

CLUES TO THE PUZZLE OF SCIENTIFIC EVIDENCE

SUSAN HAACK

University of Miami

Abstract

The evidence with respect to scientific claims is like empirical evidence generally — only more so: more complex, more dependent on instruments, etc., and usually a shared resource. Warranted scientific claims are always warranted by somebody's, or somebodies', experience, and somebody's or, somebodies', reasoning; so a theory of warrant must begin with the personal and then move to the social before it can get to grips with the impersonal sense in which we speak of a well-warranted claim or ill-founded conjecture.

The liberty of choice [of scientific concepts and theories] is of a special kind; it is not in any way similar to the liberty of a writer of fiction. Rather, it is similar to that of a man engaged in solving a well-designed word puzzle. He may, it is true, propose any word as the solution; but, there is only *one* word which really solves the puzzle in all its parts. It is a matter of faith that nature — as she is perceptible to our five senses — takes the character of such a well-formulated puzzle. The successes reaped up to now by science... give a certain encouragement to this faith. [ALBERT EINSTEIN]¹

What is scientific evidence, and how does it warrant scientific claims? That honorific usage in which “scientific evidence” is vaguely equiva-

Principia, 5(1–2) (2001), pp. 253–81. Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil. © 2002 Susan Haack. This paper is an abridged version of ch. 3 of Susan Haack, *Defending Science — Within Reason* (forthcoming).

lent to “good evidence” is more trouble than it’s worth. When I write of “scientific evidence” I mean, simply, *the evidence with respect to scientific claims and theories*. Scientific evidence, in this sense, is like the evidence with respect to empirical claims generally — only more so: more complex, and more dependent on instruments of observation and on the pooling of evidential resources.

The only way we can go about finding out what the world is like is to rely on our experience of particular things and events, and the hypotheses we devise about the kinds, structures, and laws of which those particular things and events are instances, checked against further experience and further hypotheses, and subjected to logical scrutiny. The evidence bearing on any empirical claim is the result of experience and reasoning so far, a mesh of many threads of varying strengths anchored more or less firmly in experience and woven more or less tightly into an explanatory picture. So I look at questions about evidence, warrant, etc., not in pristine logical isolation, but in the context of facts about the world and our place as inquirers in the world. And I deliberately eschew the familiar Old Deferentialist jargon of the confirmation of theories by data or by observation or basic statements,² to signal that my conception of evidence, presupposing no distinction of observational and theoretical statements, is considerably ampler than “data”; that my conception of warrant is ineliminably temporal, personal, and social; and that my account of the determinants of evidential quality is not purely formal, but worldly, and not linear, but multi-dimensional.

Scientific evidence, like empirical evidence generally, normally includes both experiential evidence and reasons, and both positive evidence and negative. It is complex and ramifying, structured — to use the analogy I have long found helpful, but only recently found anticipated by Einstein — more like a crossword puzzle than a mathematical proof. A tightly interlocking mesh of reasons (entries) well-anchored in experience (clues) can be a very strong indication of the truth of a claim or theory; that is partly why “scientific evidence” has acquired its honorific use. But where experiential anchoring is iffy, or where background beliefs are fragile or pull in different directions, there will be ambiguity and the potential to mislead.

Of course, the role of an analogy is only to suggest ideas, which then have to stand on their own feet; of course, the usefulness of one analogy by no means precludes the possibility that others will be fruitful too; and of course, an analogy is *only* an analogy. Scientific evidence isn't like a crossword puzzle in every respect: there will be nothing, for example, corresponding to the appearance of a solution in tomorrow's paper; nor (though a seventeenth-century philosopher thinking of scientists as deciphering God's Book of Nature would have thought otherwise) is there a person who designs it. Nor, unlike Kuhn's mildly denigratory talk of normal science as "puzzle-solving," is my use of the crossword analogy intended to convey any suggestion of lightness or of the merely routine. But the analogy will prove a useful guide to some central questions about what makes evidence better or worse.

All of us, in the most ordinary of everyday inquiry, depend on learned perceptual skills like reading, and many of us rely on glasses, contact lenses, or hearing aids; in the sciences, observation is often highly skilled, and often mediated by sophisticated instruments themselves dependent on theory. All of us, in the most ordinary of everyday inquiry, find ourselves reassessing the likely truth of this claim or that as new evidence comes in; scientists must revise their assessments over and over as members of the community make new experiments, conduct new tests, develop new instruments, etc. All of us, in the most ordinary of everyday inquiry, depend on what others tell us; a scientist virtually always relies on results achieved by others, from the sedimented work of earlier generations to the latest efforts of his contemporaries.

This, from a 1996 press report, conveys some idea of just how much "more so" scientific evidence can be:

A recovery team found [a 4.3-pound meteorite, designated ALH84001] in 1984. ... 4 billion years earlier, it was part of the crust of Mars. (Scientists know this because when the rock is heated, it still gives off a mix of gases unique to the Martian atmosphere.) ... From this unprepossessing piece of rock scientists have teased out ... evidence leading toward an astonishing conclusion. Team member Richard Zare, a chemist at Stanford, used lasers and an extremely sensitive detector called a mass spectrometer to

spot molecules called polycyclic aromatic hydrocarbons. PAHs result from combustion; they are found in diesel exhaust and soot. . . . But they also come from the decomposition of living organisms. The residue in ALH84001, says Zare, "very much resembles what you have when organic matter decays." . . . Under another high-tech sensor, an ultra-high-resolution transmission microscope [scientists] found that the thin black-and-white bands at the edge of the carbonates were made of mineral crystals 10 to 100 nanometers across. . . . The crystals in the meteorite were shaped like cubes and teardrops, just like those formed by bacteria on earth. [David MacKay of the Johnson Space Center says] "We have these lines of evidence. None of them in itself is definitive, but taken together the simplest explanation is early Martian life." . . . Some scientists in the field express more optimism than others . . .³

Since then there has been heated controversy over whether or not this really was evidence of early Martian life. In 1998 new chemical studies comparing organic materials in the meteorite with those found in the surrounding Antarctic ice showed that significant amounts of the organic compounds in the meteorite are terrestrial contamination; but these studies didn't examine the crucial molecules, the PAHs. Controversy seems likely to continue at least until new samples of Martian rock and soil can be brought back by robotic spacecraft.⁴

As the example suggests, warrant comes in degrees, and is relative to a time; a scientific idea, usually very speculative at first, tends either to get better warranted, or to be found untenable, as more evidence comes in or is flushed out. As the example also suggests, talk of the degree of warrant of a claim at a time, simpliciter, is shorthand for talk about how warranted the claim is at that time by the evidence possessed by some person or group of people.⁵

Since it is individuals who see, hear, etc., my account begins with the personal conception, the degree of warrant of a claim for a person at a time. The next step, distinguishing a person's experiential evidence and his reasons, and explaining how the two work together, is to articulate what makes a person's evidence with respect to a claim better or worse, and hence what makes the claim more or less warranted for him. Then, to articulate something of what is involved

in evidence-sharing, I shall need to extrapolate from the degree of warrant of a claim for a person at a time to the degree of warrant of a claim for a group of people at a time; and then to suggest an account of the impersonal conception, of the degree of warrant of a claim at a time, simpliciter. Then I will be able to say something about how the concept of warrant relates to the concepts of justification and confirmation; to explain how degree of warrant ideally relates to degree of credence; and to discriminate what is objective, and what perspectival, in the concepts of warrant, justification, and reasonableness.

