SCIENTIFIC EVIDENCE

Abstract

This chapter introduces the evidence question in philosophy of science: what *is* scientific evidence? It considers the empiricist conception of observation, and its history, in relation to both evidence and perception, as well as the challenge from the theory-ladenness of observation. It looks also at the idea of theoretical evidence and non-empiricist conceptions of observation and evidence.

1 Introduction—the evidence question in science

Science is an epistemic enterprise in which we judge the truth (or falsity or probability) of a hypothesis on the basis of an inference from evidence. To understand science, therefore, we need to take an interest in the character of the hypotheses it infers, the nature of the evidence from which it makes the inferences, and the kinds of inference it employs. Of these three components, it is the last, the nature of the inferences used by science, that has attracted most attention from philosophers of science. The theory of confirmation aims both to describe, in general terms, the inferences of science, and to explain why (or, at least, examine whether) they can provide a rational basis, given the evidence, for our doxastic attitudes towards the theories of science. The theory of confirmation thus tells us what it is for some proposition to be *evidence for* a hypothesis. (The theory of confirmation is addressed in chapters on evidential support and paradoxes of confirmation in this volume. See Achinstein (2001) for a survey as well as a detailed study of the confirmation question.)

Rather less attention has been given by philosophers of science to the nature of evidence itself. That is, the theory of confirmation has occupied philosophers' energy to the exclusion of the evidence question: which propositions may be used as evidence in a scientific inference? This question is nonetheless of crucial importance. For a correct theory of confirmation cannot tell us whether a hypothesis should rationally be believed unless we also know which propositions may legitimately be used as evidence. (Assuming

a plausible theory of confirmation, then trivially, if any proposition were allowed to be evidence, then every hypothesis would be confirmed by the evidence, and if no proposition were allowed to be evidence, then no hypothesis would be confirmed by the evidence.)

The importance of the evidence question is illustrated by the realism debate in the philosophy of science. The standard view of evidence, shared by most realists and anti-realists alike, is that evidence in science is observational, and that to observe something in science is to perceive it. (Whether and to what extent the perception involved in observation can be mediated by an instrument is disputed.) Theories in science concern the unobservable. If scientific inference is to allow us to have a positive doxastic attitude to a theory (such as believing it to be true), then it must show how a rational inference can bridge this gap between the observable (i.e. perceptible) and the unobservable. For example, such an inference must bridge the gap between what we can see in the laboratory and the existence of atoms or magnetic fields and such like. The realist and anti-realist will disagree about whether any rational inference procedure can do this. The typical realist claims that inference to the best explanation can bridge this gulf, because what we see can be best explained by the existence and behaviour of things we cannot see. For example, the realist will say that the best explanation of Brownian motion (the perceptible rapid vibrations of dust particles on the surface of water) is that water is composed of (imperceptible) molecules that are themselves in motion. Whereas the anti-realist will deny that this is a rational form of inference. Hence one major form of the realism debate centres on the theory of confirmation: can inference to the best explanation (or other rational inference) show how propositions concerning only the perceptible can confirm a theory concerning the imperceptible?

With this conception of evidence the realist faces a considerable challenge. However, if this shared assumption about evidence were wrong then matters look rather different. Scientists themselves do not use 'evidence' in a way that conforms to that assumption. They frequently publish papers discussing observations of atoms, quarks, magnetic fields, and any number of other imperceptible entities and processes. If this is a legitimate and relevant use of the concept of observation, and if such observations are evidence (as the scientists take them to be), then the challenge to the realist looks to be rather less daunting, even if it has not disappeared altogether. If some of our evidence already concerns imperceptible entities, then our ability to infer facts about other imperceptible entities looks less questionable.

The lesson of this brief discussion of the realism debate is that the nature of evidence in science is an important but largely neglected topic. It is crucial to

key debates about the possibility of scientific knowledge. Yet philosophers of science have preferred to focus on the confirmation question instead, taking for granted an empiricist conception of observation and therefore of evidence. In this survey I shall examine this empiricist conception of observation and then look at the question of the theory-ladenness of observation to which it gives rise. In so doing I assume that evidence in science is often observational. So I also ask whether *all* evidence in science is observational. In the final section, I look at alternatives to the empiricist conception of observation.

