

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

Cross-references:

007 Philosophy

016 Scientific Revolutions

032 Risk Perception

034 Public Attitudes Toward Science and Technology

041 Peer review

“Fallibility and Authority”¹

Sherrilyn Roush

Over the centuries since the modern scientific revolution that started with Copernicus, Galileo, Kepler, and Newton, two things have changed that have required reorientation of our assumptions and re-education of our reflexes. First, we have learned that even the very best science is fallible; eminently successful theories investigated and supported through the best methods, and by the best evidence available, might be not just incomplete but wrong. That is, it is possible to have a justified belief that is false. Second, we have learned that it is impossible, even for scientists, to maintain the Enlightenment ideal of “thinking for oneself” on every matter about which we want to have, and do think we have, knowledge; the volume of information involved makes us all epistemically dependent on others. (Kant 1996) Scientists in practice have adjusted to these developments much more easily than have lay people. It is also easier to adjust in scientific practice than it is to explain these matters explicitly and accurately to others. To do so it is helpful to consider our epistemological situation precisely, and to understand the broader cultural ideas and historical forces at work in modern science and its public reception.

When scientists present results and arguments in public, responses run the gamut from reflexive trust to indiscriminate mistrust. The extremes of this spectrum of response have destructive consequences, and ignorance of the way science works and of the primary scientific subject matter

¹ Sherrilyn Roush is Associate Professor of Philosophy, and a faculty member of the Group in Logic and the Methodology of Science, at U.C., Berkeley. She is the author of *Tracking Truth: Knowledge, Evidence, and Science* (Oxford University Press, 2005). Her recent papers include “The Value of Knowledge and the Pursuit of Survival,” “Second-Guessing: A Self-Help Manual,” and “Randomized Controlled Trials and the Flow of Information.” She is currently writing a book called *Rational Self-Doubt*.

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

encourages these extremes. Ignorance about how science works can be reduced somewhat by greater education of the public. However, the audience's relative ignorance of a specific topic will always and necessarily, because the scientist always has more knowledge of her subject matter than a layperson does. This creates what may seem an unsolvable problem, for how can the layperson exercise rationality in deciding to believe what a scientist says? If the layperson were in a position to evaluate whether the scientist should be believed, then he wouldn't need the scientist but could figure out the answer himself! It seems that the layperson has the choice of blind trust or blind mistrust. In considering epistemic dependence below, we will see that this problem is only apparent. This is a good thing, since the scientist herself faces the same situation; she is in a layperson-expert relation to many other scientists.

Cynicism and errors concerning the admission of fallibility

One key flashpoint in the public reception of science is the admission of fallibility. When, in addition to stating results, scientists admit that they or their predecessors have been wrong in the past or that they might be wrong now, they are often greeted with one of two curiously opposite responses. One is an abject skepticism that concludes that no one, scientists included, knows anything. Another is the confident conclusion that one is justified in believing whatever hypothesis one chooses. Both responses imply a leveling of authority, since in either case the scientist is taken to know no better than anyone else. Both the thoroughgoing skepticism and the personal confidence are usually defended by the familiar claims that "they haven't *proved* it" and "it's *just a theory*."

The word "proof" is properly reserved for mathematical and logical arguments, where the standard for success is that the premises rule out all possible alternative conclusions, that is, ensure that it is logically impossible for the conclusion to be false. Proof is an *infallibilist* standard because it requires ruling out any possibility of error in order to be counted as successful at gaining knowledge or justified belief. This standard is typically being invoked when the claim is made that scientists haven't proved their theories, for in such cases this is stated not as the obvious fact it is, but as an accusation intended to undermine the epistemic status of scientific theories. Such undermining could occur, though, only if proof, in the mathematical sense, was a standard appropriate to science.

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

By contrast, *fallibilism* says that we can have justified belief in a proposition *p* even if our evidence about *p* does not rule out the possibility that *p* is false. The reasonableness of this view depends in part on the distinction between possibility and probability; to admit that one *might* be wrong is not to say that this is *likely*. Scientists, like all human beings, are in fact fallible – might be wrong – but it does not follow from this that they do not have knowledge, unless an unrealistic infallibilist standard is assumed. That one is not ten feet tall does not imply that one must be ten feet tall in order to succeed at basketball.