Because warranted scientific claims and theories are always warranted by somebody's, or somebodies', experience, and somebody's, or somebodies', reasoning, a theory of warrant must begin with the personal, and then move to the social, before it can get to grips with the impersonal sense in which we speak of a well-warranted theory or an ill-founded conjecture. This, obviously, is about as far as it is possible to be from Popper's ideal of an "epistemology without a knowing subject." Ironically enough, however, it is almost as congenial to his analogy of scientific knowledge as like a cathedral built over the centuries by generations of masons, carpenters, glaziers, gargoyle carvers, and so on, as to mine of scientific knowledge as like part of a vast crossword gradually filled in by generations of specialists in anagrams, puns, literary allusions, and so forth.

— 1 —

Warrant — the personal conception

What determines the degree of warrant of a claim for a person at a time is the quality of his evidence with respect to that claim at that time. "His evidence" refers both to his experiential evidence (his seeing, hearing, etc., this or that, and his remembering having seen, heard, etc., this or that — his past experiential evidence), and to his reasons (other beliefs of his). There are significant asymmetries between experiential evidence and reasons, as between clues and crossword entries: most importantly, the question of warrant arises with respect to a person's reasons, as it arises with respect to crossword entries; but perceptual, etc., events and states, like clues to a crossword, neither have nor stand in need of warrant.⁶

Let me take experiential evidence first.

Both in the law and in everyday life, there is a usage in which “evidence” means “physical evidence,” and refers to the actual fingerprints, bitemarks, documents, etc. We hear reports of new evidence about a plane crash brought up from the ocean floor, or of new evidence about a crime discovered in a suspect’s apartment. My account will accommodate this usage, not directly, but in an oblique way, by taking for granted that in scientific observation, as in perception generally, we interact by means of our sensory organs with things around us — with the traces of the gases given off by that Martian meteorite when it is heated, with stuff on the slides under the microscope, with Rosalind Franklin’s X-ray diffraction photographs of DNA, and so on. So in my account, the bits of airplane, the incriminating letter, etc., are the *objects* of experiential evidence, *what* is perceived. A person’s experiential evidence is his perceptually interacting in one way or another — with the naked eye at a distance in poor light, by means of a powerful microscope in good light, etc. — with a thing or event.

Thinking of experiential evidence in science, it is natural to speak, not of perception, but of observation; and here — as when we speak of the observations made by a detective, or of a patient’s being “under observation” in hospital — the word carries a connotation of deliberateness. Scientific observation is active, selective; it calls for talent, skill, and sometimes special training or background knowledge, as well as patience and sharp eyes. Very often it is mediated by instrumentation. Experiential evidence and reasons work together, as the reasonableness of a crossword entry depends in part on its fit with the clue and in part on its fit with intersecting entries. I don’t assume a class of claims (the “observation statements” of some Old Deferentialist accounts) fully warranted by experience alone; rather, I see experiential evidence and reasons as carrying the burden in different proportions for different claims. But neither do I assume that each scientific claim has its own experiential evidence, as in a conventional crossword each entry does; often it is more like an unconventional crossword in which a clump of entries shares a clue, or a bunch of clues.

All this, obviously, takes the relevance of experience to warrant

for granted. So what about Popper's argument for the *irrelevance* of experience — that, since there can be logical relations only among statements, not between statements and events, scientists' seeing, hearing, etc., this or that can have no bearing on the warrant of scientific claims and theories? It is true that logical relations hold only among statements (or whatever the truth-bearers are); but the conclusion Popper draws — that, e.g., someone's seeing a black swan is utterly irrelevant to the reasonableness or otherwise of his accepting the statement that there is a black swan at such-and-such a place at such-and-such a time, and hence to the reasonableness or otherwise of his rejecting the statement that all swans are white — is about as thoroughly implausible as a conclusion could be; so implausible that Popper himself elides it into the quite different thesis I have been defending, that experiential evidence is relevant but not sufficient. This doesn't yet tell us *how* experience contributes to warrant; but it does tell us that the other assumption on which Popper's argument for the irrelevance of experience depends — that warrant is a matter exclusively of logical relations among statements or propositions — must be untrue.

So, how does experience contribute to warrant? A simple answer might rely on the old idea that, while the meanings of many words are learned by verbal definition in terms of other words, the meanings of observational words are learned by ostensive definition, as the language-learner hears the word used by someone pointing out something to which it applies.⁷ So a person's seeing a dog warrants the truth of his belief that there's a dog present in virtue of the fact that "dog" is ostensively defined in such a way as to guarantee that it is appropriate to use it in just such observable circumstances as these. This picture, with its simple division of terms into observational and other, and of definitions into ostensive and verbal, won't do as it stands. Language is far subtler than that, the interconnections of words with observable circumstances and among themselves much more tangled — as the language-learner soon discovers as he masters "toy dog," "looks like a dog," etc., and learns more about what the truth of "it's a dog" requires and what it precludes. Nevertheless, the central idea seems right: our perceptual interactions with the world give some degree of warrant to claims about the world

because of the connections of words with the world and with each other that we learn as we learn language.

Perhaps we can preserve this central idea while remedying the deficiencies of the simple dichotomy of ostensive versus verbal definitions. Even a very simple correction, replacing the dichotomy of observational versus theoretical predicates by a continuum of more and less observational, less and more theoretical, would be an improvement. But it would be better to make room for the possibility of different speakers learning a word in different ways, and of terms that can be learned *either* by a combination of ostention and verbal explanation *or* entirely by verbal explanation. Correcting the simple contrast of ostensive versus verbal definitions, allowing for the tangled mesh of extra- and intra-linguistic connections of words, we could explain both how experiential evidence can contribute to the warrant of a claim, and how the warrant given a claim by a person's experience may be enhanced, or diminished, by his reasons.