I shall also assume a broad functionalist characterization of the concept of evidence: evidence propositions are those that when used as premises in inferences, enable those inferences to achieve our proper epistemic goals. (One might extend the notion of 'inference' so that this characterization includes other kinds of doxastic sensitivity to a content.) It is commonly thought, for example, that S's evidence consists of those propositions believed by S that can be used by S to justify her other beliefs (Conee and Feldman 2004). However, if one thinks that the S's aim *qua* believer should be knowledge then one might think that S's evidence consists in those propositions that S can use, by inference, to generate knowledge (Bird 2018).

2 The empiricist conception of observation

As we have seen, philosophers of science take observation to be *the* source of evidence in science (or at very least the principal source of evidence). Furthermore they typically have an empiricist conception of what that observational evidence is: observation is essentially perceptual in nature. Chakravartty (2007: 3–4) writes 'There are things that one can, under favourable circumstances, perceive with one's unaided senses. Let us call them "observables"'. Van Fraassen denies that particles in a cloud chamber are being observed because they cannot be seen (van Fraassen 1980: 17). For van Fraassen, 'Observation is perception, and perception is something that is possible for us without instruments' (van Fraassen 2001: 154). Realist philosophers of science are likewise disposed to regard imperceptible entities as unobservable: Papineau (1978: 17–18) writes 'The modern physical sciences …seem to have discovered all kinds of perfectly real but unobservable entities, such as magnetic fields, radio waves, and subatomic particles.'

Among earlier empiricists the 'phenomenal conception of observation' predominated. (Compare what Williamson (2000: 173) call the 'phenomenal conception of evidence'.) Mill (1843), for example, argued that what is typically called 'observation' is not strictly the observation at all. For our

observation *reports* are usually inferences from perceptions, inferences that are 'performed with the speed and certainty of an instinct' so that we cannot distinguish them from the perceptions themselves. In truth, '[o]bservation extends only to the sensations by which we recognise objects; but the propositions which we make use of, either in science or in common life, relate mostly to the objects themselves' (Mill 1843: 643). Mill also remarks that the inferred description is liable to error: it is the uninferred basic perception that is certain, and this may indeed be part of the motivation for his view. According to his view, evidence should be certain; observation provides this certain evidence, but only if an observation is the perceiver's inner sensation of things, not a fact about the thing itself.

Mill's distinction between the true observation (the sensations or perceptions) and the seeming observations (how we describe the things we experience after inference) is a major part of what was at issue in the 'protocol sentence' controversy of the Vienna Circle. The latter used two terms '*Beobachtungsatz*' (observation sentence—Schlick's preferred term) and '*Protokollsatz*' (protocol sentence—Neurath's term) interchangeably to denote evidence reports. Members of the Vienna Circle disputed the appropriate content for such reports and as a result also disagreed about the exact structure of the justification provided by such reports. (For a detailed examination of the controversy see Uebel (2007).)

The foundationalist Schlick (1934) promoted a phenomenalist conception of the protocol sentences, which reported the subjective sense-experiences of the subject (à la Mill). Such a sentence might report 'here, now, white', which, Schlick held, is incorrigible because 'white' expresses the subject's experience (rather than an objective colour property).

Carnap and Neurath rejected Schlick's conception of protocol sentences on the ground that a connection cannot be made between the protocol sentences (thus understood) and the statements of science. Neurath (1932), therefore, developed a physicalist and intersubjective conception whereby protocol sentences referred to physical objects and processes. Quine (1990: 3) took the intersubjective conception further: observation sentences report facts that competent speakers of the language would unanimously agree on, using perception. While enabling the connection to science, the physicalist, intersubjective approach entails that the basic evidence claims of science are not certain and may be corrected in the light of further scientific research. And so Neurath and, later, Quine took a coherentist approach to the nature of evidential support, contrasting with Schlick's foundationalism. For example, in place of Schlick's phenomenalist protocol sentence, a physicalist protocol sentence might report, 'this object is white'. However, should other evidence

suggest that lighting conditions are not normal, then it might be appropriate to withdraw assent to this statement. In Neurath's analogy of the boat, any plank might be replaced, none is foundational. What makes a statement an evidence statement is not just that it supports the other claims of science, but also that it contributes to a maximally coherent whole.