Infallibilism clearly supports the extreme responses to scientific testimony, for if we do require infallibility in order to count a person as justified at all, then the admission that we are not perfect implies that scientists achieve no distinctions at all among better- and worse-supported theories; analogously, if God is dead, and God was the only possible ground for morality, then everything is indeed permitted. For scientists, who have experience of the detailed work of gathering, eliciting, and judging better and worse evidence, the conclusion that there are no distinctions – that without epistemic perfection we have no knowledge at all – is foolish. However, for those without knowledge of and experience with scientific discussions, this process is a black box, so they have a harder time accepting the gradations that fallibilism requires. The infallibilist mistake thus undermines the perceived authority of scientific results.

Although assuming an infallibilist standard betrays a deep misunderstanding of the way empirical investigation works, the inference is a psychological instinct that has been with even the most intellectually sophisticated people until a century or so ago, and initially it had a clear rationale. During and after the first modern scientific revolution it was often believed and explicitly argued that the attempts at gaining knowledge of the world that came before Copernicus, Galileo, Kepler, and Newton had not only led to false theories but were not science at all; the ancients and medievals did not conduct themselves in the methodical and empirical way that the new giants had. Indeed that was why, it was thought, they had not made the astonishing leaps of progress the new physicists achieved; the reason they had come up with false conclusions was that they were not following *the scientific method*. (Bacon 2000; Descartes 1998, 2000; Newton 1999)

The modern revolution thus made it easy to believe that because we *did* follow the scientific method our conclusions would not later be

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

found to be far off the mark. Thus, a presumption existed that the new physical theories, of mechanics and subsequently those of electromagnetism and thermodynamics, would never need to be revised. They would certainly need augmentation and filling out as scientists continued to explore and accumulate knowledge of more aspects of the world, but the mechanisms were right, and what remained was mopping up. One could make mistakes of detail, of course, but if one used the scientific method one would not come up with patently false scientific theories.

With the 20th century, and the theories of relativity and quantum mechanics, this view had to yield in the face of a second massive revolution in modern physics. (Reichenbach 1942) The force model that had been used in mechanics and electricity and magnetism, while remaining just as successful in familiar domains of application, is not a correct picture of what is happening behind appearances when a planet orbits the sun or an electron interacts with a proton. This development made it impossible any longer to explain why our predecessors came up with strictly speaking false pictures by appeal to the idea that they were not doing science – what would science be if we had to say that Newton was not doing it? Rather, one had to conclude that even the greatest science is fallible; it may yield false, even if successful, models at any level of generality, even if the practitioners make no mistakes of inference or method, that is, even if they are justified in their beliefs.

Scientists are usually comfortable with this revised understanding of the way science works, physicists freely admitting, for example, that the Standard Model of particle physics is not the final theory (due to a surfeit of adjustable parameters). Physicists also freely admit that Relativity and Quantum Mechanics do not fully mesh, implying that neither is the final theory and that new concepts are probably needed. Yet despite the implicit admission that our best theories are strictly speaking false, there is every reason to think that these theories get a lot right, that they contain crucial building blocks for the next stages which will be retained in the new theories. Scientists are not only quite comfortable with the early twentieth century revision of deep theories, but are expecting, or at least dreaming for, such deep revisions to come around again.

This combination of attitudes can make fallibilism seem more puzzling, even paradoxical. How could it make sense to be confident in one's theory, and simultaneously to be confident that something in it is false, especially when one is not sure where the false parts will be found to

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

be? This is analogous to the so-called “preface paradox,” in which a person who has written a book, and is confident in each of its claims due to the arguments for each of those results, also admits in the preface of the book that it is likely that at least one of those many, many claims is wrong. How, it is asked, can one assert the conjunct propositions p_1, p_2, \dots, p_n and also assert that at least one of these conjuncts is false?

