Evidence
Lecture 1 (27. Apr 2022)
Introduction

Erik Curiel†
(date of notes: 27. Apr 2022)

Contents

1 Administration, Organization

1. recording lectures
2. anyone not getting emails?
3. who I am
4. what you can expect from me
5. what I expect from you

† **Author's address**: Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität; Black Hole Initiative, Harvard University; **email**: erik@strangebeautiful.com

6. voluntary student presentations
7. short weekly essays
8. anything else?

2 Standard Accounts and Their Discontents – or – The Indefatigable Orneriness of Evidence in the Wild

> Ahhh, it's so hard, ya' know, it's so hard to believe in anything anymore, ya' know what I mean? It's like religion, you can't really take it seriously, because it seems so mythological and it seems so arbitrary, and then on the other hand, science is just pure empiricism, and by virtue of its method it excludes metaphysics. And I guess I wouldn't believe in anything if it weren't for my Lucky Astrology Mood-Watch.
>
> Steve Martin
> "A Wild and Crazy Guy"

1. toy model of standard philosophical view of evidence, *viz.*, raising probability: $p(H|E) > P(H)$
2. what an unruly mob of philosophical sins are hidden under the seemingly innocuous 'E'!
3. example illustrating why it's sophomoric: Newtonian gravitational theory used to structure data used in evidence for GR

3 Evidence – Some Questions

fundamental question: what is evidence?

1. it's just this thing, ya know?
2. perhaps true, but not helpful, so let's break the question down into components:
   a. what can one have evidence for?
   b. what can evidence consist of?
   c. how does it act as evidence?
3. components:

   what can one have evidence for?

   a. a scientific claim about the state of the world (past, present, future, spatially bounded or not, temporally bounded or not, spatiotemporally bounded or not):
      i. different kinds of descriptions and characterizations (see `description-explanation-notes.tex` and related files)
      ii. conditional claims

   iii. existential claims

   iv. contextually universal claims

b. a scientific claim about the modal structure of the world (no adherence to realism of any kind required to make and understand this claim)

   i. conditional and counterfactual claims (IVPs, *e.g.*)

   ii. dispositional claims

   iii. universal claims

   iv. causal claims

   v. claims of possibility and necessity

   vi. claims of probability

   vii. frameworks, theories

   viii. global and architectonic structures (*e.g.*, spacetime structure *à la* Stein 1967)

c. a scientific claim about the reliability of a methodology or contraption of any relevant kind (instruments, experimental arrangments and procedures, forms of theoretical argument, and so on)

d. more generally speaking, therefore, the "hypothesis" for which one can have evidence may be a claim about the world, or a claim about our methods for investigating and studying the world (*e.g.*, about their fruitfulness)

e. such includes: frameworks (Newtonian mechanics, *e.g.*); theories; "truer method[s] of philosophy" (Newton 1726, preface to 1st edition); *etc.*

what can evidence consist of?

a. not all substantiated scientific claims can serve as evidence for anything else:

   i. Smeenk's examples of cosmology in the 1930s—figure out exactly what it was, go back and watch the video of the talk and the Q&A

   ii. sometimes, we are not in a position to use a warranted claim as evidence, because we do not know or understand enough to make use of it—conservation of momentum was not used as evidence for anything else for a *very* long time, I suspect until the 1920s-1930s in particle physics, *e.g.*, the argument for the existence of neutrinos (see email I sent to Hans about this)

   iii. are facts evidence? claims? propositions? observations? structured data?

how does it act as evidence?

a. what can evidence do? confirm, corroborate, test, explain, yield understanding, suggest new avenues of research...

b. one may be tempted to say that all the possibilities mooted are only instances of confirmation (we confirm, *e.g.*, that the clarified concept is the correct one, or something like that); but that would be to efface the importance differences among the different activities evidence engages in, to the detriment of philosophical understanding

c. more schematically, I conjecture that it often acts in the following way: it serves, in a particular investigative context, as an essential conjunct [*** what is Peirce's word for this? ***], unique up to argumentative equivalence, in a sound argument supporting what it purports to be evidence in favor of

d. break this up:

   i. "essential": not logically necessary, since the argument may not even be such as to lend itself to formalization in a system that a formally minded philosopher would countenance; rather, the argument could not go through without it

   ii. "argumentative equivalence": any other potential piece of evidence that can be used in an argument of the same type, using the same [*** most of the same, basically the same, ... ***] other conjuncts [*** Peirce's word ***], for the same conclusion, all such that one of the arguments would fail if and only if the other would (a trivial case: whether one experimentally measures 49% or 51% for the frequency of an outcome theoretically predicted to be 50%, with a margin of error of 2%)