Summary: One empiricist tradition, exemplified by Mill and Schlick takes observations to provide a foundation for science in certainty by identifying observations with reports of inner experience. This, however, makes observation very distant from anything employed in scientific practice. Another empiricist tradition, exemplified by Neurath and Quine, gives up on certainty, allowing observation statements to be corrigible, intersubjective reports of unmediated sensory contact with the world.

3 Theory-ladenness of observation

The shift to a physicalist conception of protocol sentences made the connection to science easier, in that this avoids the potential scepticism or solipsism of Schlick's phenomenalist foundationalism. Nonetheless, a physicalist protocol statement remains distant from the observation reports actually made by scientists. These protocol statements report what the subject can directly perceive ('the pointer is at 12.5'), whereas the observation reports of scientists record the values of scientifically important quantities or qualities ('the current is 12.5 mA'). The correctness and justification of the physicalist protocol statement depends, we have seen, on the truth of certain propositions about the subject's environment (and the subject herself). The truth and justification of the scientist's observation report depends not only on these but depends also on further truths of a scientifically more sophisticated kind.

How, then, is the perceived fact (regarding the pointer) related to the recorded observation (regarding the current)? A typical (permanent magnet moving coil) ammeter is based on theoretical principles first developed by André-Marie Ampère, following the famous experiment performed by Hans Christian Ørsted which showed that an electric current can deflect a magnet by producing a magnetic field. The law employed by the ammeter states that the torque T on a coiled wire carrying a current I in a magnetic field B is given by: $T = NIAB \sin\theta$, where N is the number of coils, A is the area of the coil, and θ is the angle between the magnetic field and the normal to the coil. The torque can be used to deflect the pointer in the ammeter, and so the pointer indicates the current—if this equation is correct.

Hence the correctness and justification of the observation that the current is 12.5 mA are dependent on the truth of a theory: it is said that the observation is *theory-laden* or *theory-dependent*.

The traditional empiricist view is that evidence (in the form of observation) is one thing and theory is another, and that our rational beliefs in the latter depend on our possession of the former and never vice-versa. The fact that evidence in science is theory-laden is a challenge to this idea. The challenge takes on a sceptical or anti-realist turn if it can be shown that the theory is being tested against or confirmed by evidence that is dependent on that very same theory. Scientific theories are *not* in fact typically tested against observations that are laden with precisely those theories (Franklin 2015).

One response to the theory-ladenness threat to the observation–theory distinction is to require than in every case where an observation is laden with some theory, one is in a position to justify that theory by reference to other observations (Kosso 1992). There will be chains of such justifications, since those observations will themselves be theory-laden. (Note that these 'chains' will often link together because the same observation might be evidence for more than one theory, and one theory will be supported by multiple observations.)

One might think that these chains bottom out with observations that are not theory-laden. If so, should those observations be regarded as the 'real', basic, or ultimate evidence, such that the theory-laden observations are properly to be thought of as proxies for that basic evidence?

However, as our discussion of the protocol sentence controversy indicated, the observations that would be such basic evidence are not certain. The corrigibility of the physicalist protocol sentences suggests that every such report is theory-laden in the following sense. Every such report might be considered to be laden with the contingent, empirical hypothesis that no correcting condition holds. As mentioned above, one might correct or withdraw the statement 'this is white' if one discovers that lighting conditions are abnormal. So the observational claim, 'this is white' looks to be laden with the hypothesis that conditions are normal. In that sense *every* observation is theory-laden and so it is inevitable that there is indirect dependence of observation on theories that are themselves under test.

The view that the chains bottom out with observations that are not theory-laden is naturally thought of as a version of epistemological foundationalism—such observations, the basic evidence, provide the foundations for the whole structure of observations and theories. Neurath's rejection of Schlick's foundationalism is taken to be a version of coherentism.

In this picture all propositions in science, including those we label 'observations' gain support from other propositions.