Indeed, one cannot do this consistently if what it is to believe or assert is to have or present 100% confidence in each of these claims. This is not consistent because the claims in question are logically contradictory; all of the claims of your book are true if and only if not one of them is false. However, one need not, and should not, imagine belief or assertion in this way. The author may be very confident in each of the claims of her book without being 100% certain of any of them. Similarly, she need not be certain that there is a mistake somewhere in order to be highly confident that there is. This more realistic way of understanding what a person is saying when she notes her fallibility even while making bold assertions dissolves the supposed paradox (Roush 2010), and is perfectly consistent with understanding rationality as requiring that a person’s degrees of belief be probabilistically coherent, that is, relate to each other in the way required by the probability axioms. (Jeffrey 2004)

We can see this as follows. If a person has 99% confidence in each of the claims of her book, it will be not only possible but indeed obligatory for her to have high confidence that at least one of those claims is wrong. This is because the probability of a conjunction is the product of the probability of the conjuncts (if the conjuncts are independent, which can be assumed here without loss of generality), and the product of fractions less than one gets smaller very rapidly as the number of fractions increases. It would take only 59 conjuncts for a person with 99% confidence in each of them to be obliged to have approximately 95% confidence that one of them is wrong. Books typically contain more than 59 sentences. If we wonder why scientists don’t feel the need to put explicit fallibility notices in journal papers the explanation may be as simple as that papers do not contain as many claims as books do.

Of course, if a person is 100% certain of each of her claims, then no matter how many there are the product of the confidences in them will also be 100%. ($1^n = 1$, for all n .) It follows probabilistically that she is obliged to have zero confidence that any of them is false. However, if her confidence is 100%, then probability forbids her from ever revising those beliefs, no matter what empirical evidence arises. This person we are now

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

imagining has moral certainty, so she would not have been inclined to admit fallibility on these matters in the first place. This character would not have written the imagined preface, so there is no paradox in this case either.

Public cynicism about revision

Scientists are rational in expecting revision of some of their hypotheses and theories while being confident about each of them. However, in the public reception of science the phenomenon of revision is another problematic flashpoint. For many years the public was told that hormone replacement therapy (HRT) was good for middle-aged women. More recently, due to a study that surprised the medical establishment, this view has been revised to say that the therapy brings significant increase in the risk of breast cancer and stroke. (Beral et al. 2002, Beral et al. 2003) Apparent reversals of scientific claims happen frequently, if not always so dramatically, and a layperson's response to witnessing this is often to mistrust all scientific results – scientists can't make up their minds! Why should we take them seriously today when they are liable to tell us the opposite tomorrow? If one is ignorant about how science works, as well as of the subject matter, then this response has an unfortunate plausibility.

This cynical, or at least exasperated, response to revisions rests partly on mistaken assumptions about scientific inquiry and partly on unfortunate distortions wrought by the marketplace, for which science gets blamed. The first type of mistake occurs when overly generic statements of scientific results make it look as if new research contradicts old research when it does not. Second, there is a mistaken belief that revision is an indication of something bad or unreliable about science. Finally, because human beings crave simple, definite answers and sure solutions, journalists are encouraged to bring satisfaction by overstating results, and companies stand to make a profit by overselling tentative and limited conclusions about efficacy. Overstated conclusions are more likely to turn out false in the course of time and thus contribute to a mistaken impression of dramatic revision on the part of scientists. Scientists do revise claims. However, they revise less dramatically and less often than it appears.

If we state early scientific conclusions about HRT as the generic claim that it is good for middle-aged women, this will be in direct contradiction with later finding any respect in which HRT is bad for such

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

a woman. However, claims that HRT is good for treating hot flashes and preventing osteoporosis – which have largely stood the test of time – are not contradicted by the later finding that hormone therapy is bad with respect to breast cancer and stroke. Thus if the earlier claims are stated at the warranted specificity, then we see that what we had with the news of the 2002 study showing greater risks of breast cancer and stroke was not a revision but an expansion of our knowledge of the effects of this treatment.

Ignoring the specifics of a sequence of scientific results can make them look contradictory when they are not. Hormone therapy, like anything else, can be good for one thing and bad for another with no contradiction. However, the promise to women that they could be “feminine forever” and the desire of doctors to help patients in whatever way possible created a market for pharmaceutical companies to pursue their natural inclination to make money. The feedback of each of these three groups to each other created confidence in overgeneralized claims, which was unfortunate for some women and also contributed to a misleading impression that science is unreliable.