We nearly all encounter a sentence like "this is a glass of water,"⁸ in the first instance, by hearing it used in normal circumstances in which a glass of water is visible to both teacher and learner. Subsequently, however, we learn a lot of caveats and complications: a glass of water looks, smells, tastes, etc., thus and so, *provided* the observer and the circumstances of observation are normal; if the stuff in the container is *really* water, it will give such and such results under chemical analysis; etc., etc. So seeing the thing can partially, though only partially, warrant the claim that there's a glass of water present; for a normal observer in normal circumstance can tell it's a glass of water by looking, even though there is room for mistake.

A molecular biologist has to learn to read an X-ray diffraction photograph, as all of us had to learn to read. Someone who had learned the predicate "helix," ostensively or otherwise, by reference to simple examples like a telephone cord, but who had no experience of X-ray diffraction photographs, wouldn't be able to make much of Rosalind Franklin's photograph of the B form of DNA. As soon as James Watson saw it, however, he was firmly convinced that the DNA molecule is helical. And his seeing the photograph partially, but only partially, warranted this claim. For a trained observer in appropriate circumstances can tell it's a helix by looking at a (good

enough) X-ray diffraction photograph, even though there is room for mistake.

In sum: a person's seeing, etc., this or that can contribute to the warrant of a claim when key terms are learned by association with these observable circumstances — the more [the less] so, the more [the less] the meaning of those terms is exhausted by that association. Experiential evidence consists, not of propositions, but of perceptual interactions; and it contributes to warrant, not in virtue of logical relations among propositions, but in virtue of connections between words and world set up in language-learning.

— * * * —

Now let me turn to reasons.

When, earlier, I rather casually referred to a person's reasons as other beliefs of his, I hadn't forgotten that according to some philosophers, among them both Peirce and Popper, belief has no place in science. I agree that faith, in the religious sense, does not belong in science; though in their professional capacity scientists accept various claims as true, this usually is, or should be, tentative, and always in principle revisable in the light of new evidence. By my lights, however, to believe something is to accept it as true, in just this fallibilist sense;⁹ that's why I shall sometimes write of the "degree of credence" a person places in a claim or theory.

Unfortunately, it won't quite do simply to construe a person's reasons as those propositions in which he places some degree of credence, ignoring the fact that some of his beliefs are strongly held and others weakly — any more than it would do, in judging the plausibility of a crossword entry, to ignore the fact that one intersecting entry is written firmly in ink, another only faintly in pencil. If a crossword entry intersecting the entry at issue is only lightly pencilled in, it counts for less, positively or negatively, than if it is indelibly inked in; similarly, if a person gives a reason for or against a claim only a modest degree of credence, it should count for less, positively or negatively, than if he holds it very firmly. One way to accommodate this might be to treat a person who places some but less than complete credence in a proposition as giving full credence to a hedged-version of the same proposition; and so to include among his reasons

something along the lines of “there is a good chance that p ,” “it is likely that p ,” or “there is some chance that p ,” “it is possible that p .” Another would be to include the propositions without the hedges, and compensate somehow by adjusting the degree of warrant of the claim for, or against, which they are reasons. I shall take the latter route.

Unlike his experiential evidence, a person’s reasons are propositional; and so it might seem that here at least we must be squarely in the domain of logic. Not so, however. Reasons ramify, more like the entries in a crossword puzzle than the steps in a mathematical proof. The plausibility of a crossword entry depends not only on how well it fits with the clues and any intersecting entries, but also on how plausible those other entries are, independent of the entry in question, and on how much of the crossword has been completed. Similarly, the quality of a person’s evidence with respect to a claim depends not only on how supportive his reasons are of that claim, but also on how warranted those reasons are, independent of the claim in question, and on how much of the relevant evidence his evidence includes. Moreover, as it turns out not even supportiveness — not even conclusiveness, the limit case of supportiveness — is quite simply a matter of logic.

— * * * —

For reasons to be conclusive with respect to a claim — i.e., to support it to the highest possible degree — it is not sufficient that they deductively imply the claim. For inconsistent propositions deductively imply any proposition whatever (from p and not- p , q follows, whatever q may be);¹⁰ but inconsistent reasons aren’t conclusive evidence for anything, let alone for everything (p and not- p isn’t conclusive evidence for any q , let alone for every q).¹¹ For example, suppose the evidence is: that the murderer is either Smith or Jones; that whoever committed the murder is left-handed; that Smith is right-handed; and that Jones is right-handed. This deductively implies that Jones did it; *and* that Smith did it; *and* that aliens did it. But it is certainly not conclusive evidence for any of these claims, let alone for all of them. However, if the evidence were: that the murderer is either Smith or Jones, that whoever committed the murder is left-

handed, that Smith is right-handed, and that Jones is left-handed, it would be conclusive with respect to the claim that Jones did it. So conclusiveness requires that the evidence deductively imply the claim in question, *but not also its negation*; i.e., that it deductively imply that claim differentially, and not just in virtue of the fact that, being inconsistent, it implies every proposition whatsoever.¹²

The principle that everything follows deductively from a contradiction is a principle of classical logic. So non-classical logicians may object that while the inference from “ p and not- p ” to an arbitrary “ q ” is valid in classical logic, there is a whole range of non-classical systems — paraconsistent logics, relevance logics, connexivist logics, etc., etc. — in which this inference is not valid; and propose that we close the gap between conclusiveness of evidence and deductive implication by resorting to such a logic. I suspect that the motivation for such non-standard systems derives at least in part from a confusion of logical with epistemological issues; but I don’t rule out the possibility that they might shed some light on how inconsistent evidence could, in some circumstances, be better than simply indifferent with respect to supportiveness.¹³

Again, lawyers might object that inconsistent testimony can be extremely informative. Indeed it can; but that witness A says that p , while witness B says that not- p , does not constitute inconsistent evidence in the sense at issue here (i.e., evidence of the form “ p and not- p ”). Granted, a person who is aware of an inconsistency in his evidence with respect to some claim is in something like the position of a lawyer faced with inconsistent testimony; and if he is sensible he will try to identify the background beliefs responsible for the inconsistency, and assess which are better-warranted. Witness A saw the murder from close up, a juror might reason, B only from a distance, so A ’s testimony is likelier to be right; or: A is the defendant’s brother-in-law, while B is a stranger to him, so B has less reason to lie. A scientist who realizes that there is an inconsistency in his evidence may reason in a similar way: “my confidence that DNA is composed of the four nucleotides in regular order is less well-warranted than my confidence that bacterial virulence is contained in nucleic acid, not protein; so of my evidence that DNA is genetic material, and my evidence that it isn’t, the former is likelier to be right.” But this is quite

compatible with my point, which is only that inconsistent evidence is not conclusive evidence.