The challenge for coherentism then is to explain how the system of propositions connects to the world (that is, something outside the system of propositions) at all. Some attempt this from within coherentism by identifying a class of 'cognitively spontaneous beliefs'—these are intended to be those beliefs that arise directly from experience. These, it is claimed, can gain coherentist justification when in the system there is a proposition asserting the general reliability of these cognitively spontaneous beliefs (BonJour 1985).

A different approach is to combine elements of both foundationalism and coherentism, as in Haack's (1993) *foundherentism*. The latter, like the empiricist foundationalism we have been considering, gives a foundational role to experiential beliefs. But like coherentism, foundherentism accepts that such beliefs may get some degree of justification from other beliefs.

We have seen that a typical observation statement in science is laden with scientific theory. It is tempting to see it as an inference from what one perceives (e.g. of an instrument) plus background scientific theory. Even what one perceives is theory-laden, since it can be seen as an inference from the subjective experience (how things appear to be) plus some minimal 'theory' such as the belief that conditions are normal for perception. Cognitive psychology suggests that even the subjective appearances of things can be considered an inference that may be theory-laden. Helmholtz (1867) had argued that perceptual experience is the result of unconscious inferences. Similarly, according to New Look psychology, visual perception (for example) may be understood as a sub-personal inference from the pattern of retinal stimulation supplemental by background knowledge. Jerome Bruner and colleagues (Bruner and Postman 1949; Bruner et al. 1951) showed that perceptual experiences can be affected by prior conditioning. For example subjects used to standard playing cards found it difficult to judge the colour of an anomalous card (a five of hearts where the hearts are black, not red). More recently, predictive processing (Hohwy 2018) regards the brain and its perceptual mechanisms as constantly forming models of the world in the light of background information. (Fodor (1984) notes, however, that perceptual experience is not always open to penetration by just any belief.)

Summary: The traditional view has been that theories are tested by comparing their predictions to observations. So observation is epistemologically prior to theory. However, this relationship seems often to be reversed. Many observation statements in science are such that in order for those propositions to be known to be true, some piece of scientific theory needs to be known to be true also. One might think that there is a class of observation statements

for which this is not the case, and that these are the basic (or 'real') observation statements. But, as we saw in the preceding section, any non-subjective observation claim is potentially corrigible. So, it could be said that every observation statement is theory-laden, at the very least with some theory to the effect that conditions are not such that the statement needs correcting (e.g. 'observational conditions are normal').

4 Theoretical evidence

If one accepts the claim that observation is at least sometimes theory-laden, then one ought to accept that evidence can be at least partly theoretical. Perhaps, then, evidence can sometimes be *entirely* theoretical?

Scientists often talk of 'theoretical evidence': witness published papers with titles such as 'Experimental and theoretical evidence for the existence of absolute acoustic band gaps in two-dimensional solid phononic crystals', 'Empirical and theoretical evidence of economic chaos', 'Theoretical evidence for a dense fluid precursor to crystallization', and 'Observational and theoretical evidence in support of a significant in-situ photochemical source of tropospheric ozone'. Nonetheless, the concept of theoretical evidence has received no discussion among philosophers. Indeed to many the idea may appear to be an oxymoron precisely because, they hold, to say that scientific evidence is observational is to express a tautology.

Be that as it may, theoretical propositions are used in scientific inferences to justify scientific beliefs and to extend scientific knowledge. And so, by the functional characterization of 'evidence' I articulated at the end of Section 1, such propositions should be counted as evidence. A typical use of theoretical evidence occurs when scientists make predictions regarding the existence (or nature) of some phenomenon from a theory (typically with auxiliary assumptions). (A phenomenon in this sense is a repeatable empirical state of affairs or process—such as the photoelectric effect, the sublimation of carbon dioxide, population cycles in ecology. A phenomenon may not have sufficient generality or explanatory power to be called a 'law', though some phenomena, such as the elliptical orbits of planets, may.)

The standard picture of scientific practice suggests that if the predicted phenomenon is confirmed by observations (e.g. experimental measurements) then it is the observations that are evidence for the existence of the phenomenon, and thereby for the theory. However, if the theory is itself well-established, its epistemic status may be more robust than that of the predicted phenomenon. In which case our reasons for thinking that the phenomenon

exists will include the fact that a well-established theory predicts its existence. Since the theory provides a reason to believe in the phenomenon, it provides evidence for the phenomenon.