There is another way that scientific claims can be revised which involves a merely apparent contradiction of the previous view. One may find, for example, that HRT increases the risk of cardiovascular disease on average. However, one may also, or subsequently, find that the average hides important variation, for example that though HRT given in later years increases this risk, when the therapy is given in early menopause it *decreases* the risk. Or we may find that different preparations of the hormones – including or not including progestagen, for example – makes a big difference to the risks. The greater specificity of the new results can support different recommendations than practitioners gave earlier – an early menopausal woman now having benefit and much less risk than expected. However, in neither case do the new findings contradict the previous claims. It remains true that HRT carries this risk of cardiovascular disease *on average*, even while it carries benefit for people with more specific traits, and relatively high risk for those that lack those traits. Careful reporting of results at the warranted level of specificity can help to avoid the false impression that scientists are changing their minds.

Of course, the kind of revision in which a newly accepted hypothesis or theory flatly contradicts, and does not redeem, the old one also happens. Wegner’s geological theory of continental drift was directly denied for some time, and later accepted. This kind of revision is perhaps

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

hardest for the layperson to resist becoming cynical about. How much can we trust scientists' current claims when they change their minds 180 degrees even about big things? However, revisions, even dramatic ones, are a straightforward and indeed necessary consequence of progress in empirical investigation. This is because induction – inference from incomplete evidence, as in empirical science – is *non-monotonic*: additional evidence can undermine the previous claim that the total evidence supports a given hypothesis, and thereby undermine our justification for asserting it. As the total set of evidence increases, the probability of a given hypothesis relative to the total evidence set can go up, and down, and even up and down again and again. In such cases changing our confidence accordingly at each new stage is the only appropriate response to the new evidence. (Koons 2009)

This erodability of inductive inference stands in contrast to deductive inference: from the facts that all men are mortal and that Socrates is a man it follows that Socrates is mortal. Adding further premises, that is, new evidence, can never undermine the legitimacy of that inference. We might discover that not all men are mortal, but that would involve a rejection of one of the premises. Monotonicity means that there is no possible additional fact that could show that it does not follow from the premises that Socrates is a man and all men are mortal that Socrates is mortal.

Importantly, though it is a consequence of the erosion of an induction that we are no longer justified in believing the conclusion as strongly as we were before, the erodability of induction attaches to the inference, not the conclusion. The new evidence that undermines our previous justification need not falsify the conclusion. This is why the addition of new evidence at a still later stage could increase the probability of the conclusion again.

For example, if our only evidence was thousands of swans, all of which were white, as it was for Europeans before they encountered Australia, we would be justified in concluding that all swans are white, even though as a matter of fact they are not. We can find the conclusion false by going to Australia, but our justification for believing all swans are white can be undermined without finding a counterexample. For example, we may notice that color is often variable within other species. This does not tell us there is a black swan, but it does say that observing a bunch of white swans does not by itself justify the same degree of belief that they all are white that it did before. This phenomenon, called *cross-induction*,

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

in which a legitimate inductive inference is undermined by further evidence that does not necessarily contradict the conclusion, is crucial to understanding why the fact that scientists revise their hypotheses and theories does not mean their assertions are not or were not justified, or that we are not justified basing our decisions upon them.

Whether one is justified or not in making an assertion does not depend on whether an all-knowing God would have made the assertion, but on whether one has drawn an appropriate conclusion from the evidence one has. Thus, justification of a belief is relative to a body of evidence. This means that even when new evidence makes a different conclusion justified, this does not show that one had not been justified in one's former belief; one had a different evidence set then, and different conclusions are justified when the bodies of evidence are different. Thus, the fact that scientists revise their hypotheses and theories – that they used to believe one thing and now believe another – does not imply that any of their beliefs are not justified at the time they are held. Different beliefs can be justified at the two times as long as the scientists have different sets of evidence at the two different times.

Given the erodability of the best inductive inference, and given that there is an infinite amount of possible empirical evidence we will not reach the end of, it would be suspicious if scientists did not revise their views over time; it would suggest they were not increasing their store of evidence sufficiently or sufficiently openly and rigorously. Thus, seeing scientists revise their views is a good thing and should be reassuring.