— * * * —

Against the background of the familiar quarrels between the inductivist and deductivist wings of the Old Deferentialism, it may seem that to acknowledge that there is such a thing as supportive-but-not-conclusive evidence must be to declare allegiance to the inductivist party. Not so, however. There is supportive-but-not-conclusive evidence; but there is no syntactically characterizable inductive logic, for supportiveness is not a purely formal matter.

David Mackay observes that, though the evidence derived from that meteorite is not definitive, “the simplest explanation is early Martian life.” He takes for granted that the evidence so far supports the idea of early Martian life because there having been bacterial life on Mars long ago would explain how things come to be as the evidence says. And, whether or not he is right about bacterial life on Mars, he is right to assume a connection between supportiveness and explanation.

The connection is not, however, simply that evidence supports a claim in virtue of the claim’s being the best explanation of the evidence. Supportiveness of evidence is not categorical, but a matter of degree. That there is a significantly greater incidence of lung cancer in smokers than in non-smokers, for example, supports the claim that smoking causes lung cancer; but the degree of support is very significantly enhanced by additional evidence of specific genetic damage connected to lung cancer and caused by smoking. Moreover, there is “mutual reinforcement between an explanation and what it explains.”¹⁴ In the example just given, the evidence supports the claim in virtue of the claim’s potential to explain the evidence. But the explanatory connection may go either way; in other cases it is a matter, rather, of the evidence potentially explaining the claim. That there is a trough of low pressure moving in a south-easterly direction, for example, supports the claim that Hurricane Floyd will turn north before it reaches the South Florida coast, because there being such a trough of low pressure would explain the hurricane’s turning north. So “inference to the best explanation” is too one-directional, and

captures only a small part of a larger picture in which degree of supportiveness of evidence is tied to degree of explanatory integration of the evidence with the claim in question.

Explanatory integration is a pretty concept, but not easy to spell out. But it is clear, at any rate, that neither explanation nor, a fortiori, explanatory integration or supportiveness of evidence, can be narrowly logical concepts. For explanation, like prediction, requires the classification of things into real kinds. Knowing that geese migrate south as the weather cools, we predict that when the weather gets cooler this goose will fly south, and explain that this goose flew south because the weather got cooler — which is only possible because classifying something as a goose identifies it as of a kind members of which behave thus and so. There is the same covert generality in the previous examples: e.g., if “trough of low pressure” and “hurricane” didn’t pick out real meteorological phenomena connected by real laws, the prediction would be unjustified and the appearance of explanatoriness bogus. Explanatoriness is not a purely logical, but a worldly, concept.

So if we think of supportiveness as a relation among sentences, it will be a vocabulary-sensitive relation, requiring kind-identifying predicates; in other words, it will not be syntactic, a matter of form alone, but broadly semantic, depending on the extensions of the predicates involved. (The point is masked, but not obviated, if we think of supportiveness as, rather, a relation among propositions.) This suggests why scientists so often find themselves obliged to modify the vocabulary of their field, shifting the use of old terms or introducing new ones: a vocabulary can not only be more or less convenient or more or less transparent in meaning, but also — most importantly — more or less successful at identifying kinds of thing, stuff, or phenomena.

— * * * —

How plausible a crossword entry is depends not only on how well it fits with the clue and any already-completed intersecting entries, but also on how plausible those other entries are, independent of the entry in question, and how much of the crossword has been completed. Analogously, the degree of warrant of a claim for a person

at a time depends not only on how supportive his evidence is, but also on how comprehensive it is, and on how secure his reasons are, independent of the claim itself.

A person's evidence is better evidence with respect to a claim, the more [less] warranted his reasons for [against] the claim in question are, independently of any support given them by that claim itself. This clause suggests a possible solution to the problem of weakly-believed reasons. By including among a person's reasons all those propositions he gives some degree of credence, I avoided the difficulty of accommodating hedged propositions like "possibly *p*," or "maybe *q*" into my account of explanatory integration; but didn't I thereby give partially-believed reasons more weight than I should? No: the unhedged "*p*" or "*q*" included as proxy will be less independently secure than the hedged "possibly *p*" or "maybe *q*" that would accurately represent the person's degree of credence; and so the independent security clause will compensate for my including the stronger proposition by ensuring that a reason for [against] a claim contributes less to its warrant, or lowers its warrant less the less firmly it is held.

Although the independent security clause mentions warrant, there is no vicious circularity. In a crossword, the reasonableness of an entry depends in part on its fit with other entries, and hence on how reasonable they are, independent of the entry in question. Similarly, the warrant of a claim depends in part on the warrant of other claims that support it, independent of any support given to them by the claim itself. This interlocking of mutually supportive claims and theories no more conceals a vicious circle than the interlocking of mutually supportive crossword entries does. Nor does it threaten us with an infinite regress, or leave the whole mesh floating in mid-air; for experiential evidence, which stands in no need of warrant, serves as anchor for scientific claims, as clues do for crossword entries.

The quality of a person's evidence, and hence the degree of warrant of a claim for him, also depends on how much of the relevant evidence his evidence includes. Comprehensiveness is one of the determinants of evidential quality, not an afterthought to be relegated to methodology. (Stating the comprehensiveness requirement precisely, however, would call for an extension of the conception of

evidence on which I have relied thus far; for in this context “all the relevant evidence” has to mean something like “answers to all the relevant questions.”) Even if it strongly supports the claim in question, even if it is very secure itself, evidence is poorer insofar as relevant information is missing. Weakness on the dimension of comprehensiveness is apt to make evidence misleading, i.e., supportive of a false conclusion; and since the evidence with respect to a scientific claim is never absolutely comprehensive, there is always the possibility that, as new evidence comes in, the evidence so far will turn out to have been misleading.

Because the determinants of evidential quality are multi-dimensional, and one claim may do well on one dimension and another on another, there is no guarantee of a linear ordering of rival claims with respect to degrees of warrant. Moreover, the three dimensions interact. Evidence which is poor on the dimension of comprehensiveness is often also poor on the dimension of supportiveness; while evidence which is highly supportive of a claim is often rather lacking in independent security.