5 Non-empiricist conceptions of observation and evidence

The standard, empiricist, philosopher of science takes evidence to be observational and takes observation to be perceptual. Thus philosophers of science concerned with the nature of the evidence on which theories rest find themselves discussing the theory-ladenness of observation, where the latter is regarded as a matter of the influence of theory (or background belief or prior experience) on *perception*. (See, for example, Fodor's (1984) 'Observation reconsidered'.) A small number of philosophers of science have offered non-empiricist ways of thinking about observation and evidence.

Bogen and Woodward (1988) reject the idea that the evidence with which scientists test theories is the 'observable' as (empiricist) philosophers of science have traditionally construed 'observation' (i.e. in terms of what is perceptible). They draw a distinction between 'data' and 'phenomena'. 'Examples of data include bubble chamber photographs, patterns of discharge in electronic particle detectors and records of reaction times and error rates in various psychological experiments. Examples of phenomena, for which the above data might provide evidence, include weak neutral currents, the decay of the proton, and chunking and recency effects in human memory' (Bogen and Woodward 1988: 306). While data are for the most part observable, it is not data that provide evidence *for theories,* it is phenomena. On the other hand, phenomena are not observable, they are inferred. Bogen and Woodward discuss the example of the melting point of lead. Contra Nagel (1961: 79), they point out that this is not directly 'observable', since the experimenter has to take a number of readings from the thermometer, which will not all be the same. These readings are the data, and from these the scientist calculates a mean value and a standard deviation. This gives us the melting point of lead, and that is a phenomenon that can be used to test some theory, for example concerning the internal bonds in lead and how they depend on the structure of lead atoms.

Although Bogen and Woodward, using the standard philosopher of science's conception of observation, deny that *observations* are evidence for theories, what is more important is that they deny that the evidence for theories typically concerns what is *perceptible*. On the other hand, scientists' use of

the term 'observation' does permit them to talk about the phenomena being observable, in papers such as 'Observation of Gravitational Waves from a Binary Black Hole Merger', where such observations (extracted with a great deal of inference from the data) provide evidence for general relativity.

Dudley Shapere (1982) points out that the philosopher's use of 'observation' takes the latter term to play a double role. First, 'observation' denotes a special kind of *perception*. Secondly, 'observation' picks out an *evidential* role: observation is what provides a foundation for knowledge or well-grounded belief in science. That these two roles coincide is a central tenet of empiricist philosophy of science: only perceptual experience grounds belief and knowledge. Shapere argues that in modern science these two roles have come apart.

What scientists take to fulfil the evidential role is not always (nor even typically) perceptual in nature (as we have seen Bogen and Woodward argue; see also Bird 2022). Shapere notes that what scientists call the 'direct observation' of the core of the Sun is a complex process whereby neutrinos produced in the Sun's core are captured in a vat of cleaning fluid at the bottom of a mineshaft. This 'capture' of neutrinos is a matter of a small proportion of the solar neutrinos passing through the cleaning fluid interacting with chlorine atoms to produce argon atoms. The latter can be counted because they are themselves radioactive, and their decay products can be quantified using standard techniques. Perception plays no significant role in this process of observation. On the other hand, inference and theory are important. To produce this observation of the solar core requires a great deal of inference from (well-grounded) theory (e.g. concerning radioactive decay) as well as other experimental results (e.g. which tell us how much argon is produced in the cleaning fluid by other processes).

The conclusion from Shapere's example and many others like it is that, as 'observation' is used by scientists, observation is not perceptual nor even prior to inference. What then is observation? Shapere himself analyzes observation and observability in terms of information transfer: an entity x is directly observed if there is direct transfer of information from x to a suitable receptor. 'Direct transfer' here means 'transfer without interference'. For example, solar neutrinos travel directly, without interference, from the Sun's core to the cleaning fluid in the detector, whereas any photons produced in the Sun's core are quickly absorbed by atomic nuclei and re-emitted, and so any information transfer is indirect, and the character of the information is itself altered in the process. (Photons do carry information directly from the outermost layer of the Sun, the photosphere, since from there to the Earth's

surface, there is little interference. The photosphere *is* thus observable by the light it produces.)