Our evidence is always incomplete, and thus our most justified beliefs are permanently liable to revision. This can be a difficult phenomenon to accept, but everyday life is also full of examples of it. It can be dramatic to discover evidence that a close friend is untrustworthy. Yet we do not consider that we were unjustified in formerly believing the person was trustworthy (unless the new evidence was something we think we should have seen all along). We think we were unlucky, but not unjustified, to put ourselves in dependence on that person. It is the same when we make decisions on the basis of the testimony of the best scientists. We take a gamble, and may not in the end get what we want, but that does not mean we or they were not justified in taking that risk. It is also possible and common to discover evidence that the evidence against our trusted friend does not have the significance we had, quite rationally, attached to it. Perhaps the stealthy phone calls to one's husband

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

turn out to be followed by a surprise party that the two of them had been organizing. This would provide a cross-induction.

Once again, though old evidence combined with the new evidence does not support one's old conclusion that the friend was untrustworthy, it does not say that that conclusion was not justified at the earlier time on the basis of the smaller evidence set. Moreover, the new evidence does not logically imply that one's suspicious hypothesis about the friend was false. Further evidence – such as stealthy phone calls continuing after the surprise party – could make a new total evidence set that supports the original conclusion that one's friend is untrustworthy. This is our life with incomplete evidence; changing our beliefs, and changing them again, is sometimes the rational thing to do. (Roush 2010)

Of course, in the absence of continued stealthy phone calls, or similarly suspicious behavior, it would be quite irrational to withhold trust in one's friend after the party. The evidence set one has does not support that mistrust; untrustworthiness is merely a logical possibility. Similarly, it is irrational to withhold trust, of the appropriate probabilistic sort, in an announced scientific result in the absence of evidence that specifically casts doubt on the justification of that result. The blank observation that scientists have been wrong before does not justify skepticism about specific conclusions. The generic fact merely that one was wrong before about one's friend does not justify believing now that she is untrustworthy: indeed, what one had been wrong about before was precisely that she was *untrustworthy*, and *that* error certainly does not give one reason to believe she is untrustworthy now.

Probability and responsibility

If hormone therapy is unqualifiedly good, or unqualifiedly bad, then a doctor's decision whether to recommend it and a patient's decision whether to undergo it are easy. Another reason hazardous overstatements of scientific results are attractive is a wish that our decisions would be simple, even that the facts, perhaps even the scientist, would make our decisions for us. However, it is never possible for the facts or the scientist (qua scientist) to make our decisions for us, even when our beliefs are certain and unqualified. The qualifications that accompany responsible reports of scientific results simply bring this into sharper focus.

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

Hormone therapy is good for some things and bad for others. That is, it has both benefits and risks. To take the treatment is to accept both, and whether one should do this depends on how much one values the benefits and how much one dislikes the risks. No fact about the world determines these evaluations. When the world will not satisfy all of a person's goals at once – when a complex of benefits and risks is in question -- it is obvious that the subject has a burden to decide on a ranking of her preferences. (Peterson 2009)

An analogous point applies to the fact that, especially in the domains most relevant to policy and to the decisions of ordinary life, scientists typically present their results as probabilities, not certain predictions about the effects of one thing upon another. Probability can generate frustration and the feeling that one is not being given the answer to the question, that the scientist is hedging her bets to avoid responsibility. However, the report of probabilities is typically not hedging but an effort to be as accurate as possible about what is known and to assert no more than is known. Any other form of address would be misleading, and indeed *irresponsible*. (Howson and Urbach 1996, Talbott 2008) Frustration at this situation may come again from a wish or expectation that scientists would make our practical decisions for us. But even if there were not probabilities but certainties, the question what to do would not follow from facts alone. Suppose the following were certainties: if a rainforest is lost, then three cities will be under water within 20 years, and saving a rainforest has zero cost. Then the decision what to do is obvious, but only because we place greater than zero value on cities, a value judgment that does not depend on scientific testimony.

Authority and rationality

Over the period since Copernicus there has been a strain of belief that has not been revised, especially among the general public. This is the conviction that it is epistemically superior to think for oneself, to depend on one's own reasoning in determining what to believe. A romantic view of science, in which the genius does everything in isolation and independently of everyone else, is also a part of this legacy. The idea has an august lineage. Long before the 18th century Enlightenment – the Age of Reason in which we were exhorted to apply our own minds instead of trusting in authority – Martin Luther declared “Here I stand, I can do no other,” setting his own conscience against the deliverances of the church institution. Galileo similarly whispered about the earth “Yet it moves,” at

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

the trial for his heliocentric heresy in which he had officially recanted. René Descartes famously instructed us how to build the foundations of all science on the knowledge simply that one thinks, and thought it imperative in building that foundation that one let go of beliefs held merely by habit, custom, and trust in others. One must evaluate every brick of the edifice oneself if it is to continue sturdy. (Descartes 1984, 1998) Responsibility for and authority over one's own beliefs – freedom of conscience – has come to have an even more robust life outside of science in political institutions, the most obvious manifestation of which is the first amendment to the U.S. Constitution, which protects the individual's freedom of speech, and expression more broadly.