   iii. "sound": relative to investigative context

   iv. "supporting": conducing to conditional assertability, conditional acceptance, *not* necessarily belief

   v. general notes: necessary involvement of theory in bringing evidence to bear, often in the person of, or at least grounded in, schematizing the observer; no necessary danger of theory-ladenness in any pernicious sense

e. I do not claim that every instance of what could reasonably be called 'evidence' in every case of a human activity that could reasonably be called 'scientific' must conform even loosely to this schematic characterization; I do claim that: the schematic characterization captures much if not most of what is commonly agreed to be evidence in physics; it likely does so in many if not most other sciences; and, most important, it gives us insight into how scientific reasoning works, why scientific reasoning works, and the structure of scientific knowledge

f. here is a funny kind of example that may not conform to this schema: Röntgen notices that crystals in a darkened room near gas discharge tubes begin to glow green—what was this evidence for, and how could it possibly be evidence of anything? Was it evidence? It surely seems to be in some sense—evidence that something unknown is in operation that needs further investigation

4. some consequences:

   a. the communal character of scientific knowledge and reasoning; think of Popper's marvelous remark Popper (1959, part II, ch. 5) to the effect that an observation sentence is not an autobiographical report, but a conventional statement of socially accepted scientific fact

   b. there are similar facts about theoretical reasoning one must attend to

   c. think about how cryptic and obtuse Heisenberg's original 1925 arguments were, but they were "accepted" in the sense as to spur others to recapitulate their essence in more perspicuous, more fruitful, more general ways, to use them as springboards for new directions of research, *etc.*

4 Theory and Experiment As Part of Evidence (Or: What Allows Something to Act as Evidence)

1. The following discussion is supposed to show how purely theoretical components (the calculations of the perturbative effects of the other planets and of the shape of the Sun) often (always?) play a role in the constuction of evidence, and also how, at the same time, theoretical components enter the construction of the "measured" or "experimental" or "observational" components (in this case, the reliance on the hypothesis of stability and regularity to justify the averaging and extrapolating calculations performed to derive the precession *per century*—which is obviously something one cannot observe by simple telescope and stop-watch alone). And also how complex all those relationships can be, both in their own mixture of observation, model and calculation, and also using a theory one believes to be false to construct evidence for one that one hopes is better.

   Formal accounts can't account for the complexity of the confirmational relations among theories: we actually use Newtonian gravitational theory to construct a lot of the evidence that we use to measure the goodness of general relativity (precession of Mercury's perihelion: the calculation of its actual value uses Newtonian gravitational theory, one then argues that Newtonian gravitational theory can't account for it, but general relativity can; calculation and use of flattening of the velocity curves for spiral galaxies has a similar structure). The precession example is instructive: the calculation of Mercury's *total* precession is monstrously complex, involving perturbative contributions from the effects on Mercury's orbit of Jupiter and the other planets, the oblateness of the sun, and so on, and we then *subtract* all that from the precession calculated from observational measurements; all that, however, calculations of perturbative effects and the calculated precession based on observations, are *averaged* quantities, whose suitability for use as evidence depends on our belief in the regularity and stability of the phenomena at issue; but we don't have *the foggiest clue* how to do or show *any* of that (calculations, demonstrations of regularity and stability) in even perturbative or approximative models in general relativity—the linearized Schwarzschild model we use to calculate Mercury's anomalous perihelion certainly does not suffice.

   It's interesting to work through in detail how that calculation is supposed to work to provide confirmatory support for general relativity. The Sun-Mercury system is modeled as a reduced 2-body system, with no perturbative influences from anything else (including, *e.g.*, the oblateness of the sun); Mercury's orbit is calculated, and it is shown that the orbit is not closed, but precesses; that precessions then matches what cannot be accounted for by Newtonian models and means. But the value arrived at by complex calculation based on observation is not that anomalous value, but is rather a total value, from which the "Newtonian" contribution is subtracted. But we have no idea how to recover that Newtonian contribution in general relativity. So is general relativity really "predicting Mercury's orbit"?

No. It is returning the difference between, on the one hand, the value derived by a complex Newtonian calculation based on Newtonian models of the Solar System (including fine details of the sun's shape) and, on the other, the value derived by a complex Newtonian calculation involving observed values themselves derived by the use of Newtonian theory to model and validate the observations.

