying a track record is one way of estimating reliability but it is not the only way. For example, we can reasonably judge a machine to be reliable at doing its job through knowledge of its mechanism, or we can justifiably judge it as more reliable than another machine via our knowledge of the two mechanisms and of how good such mechanisms could possibly be at delivering the result we want. Another kind of painfully simple example to make the point: we know the valences of Hydrogen and Carbon. Say we know the valence needed at the active site of any enzyme that is going to catalyze a chemical reaction we want to know how to produce. Suppose Carbon has that valence and Hydrogen does not. A given molecule with a Carbon sidechain may not work, but we can see that its chance of doing so is fantastically higher than that of a molecule with only Hydrogen sidechains.

We see an analogous thing in comparing our science to that of our predecessors, if we think of methods as mechanisms for leading us to believe the truth and avoid falsehood.¹⁴ Over historical time, our reliability is potentially and probably different from that of our predecessors because we use different methods. Here I use the term “method” broadly, so that it includes, for example, techniques that do not take the form of rules, and also includes techniques that are specific to a given subject matter. The greater the historical time between us and them, the greater the probable difference in methods between us and our predecessors. There is something to the stereotype that the Ancients preferred speculation to observation, and that speculation combined with observation is better than speculation alone. There is something right about the idea that Bacon’s interventional experimental method, and his rules for safeguarding against psychological prejudices, get us more information than mere casual observation, at least, if any method gets us anything at all. There is something to the idea that following the refinements of experimental methods like those described by John Stuart Mill and William Jevons gives us more accurate and safer beliefs. The 20th century brought a massive number of discoveries and refinements in statistical methods, and the new century has brought some further ones already. Prima facie it looks pretty good for a claim that in reliability we are not as bad off as our predecessors.

¹⁴ Kitcher’s (2000b) argument that scientists use what he calls the Galilean Strategy of generalizing the everyday inference from success to truth aims to establish the legitimacy of using an assumption that the two very general properties, success and truth, are linked. But he has moved from discussion of historical matching between instantiations of the properties, to examining the legitimacy of what we could call a “method” that he thinks scientists use for relating the two.

Consider some examples of methods, e.g., simple induction and counterinduction. We may be able to show that induction brings us truths in the long run if any method (e.g. counterinduction) does. (Reichenbach 1949, Salmon 1967) We may be able to show that the straight rule of induction is faster than any other method of ampliative inference. (Juhl 1994) But it does not seem that we can show flatly that induction is likely, or even more likely than counterinduction, to bring us true beliefs. It may be that neither is getting onto the world. However, if the pessimist is going to hang his case on the fact that every scientific enterprise faces the problem of induction generally, he could have left off his induction over the history of science. We can show that problem through one toy case. Moreover, the pessimist usually does want to admit that science gets us some more restricted kind of knowledge than that of theories, since otherwise he is the reviled “radical philosophical skeptic.” Hanging his case on the basic problem of induction would force the pessimist to conclude that justified ampliative inference is impossible, and the scientific enterprise is pointless. He would also, incidentally, undermine his own inductive argument.

Let us assume that the world is susceptible to inductive procedures. I will argue that some methods are more reliable than others – where reliability is as defined above, the probability that q is true given that one has come to believe it by that method. In particular, some of the methods we use are more reliable than any of the methods our predecessors used. To show that a method is more reliable than another method, it is sufficient to show that there are potential errors that the first method is more likely to catch than the second method is, and that there are not errors that the second method is more likely to catch than the first method is. If this is to be sufficient for defending ourselves against the foibles of our predecessors, we must assume that the application of the first method is just as competent as the application of the second. However, the pessimist hasn’t shown that our application of method to the world is worse than that of our predecessors, so we are free to make this assumption.

First let us consider methods pairwise on a question about a connection between two properties, F and G . We could construe the question as whether the one causes the other, or simply whether the correlation will continue as we look at more cases with F . We will apply both methods to a finite data set in which every case with property F also has property G . Now consider induction by generalization. This simple rule says that if all the cases you have seen having property F have property G too, then infer the generalization All F -cases have G . This rule is subject to a simple kind of error, illustrated long ago by Bertrand Russell in the example of a man who’s just jumped out the top floor of a 60 story building. As he passes the 40th floor, he says “So far, so good.” Or consider the chicken who on waking every morning is greeted and fed by the farmer. One morning things change dramatically and the chicken loses his head. In both cases there is information we know that makes the inference or expectation invalid. It is not merely that the conclusion was false, but that adding to the premises makes the inference unwarranted.