The idea of the individual's freedom of conscience is a powerful and beneficial meme, but there are two distinctions that tend to be overlooked, with destructive consequences for the reception of science. One is the difference, discussed above, between values, or preferences, and matters of fact. Whether preserving as many species as possible is a worthy goal is a question of values. What will happen to the rainforests if CO₂ is not decreased is a question of fact. Most discussions involve complex combinations of the two kinds of belief – in deciding whether preserving as many species as possible is a worthy goal one will consider other factual beliefs one has, such as whether the survival of human beings depends on it. Identifying and distinguishing fact versus value questions within a discussion is useful because the tools for addressing the two kinds of issues are usually different, and the scientist is not per se responsible for our value decisions.

Another neglected distinction is between legality and rationality; the former does not imply the latter, and this becomes especially apparent when our presumption of freedom of conscience bleeds from the domain of values into that of facts. There the idea is problematic, and on some understandings of the claim it is plain wrong. If one acknowledges that scientists have more knowledge than one does about a topic, as is most often the case where their expertise is relevant at all, then it is not rational to form an opinion based solely on one's own cogitations. Thanks to the first amendment it is perfectly *legal* for one to believe and announce just whatever one wants to about the way the world is. That does not mean that just any belief one professes is *rational*.

We would not have most of the knowledge we typically assume we have about the world without placing ourselves in dependence on others in forming our beliefs. One thinks one knows that smoking causes cancer,

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

but observing that the two smokers one knew got cancer has no evidential significance to the question. Since one did not carry out studies on the possible causal connection between smoking and cancer, one's legitimate confidence about this connection comes from accepting the results and trusting the expertise of other people. Similarly, one knows that the Grand Canyon exists, and that cheetahs run fast, even if one has never laid eyes on either phenomenon, and this is because of trust that other people have verified these things.

Every scientist today is likewise epistemically dependent on other scientists in order to be able to make discoveries at all. For if a scientist had to repeat every experiment of the predecessors in her area in order to be counted as knowing the results they reported, she would never have the time to do a new experiment furthering our understanding. Scientists could not build on the discoveries of others and our knowledge would not grow. Descartes notwithstanding, there is nothing irrational or deficient in this epistemic interdependence among people. The volume and complexity of information that science deals with in its current state is staggering; it would be impossible even for a Galileo to hold it all in his mind, much less to evaluate step by step. (Hardwig 1985)

Epistemic interdependence among scientists can lead observers to the conclusion that the results they offer are not so much true as merely a consensus opinion, and thus not necessarily better than the belief that the earth was flat, which was also once held by a majority. Thereby once again the reports of scientists may be taken to be no more credit-worthy than one's own opinions. However, while being in a group makes group-think possible it does not make it necessary. In fact, there are both cultural and formal safeguards against mutual internal reinforcement of the opinions of scientists in a community. For example, scientists are rewarded for coming to hypotheses and conclusions that are distinct from and even challenge those of their rivals, and peer review of publication puts work under the scrutiny and power of one's rivals. (Lamont 2010, Kassirer and Champion 1994) Such mechanisms tend to prevent consensus from coming too cheaply. Any particular safeguard may fail at a given time, of course, but the mere fact that scientists form a mutually interacting and interdependent group does not mean that their conclusions are unrelated to the truth. If cooperation within a group takes a healthy and mutually critical form, then the group can be expected to be more likely to achieve its goal than the same set of individuals could if they were not interacting. (Kitcher 2001)

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

If one is in an epistemically inferior position then trusting others' testimony is the only way it is *possible* to be rational, but how can this trust actually be rational? Since one does not know the subject matter, how can one know whether the expert does either? It appears that this trust would have to be adopted on blind faith. This skeptical conundrum contributes to resistance to seeing ourselves as dependent on the knowledge of others.