This sheds some light on an old disagreement about the status of negative evidence. Popper’s thesis that scientific claims can be falsified, but not verified or confirmed, derives in part from his criterion of demarcation (equating “scientific” with “falsifiable”), and in part from the assumption that a single negative instance falsifies a generalization. The position associated with Quine and Duhem, by contrast, is that scientific laws and generalizations are no more decisively falsifiable than they are decisively verifiable. The demarcation issue aside, there is disagreement about the evidential relation of negative instances to generalizations. Popper is right, of course, that negative evidence needs to be taken into account; and right that, for instance, that there is a black swan at Perth airport at such and such a time is a conclusive reason against the generalization that all swans are white. But his critics are right to insist that this doesn’t settle the matter. Conclusiveness is the highest degree of supportiveness; but supportiveness is only one dimension of evidential quality. It does not guarantee decisiveness, which would require in addition that the conclusive reason be perfectly independently secure, and that it be all the relevant evidence.¹⁵

— 2 —

Warrant — the social conception

Now let me turn to the warrant of a claim for a group of people.¹⁶

In 1954 George Gamow set up the RNA Tie Club, a group of 20 people — one for each amino acid — devoted to figuring out the structure of RNA and the way it builds proteins. Each member was to have a black RNA tie embroidered with a green sugar-phosphate chain and yellow purines and pyrimidines, and a club tie-pin carrying the three-letter abbreviation of his assigned amino acid; later there was even RNA Tie Club stationery, with a list of officers (“Geo Gamow, Synthesizer, Jim Watson, Optimist, Francis Crick, Pessimist, . . .”).¹⁷ Very few scientific communities, however, are as definitely identifiable as this; the notion of a scientific community is notoriously vague, and specifying criteria for what is to count as a scientific community, let alone for what is to count as one such community, is a formidably difficult task.

In fact, “the” scientific community to which philosophers of science sometimes optimistically refer is probably more mythical than real; the reality is a constantly shifting congeries of sub-communities, some tightly interconnected and some loosely, some nested and some overlapping, some short-lived and some persisting through several generations of workers. So it is just as well that I can sidestep the awkward problems about the individuation of scientific communities and sub-communities, because my present task is to specify on what the degree of warrant of a claim depends for *any* collection of scientists, whether that collection is a closely-knit sub-community or a scattered or gerrymandered group.

“One man’s experience is nothing if it stands alone,” wrote C. S. Peirce,¹⁸ who saw its engagement of many people, within and across generations, as one of the great strengths of the scientific enterprise. He was right; and not least because this enables the sciences to extend their evidential reach far beyond that of any individual. But it is not an unmixed blessing. For in any group of scientists there will likely be disagreements both about the claim the warrant of which is at issue, and about the reasons for or against that claim; experiential evidence, furthermore, is always some individual’s experi-

ential evidence; and in even the most tightly-knit group of scientists there will be failures of communication, with each member having only imperfect access to others' evidence.

Given that different individuals within a group of scientists may disagree not only in the degree of credence they give the claim in question, but also in their background beliefs, we can't construe the group's evidence as a simple sum of all the members' evidence. But the crossword analogy suggests a way to overcome this first difficulty. Think of several people working on the same crossword, agreeing that 2 down is "egregious," and 3 across "gigantic," but disagreeing about 4 across, which some think is "intent," and others think is "intern." What would determine how reasonable, given the evidence possessed by this group of people, an entry which depends on 4 across is? Presumably, how reasonable it is if the disputed entry is *either* "intent" or "intern" (or equivalently, since the rival entries agree in their first letters, if the last letters are either "nt" or "rn.") Similarly, where there is disagreement in background beliefs within a scientific community, the best approach may be to construe the group's evidence as including not the conjunction of the rival background beliefs, but their disjunction. However, this one-size-fits-all solution will need considerable adjustment to accommodate disagreements of different shapes and sizes: the community may, for example, be more or less evenly divided, or there may be just one dissenter.

It is always an individual person who sees, hears, remembers, etc. In scientific work, however, many people may make observations of the same thing or event; of an eclipse from observatories in the northern and in the southern hemispheres, for example. By observing the same thing or event from different places, scientists have access to more of the information the thing or event affords. And by having several people make the same observation, they can discriminate the eccentricities of a particular individual's perceptions from what can be perceived by all normal observers. Sometimes one person claims to be able to see what no-one else can: all the observations supposedly confirming that a homeopathic dilution of bee-venom degranulates blood cells, apparently, were made by one observer, Elisabeth Davenas. In such circumstances, either the person involved is an especially talented observer (as Jacques Benveniste maintains

Mlle. Davenas is), or else he or she is, as we say, “seeing things” (as John Maddox and the team he sent from *Nature* to investigate the work of Benveniste’s lab maintain Mlle. Davenas must be).¹⁹

In relying on others’ observations, scientists depend on those others’ perceptual competence, on the working of the instruments on which they rely, and on the honesty and accuracy of their reports. It is a matter, not simply of mutual trust, but of justified mutual confidence (usually grounded implicitly in the observer’s, or his instrument’s, credentials). Scientists will reasonably take into account that an observer’s commitment to this or that theory may make him reader to notice some aspects of what he or she sees than others; and if they have grounds for suspecting the observer of perceptual defect, instrumental failure, dishonesty, or self-deception — whether directly or, as with those homeopathy experiments, because the supposed results are so extraordinary — they may reasonably doubt the reliability of his or her observational reports. In a group of scientists, even if each has his own experiential evidence, most depend at second hand on others’. So the warrant of a claim for the group will depend in part on how reasonable each member’s confidence is in others’ reports of their observations; and in part (now I turn to the third difficulty mentioned earlier) on how good communication is within the group.

It hardly seems appropriate to allow that a claim is warranted for a group in which evidence is not shared, but merely scattered: as with two scientists centuries apart, the later quite unaware of the work of the earlier, or with rival research teams neither of which has ever seen the other’s reports. We would not count a claim as well warranted for a group of people, even if between them they possess strong evidence for it, unless that evidence is communicated among the members of the group. Only when their evidence is shared — as when the several people working on the same part of the crossword puzzle are all able to look over the others’ shoulders — can their joint evidence warrant a claim. “Efficiency of communication” covers a whole range of issues: how effectively refereeing and publishing processes ensure that good work is published quickly, and not drowned in a sea of worthless busywork; how good the means are of finding relevant material; how far conferences manage to be occa-

sions for genuine communication and mutual education rather than mere self-promotion and networking; how cogently and clearly work is presented.

So we could think of the degree of warrant of a claim for a group of scientists as the degree of warrant of that claim for a hypothetical individual whose evidence is the joint evidence of all the members of the group, only construed as including not the conjunctions but the disjunctions of disputed reasons, and discounted by some measure of the degree to which each member is justified in believing that others are reliable and trustworthy observers, and of the efficiency or inefficiency of communication within the group.

— 3 —

Warrant — the impersonal conception

Now I can say something about the impersonal conception, of the degree of warrant of a claim at a time, simpliciter.