It is hard to imagine a human being so uninformed about the world that he doesn’t know independently of his jump that the sequence of correlations he is seeing must shortly come to an end. But to get himself to a warranted belief about whether or not he will go splat, it is not enough to know this; he must also use the information. The method he used did not tell him to do that. Another rule added to it would, namely, what

Reichenbach called “cross induction.” (Reichenbach 1949, 430) This is the use of any information you have that is the basis for another induction that cuts across the trend you see and shows that it cannot, or cannot be expected to, continue. The man’s counterpart who looks for cross inductions is more likely to come to a true belief about what would happen when he jumped than the man who uses only simple induction, because the first man has used inductions about falling objects: they always stop at the earth’s surface, and the longer they fall the higher the impact, etc. This strongly suggests that the jumper’s falling experience will very quickly cease to be so pleasant.

The error that cross induction guards against is precisely one that would reduce our reliability in the sense defined above, for it would be a case of the thing you believe (that your fall is not dangerous) being false. The point here generalizes since any induction is subject to potential cross inductions due to the non-monotonicity of ampliative evidential support; unlike a deductive inference, an inductive inference is erodable by addition of premises. It does not seem that there can be potential errors of the method of induction plus cross induction that are caught by the method of induction alone, since the former method incorporates the latter, and the interaction between the two parts of the former method does not seem to create new error possibilities. It is true that any cross-induction is also susceptible to undermining by further evidence, but that is a new error whose discovery will require that we find a new cross-induction against the original cross-induction. Induction with cross-induction is more reliable than induction alone. Moreover, the idea that this is so only by our own lights has a hollow ring.

A contemporary Bayesian approach to method incorporates this rule in the following way. An idealized agent’s beliefs are represented as a total function from all the propositions of a language to numbers between zero and one that represent strength of belief, that is, as a total probability function. This means that this ideal agent has degrees of belief about all matters, and most importantly here, that all the beliefs that any agent actually has are expected to be taken into account in any application of any rule that falls out of the probability axioms. The probability calculus incorporates a general, rigorous version of the cross-induction rule, and it brings more. Abiding by it allows us to expose equivalence relationships between different formulations of the same and related information, thus helping to identify cross-inductions we might have otherwise missed. The quantitative aspect of the Bayesian approach yields rules for weighing how far a cross induction undermines an induction. In the case of the flying man the undermining is complete. Not so in many other cases. In coming to a conclusion in such cases, we are better off if we can tell whether the first induction or the cross induction is *more* powerful, and this approach provides ways of doing that. This is only to mention some specific features of Bayesian method that are relevant to the execution of cross inductions. Advantages of further features and methods would take several books to describe, but we don’t need to consider those to make the point that it is possible to argue that some methods are more reliable than others, and in particular that some of our methods are better than some, and in some cases any, of those of our predecessors.

Another problem with simple induction noticed probably first by Francis Bacon, and dealt with increasingly rigorously and expansively by John Stuart Mill, Charles Sanders Peirce, Neyman and Pearson, Sir Ronald Fisher, and many others who have followed, is that in the investigation of how potent an observed association between F and G is, one must deal with the possibility that the effect G was brought about by

something other than F that one was not aware of. This is now addressed through an enormous and growing list of techniques: methods for insuring a representative sample of F's, the setup of a control group that doesn't have F in order to see whether F was really what made the difference to G's presence, the randomization of the control group to attempt to insure that variables other than F are not causing any correlation seen.¹⁵ And the foregoing are primitive descriptions of old techniques for squeezing information out of experiment and observation, that have been enriched and supplemented in the last fifty years by entire fields of researchers, in Departments of Statistics for example, and even among philosophers and computer scientists in the case, for example, of causal net programming.¹⁶

Another kind of advance has come from various methods that enable the handling of *more* data. It is hard to overestimate, for example, the advances from the inception of the computer to its latest sophistications. In the beginning, when the computer was invented, the Allies would never have cracked the Germans' secret code Enigma had they not had a computational device that could check thousands of possibilities in real time, since Enigma's translation manual changed every day. A lot of the use of computers, even today, is for scanning of a possibility space to discover points that have desirable properties. That discovery feature of computing is relevant to our reliability issue here. The Allies had to have beliefs every day that some routes, areas, ships, etc., were safe, since otherwise they could have had no offensive operations. Cracking Enigma meant that they were able to succeed at a higher rate of finding certain areas clear when they believed they were, that is, that they were more reliable.