Summary: Empiricists give a central role to perception. Yet the propositions that scientists use as evidence propositions, propositions against which the claims of theories are tested or which are used to justify theoretical claims, are very rarely propositions that can be known by perception. The empiricist conception of scientific evidence is difficult to maintain. Empiricist philosophy of science uses 'observation' to denote both what can be known by perception and what is used to test and justify theories. Since nothing plays both roles we have to choose whether to use 'observation' as the philosophers do (to denote scientific propositions known by perception) or as scientists do (to denote a certain class of scientific evidence propositions).

6 Conclusion

Empiricism has dominated thinking about evidence by philosophers of science. Moderate empiricists hold that it is possible to justify theoretical claims by inference from or testing against evidence claims, where the latter are what *philosophers of science* typically call 'observations'—propositions that are closely linked to perception. Modern science presents a challenge to this picture, since the propositions used by scientists in testing and justifying theories are not perceptual at all. Indeed, they are heavily theory-laden. Some of these evidence propositions are themselves theoretical claims—theoretical evidence. Some are the results of experiments and the like. *Scientists* call these propositions 'observations' or 'observational evidence' despite their lack of connection to perception. Such observations are typically the result of inferences from data (which are often collected automatically). Some philosophers have proposed post-empiricist conceptions of data, phenomena, and observation in work that represents an important new direction for the philosophical understanding of scientific evidence.

References

Achinstein, P. 2001. *The Book of Evidence*. New York, NY: Oxford University Press.

Bird, A. 2018. Evidence and inference. *Philosophy and Phenomenological Research* **96**: 299–317.

Bird, A. 2022. *Knowing Science*. Oxford: Oxford University Press.

Bogen, J. and J. Woodward 1988. Saving the phenomena. *Philosophical Review* **97**: 302–52.

BonJour, L. 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Bruner, J. S. and L. Postman 1949. On the perception of incongruity: A paradigm. *Journal of Personality* **18**: 206–23.

Bruner, J. S., L. Postman, and J. Rodrigues 1951. Expectation and the perception of color. *The American Journal of Psychology* **64**: 216–27.

Chakravartty, A. 2007. *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.

Conee, E. and R. Feldman 2004. *Evidentialism: Essays in Epistemology*. Oxford: Oxford University Press.

Fodor, J. A. 1984. Observation reconsidered. *Philosophy of Science* **51**: 23–43.

Franklin, A. 2015. The theory-ladenness of experiment. *Journal for General Philosophy of Science* **46**: 155–66.

Haack, S. 1993. *Evidence and Inquiry*. Oxford: Blackwell.

Hohwy, J. 2018. The predictive processing hypothesis. In A. Newen, L. D. Bruin, and S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.

Kosso, P. 1992. Observation of the past. *History and Theory* **31**: 21–36.

Mill, J. S. 1843. *A System of Logic, Ratiocinative and Inductive*. John W. Parker. Reference to the 1974 edition, J. M. Robson ed. Toronto: University of Toronto Press.

Nagel, E. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace & World.

Neurath, O. 1932. Sozialbehaviorismus. *Sociologicus* **8**: 281–8.

Papineau, D. 1978. *For Science in the Social Sciences*. Basingstoke: Macmillan.

Quine, W. V. 1990. *Pursuit of Truth*. Cambridge, MA: Harvard University Press.

Schlick, M. 1934. Über das Fundament der Erkenntnis. *Erkenntnis* **4**: 79–99.

Shapere, D. 1982. The concept of observation in science and philosophy. *Philosophy of Science* **49**: 485–525.

Uebel, T. 2007. *Empiricism at the Crossroads: The Vienna Circle's Protocol-Sentence Debate*. Chicago, IL: Open Court.

van Fraassen, B. 1980. *The Scientific Image*. Oxford: Oxford University Press.

van Fraassen, B. 2001. Constructive empiricism now. *Philosophical Studies* **106**: 151–70.

von Helmholtz, H. 1867. *Handbuch der Physiologishen Optik*. Leipzig: Leopold Voss.

Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.