When, looking at science from the outside, you wonder which claims and theories are well and which poorly warranted, it is this impersonal conception which is most salient. But to say that a claim or theory is well or poorly warranted at a time must be understood as an elliptical way of saying that it is well or poorly warranted by the evidence possessed by some person or some group of people at that time. And since a claim may be well-warranted for this group or person, but poorly warranted for that group or person, the question is on *whose* evidence “impersonal” warrant is appropriately taken implicitly to depend.

I shall construe it as depending on the evidence of the person or group of people whose evidence is, in a certain sense, the best. “In a certain sense,” because in this context, “best evidence,” means “best indicator of the likely truth of the claim or theory in question.” This is close to the legal conception of “best evidence,”²⁰ and should not be confused with “best evidence” in the sense of “evidence which gives the highest degree of warrant to the claim or theory in question.” The difference is that in the sense relevant here what matters is how secure and how comprehensive the evidence is, *whether it is supportive or undermining, favorable or unfavorable* to the truth of the claim in question.

When there is efficient communication within a group, the group's shared evidence may be better than that of any individual member; but when communication within the group is poor, an individual's evidence may be the best. For example, after trying unsuccessfully to interest Karl Nägeli in his work, Gregor Mendel published his "Experiments in Plant Hybridization" in the journal of the Society of Natural Science in Brunn, Moravia, where it languished unread for decades; so for a time his evidence with respect to the particulate theory of inheritance was much better than anyone else's — in both senses. The theory was more warranted for him than for anyone else, but his was also the best evidence in the sense at issue here: i.e., the best indicator of the likely truth of the theory. It was also, in the relevant sense, the best evidence with respect to the rival, blending theory; to which, however, it was unfavorable.

Again: in 1944 Oswald Avery wasn't ready to say in his scientific publications, as he suggested in a letter to his brother, that his experiments indicated that DNA, not protein, is the genetic material. So for a time his evidence with respect to the protein-as-genetic-material hypothesis was, in the sense presently at issue, better than anyone else's: i.e., the best indicator of the likely truth of the hypothesis — but unfavorable. His was also, in the relevant sense, the best evidence with respect to the rival DNA-as-genetic-material hypothesis; to which, however, it was favorable. Much later, however, in his textbook on molecular biology, Watson would refer to "Avery's bombshell," the "first real proof of the genetic role for nucleic acids";²¹ and in the Rockefeller University calendar for the academic year 2001, beside a rather charming photograph of Avery in a funny hat at Rockefeller's 1940 Christmas party, the text informs us that it was Avery, with his colleagues Colin MacLeod and Maclyn McCarty, who "showed for the first time that genes are made of DNA."

— 4 —

Warrant, justification, and confirmation

Now I can tackle the question of the relation of warrant to such other concepts as justification and confirmation.

For a claim to be warranted to some degree, I shall require (not that the evidence indicate that the claim is more likely than not, but) only that the evidence indicate that the claim is non-negligibly likely. The claim that p is well warranted for an individual if his evidence strongly indicates that p ; the claim is fairly warranted for him if his evidence fairly strongly indicates that p ; it is weakly warranted for him if his evidence weakly indicates that p ; and it is unwarranted for him if his evidence does not indicate that p — whether because it indicates that not- p , or because it is too impoverished even weakly to indicate *either* p or not- p .

That a claim is highly warranted for a person doesn't guarantee that he is in good epistemic shape with respect to that claim. A scientist may accept a claim, with greater or lesser confidence, as true; or accept its negation, with greater or lesser confidence, as true; or give no credence either to the claim or to its negation. Ideally, he would give p the degree of credence it deserves. But he may fall short of this ideal either because p is well-warranted for him, but he gives it too little credence, or because p is poorly warranted for him, but he gives it too much credence. These failings may be described, respectively, as under-belief and over-belief.

Moreover, it may not be the evidence a scientist possesses that moves him to give a claim whatever degree of credence he does. He may give some degree of credence to a claim because he is impressed by the fact that an influential figure in his profession has endorsed it, or because he very much wants things to be as the claim says, or, etc. In such a case I shall say that, even if the claim is warranted for him, he is not justified in giving it the degree of credence he does. (Note to Karl Popper: justification, in the sense just explained, is a partly causal notion; and experiential evidence can contribute to the justification of a person's belief precisely by contributing causally to his accepting it.)

At any time, some scientific claims and theories are well warranted; others are warranted poorly, if at all; and many lie somewhere in between. Sometimes several competing claims may all be warranted to some degree. When no-one has good enough evidence either way, a claim and its negation may be both unwarranted; in which case, the best option is — admitting that at the moment we

just don't know — to seek out more evidence, and to rack our brains for other candidate hypotheses.

Most scientific claims and theories start out as informed but highly speculative conjectures; some seem for a while to be close to certain, and then turn out to have been wrong after all; a few seem for a while to be out of the running, and then turn out to have been right after all. Many, eventually, are seen to have been right in part, but also wrong in part. Some mutate, shifting in content to stand up to new evidence in an adapted form. Ideally, the degree of credence given a claim by the relevant scientific sub-community at a time — assuming we can give some sense to this not entirely straightforward idea — would be appropriately correlated with the degree of warrant of the claim at that time. The processes by which a scientific community collects, sifts, and weighs evidence are fallible and imperfect, so the ideal is by no means always achieved; but they are good enough that it is a reasonable bet that much of the science in the textbooks is right, while only a fraction of today's frontier science will survive, and most will eventually turn out to have been mistaken. Only a reasonable bet, however; all the stuff in the textbooks was once speculative frontier science, and textbook science sometimes turns out to be embarrassingly wrong.

— * * * —

I shall say that a claim is confirmed when additional evidence raises its degree of warrant, the degree of confirmation depending on the increment of warrant. Thus construed, the concept of confirmation is not only distinct from, but presupposes, the concepts of warrant and supportiveness. Some Old Deferentialists, however, used "confirm" indifferently for supportiveness, warrant, and confirmation. The confusions such ambiguities generate linger on in, for example, the still-common idea that evidence already possessed at the time a theory was proposed cannot support it; which seems plausible only if supportiveness and confirmation are run together.²²

In my usage, we can describe evidence as confirming a claim (i) when new evidence, i.e., evidence not previously possessed by anyone, raises its degree of warrant; (ii) retrospectively, when a claim

previously already warranted to some degree became more warranted when such-and-such then new, but now familiar, evidence came in; or (iii) to assess the degree of warrant of a claim first relative to such-and-such evidence, and then relative to that evidence plus additional evidence not previously included in the reckoning.