Again, this is a thoroughly primitive example compared to more recent advances. Today's computing is not only faster but has techniques that predecessors didn't imagine. At any given time, then, we have more evidence, and more and better techniques for squeezing it, and an ability to do that faster, than our predecessors' did. One might object that this computing progress has not yet been shown to be relevant because I haven't given a reason to think that more evidence makes our conclusions more reliable. But that objection rests on the basic problem of induction. As pointed out above, if the pessimist about scientific theories rests his case on that, then he didn't need to beat around the bush with the talk about the history of science.

More is all we need

More, and more specific, investigation is needed along these lines of comparing methods over the history of science if we are to see in detail where we stand, and this, I submit, is where the discussion of realism and anti-realism about our scientific theories is best pursued. At a gross level we know that our predecessors were not using certain methods before a certain point because those methods did not exist at that time. But there are many questions about what they actually were doing in each case, and how it compares to what we do for those same topics or related ones. Even if better methods are available

¹⁵ Bayesians do not believe that the procedure of randomization gives an advantage over simply insuring the matching of traits, but both have sophisticated methods for control.

¹⁶ The reader is advised to look up "Design of Experiment" in Wikipedia to see a sampling of the techniques, and that article will be superficial and out of date.

today, are they being used, and used well? And is our confidence actually tailored to their reliability? And even if better general methods are available today they may, for substantive reasons, have no added reliability value for a given topic of investigation. In such a case the calibration imperative would counsel no more confidence in our conclusions than we think our predecessors had a right to. With advances in methods we also know that there can be trade-offs in applying them to any given case. We can be forced to lose or forego certain kinds of information when we use a method that enhances our ability to get other kinds. (MacCoun 1998) There is a lot more to say at the general level too, expanding on the primitive comparisons I have made above. However, though an expansion of that might help to identify and describe methods from the past, and to judge their reliability levels, the general comparison of methods won't be nearly as informative in the decision how confident to be in our beliefs in our specific theories, as it will be when combined with investigation of the particular methods used in the history of research those particular theories emerged from.

The effort to be calibrated requires that we tailor our confidence in our specific contemporary theories to our reliability in coming to espouse them. Methods may have changed generally for the better over historical time, but unless they have changed in the right way between a theory we currently espouse and its immediate predecessor that we regard as falsified, we appear to be vulnerable to the pessimist's argument. Thus we might wonder about the Special Theory of Relativity. We think that its predecessor, Newton's theory, is strictly speaking wrong. But if no new methods were used in the Michelson-Morley experiment that tested for variation in the speed of light over reference frames than had been used in previous belief-formation about the speed of light, then the pessimist could argue that our reliability level can be no better than our frequently failing predecessors, and thus via the calibration imperative that we should reduce our confidence in Special Relativity.

What the relevant methods were and whether the ones supporting the constancy of the speed of light were better, are questions that the pessimist should, and needs to, take up if he means to undermine our confidence in the constancy of the speed of light, one of the building blocks of Special Relativity. This is because the investigations of the long past and the recent past are *prima facie* different. In fact, I think we are on quite good ground even with the Michelson-Morley experiment itself, though it was only the beginning of an increasingly rigorous series of experiments, because that the speed of light was different over reference frames was not so much investigated differently in a direct way but assumed largely for reasons of theoretical coherence by those who came before Michelson, Morley, and their immediate predecessors. The method of assumption seems to be less reliable than the method of experimentation, unless we step into the night of inductive skepticism in which all cows are black.