Fortunately, the problem is merely apparent. An analogy will expose the mistake in this line of reasoning. Suppose one is not a great cook but likes fine food, so one goes to restaurants, sometimes restaurants one has not been to before. One trusts that neither the restaurateur, nor the chef, nor the waiter is trying to poison one, and that all are competent at delivering food that is at least safe. One believes this even though these people are perfect strangers. Often one is even confident that the food will be of high quality despite never having been to this restaurant before. How could one be justified in these judgments when one has never tasted their products? Even if one watched the preparation, that would give little grounds for judging the quality, since one is not a cook oneself. We frequently trust our very lives to perfect strangers without having personally verified any of our assumptions about them or their expertise.

The point is not to become skeptical of unfamiliar restaurants, but rather to recognize how rational our trust in such cases can be and usually is. We do not have to know the chef, or know what he knows, and we do not have to taste the food, in order to have reasons to believe that we will not be poisoned and that the food will be something between tolerable and very fine. Typically one would not have trust that the food will be high quality without first having testimony from a restaurant review, or a friend known to have sufficiently high standards, that the place served good food. As for poisoning, the very fact that one does not know the restaurateur or her staff means that they probably have no incentive to kill one. And though one does not normally think about it, one reason to trust that the food will not cause illness is the knowledge that the health department makes regular inspections, and code violations carry the threat of shutting the establishment down. And if more than a few people actually got sick at a restaurant, the news would travel and might put the restaurant out of business before the reports of illness even got to one.

The incentive structure all of this creates gives you strong reasons to feel confident that the food will not make you sick. That we are epistemically dependent on other people's judgments does not imply that

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

we are irrational or trust blindly without reasons. We can evaluate whether to believe a given witness, and evaluate the general system that produces these witnesses, via cues that are accessible to those who are not the witnesses themselves. These cues do not rule out the logical possibility that the restaurant will serve you bad food, or even that it will poison you. However, as fallibilists we do not require fulfilling that standard to be counted as justified.

The fact that I do not judge everything for myself – that I don't know what the chef knows, that I do not taste the food before I first taste the food – does not imply that I do not judge *anything* for myself. For example, I will be the judge of whether a given friend has dining standards high enough or similar enough to mine to be taken as an authority on where to eat. I will have shared meals with her before and thereby build up evidence about her tastes. One similarly judges a food critic by either his track record or reputation. I do not verify with my own eyes or tongue every claim or meal that I trust. Nevertheless, I remain the ultimate judge of what to believe or eat, in an obvious and important sense: I decide how to weigh my sources of information for their credibility, and I can do this on the basis of evidence and arguments. Epistemic dependence does not imply that I do not or should not think for myself, even in the extreme case where I know nothing at all about the primary subject matter. Thinking for oneself remains a rational obligation, even when the deliberation is only about whom to trust.

On any matter except what one immediately senses in the here and now, one's evidence and reasons for belief will always be indirect. The kind of rationality a layperson exercises in deciding to believe a scientific expert is not different in kind from that we exercise continuously in daily life. That the best justification one can get for believing a statement of physics is to have good reasons for believing that the scientific expert has good reasons to believe that statement, is no different in kind from the fact that the best one can do for justification of one's belief that one was born on a certain day is to have good reasons to believe that one's parents have good reasons to believe one was born on that day.

There is a difference of degree, of course. We are more conscious of the concession of authority that we make in deferring to scientists because the disparity of expertise is extreme. We know what it is like to experience the day on which a person was born, and can thus trust a parent's testimony with a sense that we know on what sort of basis he or she knows the day. One who has not engaged in science does not have

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

much to go on to imagine what kind of thing scientists do to procure their evidence and what good grounds for scientific claims look like. This can lead to discomfort, but does not mean that trust is irrational. Some people cannot imagine how to cook, yet they can have good reasons to trust chefs.