This suggests a way of approaching an old controversy about whether true predictions are especially confirmatory. On the one hand, it is certainly impressive when astronomers predict that Halley's comet will reappear or that the sun will be eclipsed at such and such a future time, and turn out to be right. On the other hand, it is certainly puzzling how the fact that a statement is about the future, in and of itself, could endow it with any special epistemological importance. The explanation is that, though the intuition that successful prediction can be strongly confirmatory is correct, the reason is not simply that it is a true prediction. Verification of a prediction derived from a claim is always new evidence, in the sense required by (i) or (retrospectively) by (ii). However, new evidence may concern past events, and not only future ones; e.g., if an astronomical calculation has the consequence that there was a solar eclipse at such and such a time in ancient history, and then new evidence is found that in fact there was, this true "postdiction" confirms the theory no less than a true prediction would do. Moreover, additional evidence in the sense of (iii), even if it is not new evidence in the sense of (i) or (ii), may also be confirmatory.

Thus far, I have said only that confirmatory evidence raises the degree of warrant of a claim. In ordinary usage, however, "confirm" is quite often used comparatively, to indicate that additional evidence warrants p over some rival q . We might say that additional evidence which raises the degree of warrant of p but lowers the degree of warrant of q "confirms p over q ." In ordinary usage, again, "confirm" also often carries a suggestion that the claim confirmed is now not merely *more* warranted, but *firmly* warranted. We might say that additional evidence that raises the degree of warrant of p beyond some specified cut-off point is "strongly confirmatory."

You may have noticed that though I have talked in terms of degrees of credence, degrees of warrant, and degrees of confirma-

tion, and occasionally of likelihoods, I have thus far rather pointedly avoided "probable." By now, the reason should be pretty obvious: the classical calculus of probabilities, originally devised to represent the mathematics of games of chance, looks like a poor match for degrees of warrant. It could hardly constitute a theory of warrant, if this concept is as subtle and complex as it seems to be. Nor could it constitute a calculus of degrees of warrant; for the probability of p and the probability of not- p must add up to 1, but if there is insufficient evidence either way, neither a claim nor its negation may be warranted to any degree. For example, scientists now believe that mad-cow disease is caused by prions, protein molecules abnormally folded up in the cell,²³ but neither this claim nor its negation was even intelligible until the concept of macromolecule was developed, and neither was warranted to any degree until the significance of the folding of macromolecules began to be understood, and mad-cow disease was identified.

Naturally, given my reservations about probabilism generally, I am disinclined towards Bayesianism specifically (nor have I forgotten that even so determined a probabilist as Carnap warns of the dangers of putting too much epistemological weight on Bayes' theorem). Of course, there's nothing wrong with the theorem itself, *qua* theorem of the calculus of probabilities; and presumably, when they engage in statistical reasoning, scientists sometimes calculate probabilities in a Bayesian way. However, as even the most enthusiastic Bayesians acknowledge, degrees of credence, construed purely descriptively, need not satisfy the axioms of the calculus of probabilities; they may not be coherent. And if, as I argued above, degrees of warrant need not satisfy the axioms of that calculus either, then there is good reason (over and above familiar worries about where the priors come from) for denying that Bayes' theorem could be an adequate model of scientists' readjustments of degrees of warrant in the light of new evidence.

— * * * —

Complex and diffuse as it is, evidence is a real constraint on science. And though the degree of warrant of a claim at a time depends on the quality of some person's or some group's evidence at that time,

the quality of evidence is not subjective or community-relative, but objective.

However, it doesn't follow from the objectivity of evidential quality that it is transparent to us. In fact, judgments of the quality of evidence depend on the background beliefs of the person making the judgment; they are perspectival. If you and I are working on the same crossword, but have filled in the much-intersected 4 down differently, we will disagree about whether the fact that an entry to 12 across ends in an "F," or the fact that it ends in a "T," makes it plausible. If you and I are on the same hiring committee, but you believe that handwriting is an indication of character while I think that's all nonsense, we will disagree about whether the fact that a candidate loops his fs is relevant to whether he should be hired — though whether it is relevant depends on whether it is *true* that handwriting is an indication of character.

Quite generally, a person's judgments of the relevance of evidence, and hence of how comprehensive this evidence is, or of how well this claim explains those phenomena, and hence of how supportive it is, are bound to depend on his background assumptions. If he thinks fur color is likely to vary depending on climatic conditions, he will think it relevant to a generalization about the varieties of bear whether the evidence includes observations from the Arctic and the Antarctic; if he thinks the structure as well as the composition of a molecule determines how it functions, he will insist on asking, as Roger Kornberg reports that with the advent of structural chemistry molecular biologists began to do, "[h]ow do you do it with nuts and bolts; how do you do it with squares and blocks and the sorts of thing that we know molecules are made of?".²⁴ And so on.

When there are serious differences in background beliefs between one group of scientists and another, there will be disagreement even about what evidence is relevant to what, and about what constitutes an explanation — disagreements that will be resolved only if and when the underlying questions are resolved (or which may, as Max Planck famously observed, just fade away as the supporters of one side to the dispute retire or die off).²⁵ What has been taken for paradigm-relativity of evidential quality is a kind of epistemological illusion;

again as in the graphology example, whether evidence is relevant, whether this is a good explanation of that, how strong or weak this evidence really is, how well or poorly warranted this claim actually is, is an objective matter.

Sometimes scientists know that they don't have all the evidence relevant to a question; and sometimes they have a pretty shrewd idea what the evidence is that they need but don't have. But sometimes, given the evidence they have, they may be unable to judge, or may misjudge, whether or what additional evidence is needed. They can't always know what it is that they don't know; they may not, at a given time, even have the vocabulary to ask the questions answers to which would be relevant evidence. Nor can they always envision alternative hypotheses which, if they did occur to them, would prompt them to revise their estimates of the supportiveness of their evidence. And so on. Since evidential quality is not transparent, and scientists can only do the best they can do, a scientist may be reasonable in giving a claim a degree of credence which is disproportionate to the real, objective quality of his evidence, if that real quality is inaccessible to him. Reasonableness, so understood, is perspectival.

— * * * —

This is still quite far from a fully detailed account of the nature and structure of scientific evidence and of what makes it stronger or weaker — more like a preliminary analysis of the chemical composition of those concepts than the detailed account of their molecular structure I would ideally like. How, more specifically, do observation and background beliefs work together? What, more specifically, is involved in such concepts as explanatory integration and kinds, to which I have thus far simply helped myself? What does my long story about evidence and warrant have to do with the truth of scientific claims, or with progress in science? And what does any of this have to do with the traditional preoccupation with “scientific method”? I'm working on it.