What if we cannot determine the reliability level of our methods? The task looks more difficult the more subject-specific the method under consideration is, since usually one would need already to know or make assumptions about the subject to know whether such a method is any good. (Although if we cannot make *any* such assumptions, we are back to the basic problem of induction.) A priori arguments like those given above about general method will not be quite as easy to find here. Does it not follow that our right to our confidence is questionable? There are several reasons why this isn't a problem. One is, as explained, the burden is on the pessimist to show our methods are no better than

those of our predecessors. Another is simply that what matters here is the comparison to our predecessors, not the absolute level of our reliability, for the question is whether their reliability problems are problems for us. Moreover, although I have just offered arguments to the effect that some of our methods are more reliable than those of our predecessors, the optimist has no obligation to show that much. We have a cross-induction against the pessimistic induction over the history of science if we find that our methods are merely relevantly *different*. It is the pessimist's burden to show that despite that difference our methods nevertheless have the same (un)reliability. The optimist, it seems to me, is in pretty good shape.

The pessimist may reasonably protest that even if we show our methods to be different from our predecessors', many of our predecessors also had methods that were different from their predecessors'. A lot of good it did them since they ended up wrong a lot of the time. This won't be enough to revive our naysayer's argument, though. The argument of the last paragraph provides us with a cross induction, whose premises are true even when we take this latest pessimistic induction into account. The fact that our predecessors often had different methods from their predecessors breaks the induction over their predecessors to *them*. So the fact that they had different methods means they were not subject to a negative induction. So too for us, by the only induction that is left to make. This may also be why we tend to think of those of our predecessors who used methods that are a subset of ours as being in some real sense justified even when we also think their conclusions were wrong.

One may grant that this is all well and good for hypotheses about which associations between observations will continue, but it can say nothing about our right to be confident in our *theories*, for their claims go beyond observations. Most people have long ago rejected the theory-observation distinction, or at least its epistemological relevance, but there are further replies to this line of skepticism even if we grant the distinction. For in any case it is the skeptic's and not the optimist's burden to say why theories pose a particular problem in principle that hypotheses about the continuation of observable correlations do not. My own view is that we have not gotten very far in our right to have confidence that our most general theories are true (though we have gotten beyond observables) but this is not a matter of principle or because their claims go beyond observables, or because of a pessimistic induction over history. (Roush 2005, Chapter 6) It is what seems most plausible in light of investigation of what is required for evidence, what our methods for achieving it are as a matter of fact, and what they have so far been able to do.

No one has I believe argued successfully that we can never find methods that improve that situation. The most the pessimist philosopher has usually offered here is claims – largely based on general empiricist assumptions -- that no observational difference can show a theoretical difference. I have argued (Roush 2005, Chapter 6) on the basis of general principles of probability that the line that has been based on this idea can be sustained only if we also give up on claims that we have evidential support for hypotheses about observables (Roush manuscript BJPS). This is roughly because of the heavily theory-loaded assumptions we must make about the conditions of the world in order to have enough information for an evaluation of our evidence even as it pertains to generalizations about which observable correlations will continue. That is, the claim that

observations can't confirm or undermine theoretical claims can be sustained only if we give up on doing science at all. (Roush 2005, 197-209)

The argument of this paper also stands against a new pessimistic induction developed recently by Kyle Stanford. (2006; see also 2000a, 2000b, 2003) The inference he describes involves not observables but conceivables. Stanford points out that there were alternative possibilities to our predecessors' theories that were conceivable, and that showed our predecessors' theories false, but which they did not conceive of. We know these possibilities were conceivable, because we have since conceived them. However, there is no reason to think we are different from them; there must also be conceivables we have not conceived. And, roughly, we can't show that they don't undermine our theories because we're not conceiving of them.

Stanford focuses not on the challenge posed by beliefs of a certain sort, those about unobservables, but on ways at arriving at beliefs in science, especially the method of elimination of alternative explanations of phenomena. Thus, the argument of this paper is directly relevant. To perform this induction the cases of our predecessors and ourselves must be relevantly similar in the basis property, which here is proposed to be the affliction by unconceived conceivables relevant to the theoretical question the relevant people in each case are investigating. This property is second-order and affects our predecessors' reliability, because as we know their affliction was partly responsible for their believing false theories. However, granting that we also are subject to unconceived conceivables, the question is whether our predecessors' unreliability that was due to this property says anything about us. For it to do so there must not be any other properties to show that we may be different from those predecessors in our reliability judging possibility spaces. But there have been a lot of discoveries of new methods for dealing with large theoretical possibility spaces. In particular, before the early 20th century possible alternative explanations for phenomena and experiments were ruled out seriatim. The early 20th century saw the blossoming of a first round of techniques for ruling out large classes of theories without *conceiving* of their members. (Roush 2005, 218-223) Since then, of course, there have arisen more, and more sophisticated techniques.