In summary, a mistaken infallibilist view of knowledge, in which one imagines that in order to be justified one is required to have evidence that makes it logically impossible to be in error, is the source of much cynicism and confusion about science and its expert testimony. Even though it remains possible that we are wrong, we are often able to evaluate how probable or improbable such error is. After admitting our fallibility our options are not believing nothing or believing whatever we want. Another source of cynicism comes from witnessing scientists revising their conclusions. However, there is less outright retraction than meets the eye, and given the necessarily non-monotonic nature of empirical inference, it would be highly suspicious if scientists did not revise their conclusions since it would suggest they were not acquiring new, diverse evidence and evaluating it critically. Witnessing revision should thus be reassuring. Finally, no individual can escape dependence on other people for the justification of claims about matters of fact. However, although scientists do in general deserve to be granted authority over their topics of expertise, in virtue of the fact that they *know more*, this does not in fact undermine the individual's ability and responsibility to think for himself. He must evaluate the credibility of sources using the cues available for this purpose, such as the incentive structure of the research and whether there is consensus in the community. A person can have his own good reasons to believe, but he also can have good reason to believe in virtue of having good reason to believe that others have good reason to believe.

References

Bacon, Francis 2000. *The New Organon, Or True Directions in the Interpretation of Nature* (1620). Edited by Lisa Jardine and Michael Silverthorne. Cambridge: Cambridge University Press.

Beral V, Banks E, Reeves G. 2002. "Evidence from randomised trials on the long-term effects of hormone replacement therapy," *The Lancet*. Volume 360, Issue 9337: 942-4.

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

Beral, V., Banks, E., Reeves, G., Bull, D. 2003. "Breast cancer and hormone-replacement therapy: the Million Women Study," *The Lancet*, Volume 362, Issue 9392: 1330-1331.

Brown, Tracey 2006. "I don't know what to believe:' understanding peer review is key to developing informed opinions about scientific research." *Nature* | doi:10.1038/nature04998 in <http://www.nature.com/nature/peerreview/debate/>

Descartes, René 1984. "Meditations on First Philosophy," in *The Philosophical Writings of Descartes*, Volume II, John Cottingham et al., translators. Cambridge: Cambridge University Press, 12-15.

----- 1998. *Discourse on the Method for Conducting One's Reason Well and for Seeking Truth in the Sciences*, third edition, Chapter 2. Indianapolis, Indiana: Hackett Publishing Company.

----- 2000. *Rules for the Direction of the Mind*. Indianapolis, Indiana: Bobbs-Merrill Co.

Hardwig, John 1985 "Epistemic Dependence," *The Journal of Philosophy*, Volume 82, Number 7: 335-349.

Howson, Colin, and Peter Urbach 1996. *Scientific Reasoning: The Bayesian Approach*. 2nd Edition. Chicago: Open Court.

Jeffrey, Richard 2004 *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.

Kant, Immanuel 1996. "An Answer to the Question: What is Enlightenment? (1784)," in *Practical Philosophy*, translated and edited by Mary J. Gregor. Cambridge: Cambridge University Press.

Kassirer, Jerome P., MD and Edward W. Campion, MD 1994. "Peer Review: Crude and Understudied, but Indispensable," *JAMA* 272:96-97.

Kitcher, Philip 2001. *Science, Truth, and Democracy*. New York: Oxford University Press, 2001.

Koons, Robert 2009. "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy (Winter Edition)*, Edward N. Zalta (ed.), URL =

in

Leadership in Science and Technology: A Reference Handbook;
William Sims Bainbridge, ed., SAGE Press, 2012.

<<http://plato.stanford.edu/archives/win2009/entries/reasoning-defeasible/>>.

Lamont, Michele 2010. *How Professors Think: Inside The Curious World of Academic Judgment*. Cambridge, MA: Harvard University Press.

Newton, Sir Isaac 1999. "Rules of Reasoning," in *The Principia: Mathematical Principles of Natural Philosophy*, translated by I. Bernard Cohen and Anne Whitman. 1st edition. Berkeley: University of California Press.

Peterson, Martin 2009. *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.

Reichenbach, Hans 1942. *From Copernicus to Einstein*. New York: Philosophical Alliance Book Corporation.

Roush, Sherrilyn 2010. "Optimism about the Pessimistic Induction," in *New Waves in the Philosophy of Science*, edited by P.D. Magnus and Jacob Busch. New York: Palgrave-Macmillan.

Talbott, William 2008. "Bayesian Epistemology", *The Stanford Encyclopedia of Philosophy (Fall Edition)*, Edward N. Zalta (ed.),
URL =
<<http://plato.stanford.edu/archives/fall2008/entries/epistemology-bayesian/>>.