Susan Haack
College of Arts and Sciences
Department of Philosophy
University of Miami
P.O. Box 248054, Coral Gables
Florida 33124-4670, USA
s.haack@miami.edu

Keywords

Scientific evidence; experimental evidence vs. reasons; warrant; justification; confirmation

Notes

¹ Einstein, Albert, "Physics and Reality," *The Journal of the Franklin Institute*, 221.3, 1936; reprinted in *Ideas and Opinions of Albert Einstein*, translated by Sonja Bargmann, Crown Publishers, New York, 1954, 290-323 (the quotation is from p. 295 of *Ideas and Opinions*).

² On the distinction between the Old Deferentialism and the New Cynicism in philosophy of science, see my "Knowledge and Propaganda: Reflections of an Old Feminist" (1993), reprinted in *Manifesto of a Passionate Moderate: Unfashionable Essays*, University of Chicago Press, Chicago, IL, 1998, 123-36; "Puzzling Out Science" (1995), *Manifesto*, 90-103; "Science as Social? Yes and No" (1996), *Manifesto*, 104-22; and "Defending Science — Within Reason," *Principia*, 3.2, 1999, 187-211.

³ Rogers, Adam, "Come In, Mars," *Newsweek*, October 20th, 1996, pp. 56-7.

⁴ Wilford, John Noble, "2 New Chemical Studies Find Meteorite Samples Show No Traces of Past Life on Mars," *New York Times*, January 1st 1998, p. A22.

⁵ Compare Russell: "individual percepts are the basis of all our knowledge, and no method exists by which we can begin with data. . . public to many observers" (*Human Knowledge, Its Scope and Limits*, Simon and Schuster, New York, 1948, p. 8).

⁶ This sentence must be read in the spirit of the foundherentist account offered in my *Evidence and Inquiry: Towards Reconstruction in Epistemology*, Blackwell, Oxford, 1993, chapter 4; not as saying, as a foundationalist would, that experiential beliefs support, but are not supported by, other beliefs.

⁷ Something like this is suggested by Russell in *Human Knowledge*, pp. 4, 63ff. 501, 502. Only "something like this," however; Russell is at pains to point out that ostensive definition leaves room for differences in the meaning attached to a word by one individual and by another.

⁸ The example comes from Popper, from a passage where, rather than denying the relevance of experience, he is insisting on the fallibility of "basic statements."

⁹ Both Popper and Van Fraassen, in different ways, distinguish between belief and acceptance; I do not.

¹⁰ Proof: From " $p \ \& \ \text{not-}p$ " it follows that p . From " p " it follows that p or q . From " $p \ \& \ \text{not-}p$ " it also follows that $\text{not-}p$. From " $\text{not-}p$ " and " p or q " it follows that q . QED.

¹¹ Hempel notices the problem; he suggests that contradictory observation reports could be excluded "by a slight restriction of the definition of 'observation report'," but that "there is no important reason to do so" ("Studies in the Logic of Confirmation" (1945), in *Aspects of Scientific Explanation*, 3–47, p. 31, note 48). Carnap discusses the problem in a note entitled "Remarks on the exclusion of *L*-false evidence," *Logical Foundations of Probability*, University of Chicago Press, Chicago, IL, 1950, pp. 295–6 in the second, 1962 edition.

¹² Perhaps it would be desirable to add, as a precaution against a potential parallel difficulty in case the claim in question is necessarily true, that conclusiveness requires that this evidence, *but not any evidence whatsoever*, deductively imply the claim in question. But I shall set these complications aside.

¹³ See Routley, Richard [now Richard Sylvan], R. K. Meyer, V. Plumwood, and R. T. Bradley, *Relevant Logics and Their Rivals*, Ridgeview, Atascadero, CA, 1982 (on relevance, paraconsistent, etc., logics); Haack, *Philosophy of Logics*, Cambridge University Press, Cambridge, 1978, pp. 197–203, and *Evidence and Inquiry*, pp. 83–4. I was glad to find that in *Conceptual Revolutions*, Princeton University Press, Princeton, NJ, 1992, Paul Thagard, independently, takes a line quite like mine.

¹⁴ Quine, W. V., and Joseph Ullian, *The Web of Belief*, Random House, New York, 1978, p. 79.

¹⁵ Though I arrived at it independently, this is, I now realize, in essence the diagnosis Hempel gave (in his rather confusing vocabulary of "absolute" versus "relative" verification) in 1945, in section 10 of "Studies in the Logic of Confirmation." See also Hesse, Mary, *The Structure of Scientific Inference*, University of California Press, Berkeley and Los Angeles, 1974, pp. 130–31.

¹⁶ Like Philip Kitcher (*The Advancement of Science: Science Without Legend, Objectivity Without Illusions*, Oxford University Press, New York, 1993, chapter 8), I take the epistemology of science to be social in a relatively conservative sense, as involving interactions among individuals.

¹⁷ My source is Judson, Horace Freeland, *The Eighth Day of Creation: Makers of the Revolution in Biology*, Simon and Schuster, New York, 1979, pp. 264–5.

¹⁸ *Collected Papers*, eds Hartshorne, Charles, Paul Weiss, and Arthur Burks, Harvard University Press, Cambridge, MA, 1931–58, 5.402, second footnote (the reference is by volume and paragraph number).

¹⁹ Benveniste, Jacques, *et al.*, “Human Basophil Degranulation Triggered by Very Dilute Antiserum Against IgE,” *Nature*, 330, June 30th, 1988, 816–8; Maddox, John, *et al.*, “‘High-Dilution’ Experiments an Illusion,” *Nature*, 334, July 28th, 1988, 287–90.

²⁰ “[T]he best evidence principle ... expresses the obligation of litigants to provide evidence that will best facilitate this central task of accurately resolving disputed issues of fact”: Nance, Dale, “The Best Evidence Principle,” *Iowa Law Review*, 173.2, January 1988, 227–97, p. 233.

²¹ Watson, James D., *Molecular Biology of the Gene*, W. A. Benjamin, inc., Menlo Park, CA, 3rd edition, 1976, p. 52.

²² See, for example, Laudan, Larry, “A Critique of Underdetermination” (adapted from his “Demystifying Underdetermination,” 1990) in Klee, Robert, ed., *Scientific Inquiry: Readings in the Philosophy of Science*, Oxford University Press, New York, 1999, 83–99, p. 91. See Mayo, *Error and the Growth of Experimental Knowledge*, University of Chicago Press, Chicago, IL, 1996, pp. 206ff., for a discussion of this idea in the context of Popper’s conception of severity of tests.

²³ Cowley, Geoffrey, “Cannibals to Cows: The Path of a Deadly Disease,” *Newsweek*, March 12th, 2001, 53–61, p. 54.

²⁴ Judson, *The Eighth Day of Creation*, p. 495.

²⁵ Planck, Max, *Scientific Autobiography and Other Papers* (1949), Greenwood Press, New York, 1968, pp. 33–4.