Once we point out that our methods for dealing with unconceived conceivables are different in a way relevant to whether our theories are likely to be true, it is the pessimist's burden to show that despite these differences in method we are no more reliable than our forebears. But more can be argued. The fact that we, unlike our predecessors, don't need to conceive of possibilities in order to rule them out means that it is now suddenly unclear why the unconceivedness of alternative possibilities is so relevant to our reliability. So, it is not just that we have different (therefore possibly and quite plausibly better) methods for dealing with unconceived conceivables, but also that it is questionable that possibilities being unconceived is a worrisome phenomenon. Thus, our use of new techniques for dealing with possibility spaces that are relevant to whether we succeed breaks Stanford's induction in two ways. It is not the burden of the optimist to show that we have a method good enough to rule out all possible alternative theories. The question is whether our predecessors' faults should make us think we are in the same boat, and the pessimist has not shown that they should.

Thinking carefully about the pessimistic induction over the history of scientific failures shows that it must be a meta-induction if it is to make past failures *relevant* to us. It must appeal not merely to the falsity of our predecessors' theories but to the

unreliability of their ways of coming to their beliefs. But even if we grant their unreliability, nothing follows from this about whether we have a right to our confidence in our particular theories unless it is shown 1) that their unreliability is a reason to think we are unreliable and 2) that our believing we are unreliable forces a revision of our first-order beliefs on us. I have argued via our desire for calibration that 2) is true. However 1) has not been shown, since the manifest difference in methods between us and our predecessors breaks the pessimist's induction. He will have to go back to the drawing board and tell us why despite this we are not more likely to have true theories.

References

- Fine, Arthur (1996), *The Shaky Game: Einstein, Realism, and the Quantum Theory*. 2nd edition. Chicago: University of Chicago Press.
- Juhl, Cory F. (1994), "The Speed-Optimality of Reichenbach's Straight Rule of Induction," *British Journal for the Philosophy of Science* 45, pp. 857-863.
- Kitcher, Philip (2001a), *Science, Truth, and Democracy*. New York: Oxford University Press.
- (2001b), "Real Realism: The Galilean Strategy," *Philosophical Review* 110:151-197.
- Laudan, Larry (1981), "A Confutation of Convergent Realism," *Philosophy of Science* 48: 19-49.
- Leplin, Jarrett (1997), *A Novel Defense of Scientific Realism*. New York: Oxford University Press.
- MacCoun, Robert (1998), [Biases in the interpretation and use of research results](#). *Annual Review of Psychology*, 49, 259-287.
- Psillos, Stathis (1999), *Scientific Realism: How Science Tracks Truth*. New York: Routledge Press.
- Reichenbach, Hans (1949), *The Theory of Probability*. Berkeley: University of California Press.
- Roush, Sherrilyn (2005), *Tracking Truth: Knowledge, Evidence, and Science*. Oxford: Oxford University Press.
- (expected 2009), "Inconceivable Support Relations," *BJPS*.
- Salmon, Wesley C. (1967), *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh Press.

Stanford, P. Kyle (2006), *Exceeding our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press.

----- (2003), "Pyrric Victories for Scientific Realism," *Journal of Philosophy* 100: 553-572.

----- (2000b), "Refusing the Devil's Bargain: What Kind of Underdetermination Should We Take Seriously?" *Philosophy of Science* 48, Supplement: S1-S12.

----- (2000a), "An Anti-Realist Explanation of the Success of Science," *Philosophy of Science* 67: 266-284.

Tenney, E.R., MacCoun, R.J., Spellman B.A., and Hastie, R. (2007), "Calibration Trumps Confidence as Basis for Witness Credibility," *Psychological Science* 18: 46-50.