The use of, *e.g.*, Bayesian techniques to model this ("$P(h|e)$") obscures, indeed grievously distorts, the epistemic complexity of the situation. One may want to say, "well, it's only an idealization, used for a specific purpose, calculating a confidence value, everyone knows that a more sophisticated and complex treatment is required to address other questions, *e.g.*, about the nature of evidence"—but no, I submit that everyone does *not* "know" this, because this kind of idea is *never* discussed in presentations and discussions formal confirmation theory. Formal confirmation theory is never presented as an "idealization", but rather only as the complete and satisfactory answer to all questions about confirmation one could have. It is, moreover—can be—only a prodigious leap of faith to conclude that the Bayesian analysis (*e.g.*) works even to address only "narrow questions of confirmatory support", for once one realizes how complex the evidential relationships really are, it becomes evident that one must *argue* that "narrow questions of confirmatory support" really are *so* narrow as to be adequately represented and answered by such childishly simple and naive expressions as "$P(h|e)$".

2. Peirce (1992, p. 185):

> We may, however, and do desire to find formulas expressing the relations of physical phenomena which shall contain no more arbitrary numbers than changes in the scales of measurement might require. ¶ When a formula of this kind is discovered, it is no longer called an empirical formula, but a law of Nature; and is sooner or later made the basis of an hypothesis which is to explain it. These simple formulæ are not usually, if ever, exactly true, but they are none the less important for that; and the great triumph of the hypothesis comes when it explains not only the formula, but also the deviations from the formula.

See (Harper 2011; Smith 2014) on theory-mediated measurements and the process of iterated approximations in Newton. *This* is why abduction, for Peirce, is not simply the fallacy of "affirming the consequence", and why, *contra* people like Hartmann, it does not suffice to give a probabilistic (*e.g.*, Bayesian) account of abduction, for the *warrant* of the abduction comes not from the logical form used to "derive" it, but from the *process of testing it, because that process itself derives from the logical form of the abduction*—the abduction itself, the representation of it, shows one what constitutes epistemically severe tests of the abducted proposition, and that includes perhaps most importantly of all explaining the deviations of *deduced* predictions from the formula. One *needs* an abduction *already in place* in order to decide whether an observed deviation of an experiment from a deduced prediction ought to count as a refutation of the formula (hypothesis) at issue. This explains the opaque remark earlier in the paper (Peirce 1992, pp. 191–192):

> But, after all, there is an immense difference between the relation of *Baroco* and

> *Bocardo* to *Barbara* and that of Induction and Hypothesis to Deduction. *Baroco* and *Bocardo* are based upon the fact that if the truth of a conclusion necessarily follows from the truth of a premise, then the falsity of the premise follows from the falsity of the conclusion. This is always true. It is different when the inference is only probable. It by no means follows that, because the truth of a certain premise would render the truth of a conclusion probable, therefore the falsity of the conclusion renders the falsity of the premise probable. At least, this is only true, as we have seen in a former paper, when the word probable is used in one sense in the antecedent and in another in the consequent.

This is why abduction provides the grounds for the possibility of a rational procedure of belief fixation: without it, the verification of predictions cannot provide epistemic warrant for the hypotheses from which the predictions were deduced. In other words, abduction is what allows potential evidence to serve in fact as actual evidence.

3. One might think that the empirical predictions of a theory are always inferred indexically, relative to some (class of) target system(s), because a theory never makes "direct contact" with experimental data, but rather only "models" constructed from the theory (Morgan and Morrison 1999; Giere, R. N., 1988, Explaining science: A cognitive approach, University of Chicago Press, Chicago; Frigg and Hartmann, "Models in Science", *SEP*; also of possible relevance is Nguyen, J. 2017, Scientific representation and theoretical equivalence, Philosophy of Science 84(5), 982–995). I doubt this all around, unless "some (class of) target system(s)" is construed so broadly as to make it vacuous. Think of general propositions about "generic systems" treated by a theory, *e.g.* energy conservation for conservative forces in Newtonian mechanics. Also, think of my schematic models (previously called 'principle models'), which have no well defined target class of systems.

5 Theory as Evidence

1. is the equilibrating of a hot cup of tea with the air in my living room evidence for anything in thermodynamics (*e.g.*, the Minus-First Law)? I think not. At this point, thermodynamics is so deeply entrenched that it is much more likely that the theory of equilibration would be used as evidence for the tea's behavior than vice-versa.

2. Katie says (abstract to her BSPS talk "Reduction, Abstraction and Approximation"):

> In contrast, in the case of vertical reduction, the two theories describe different subject matters – it is not the case that one theory is more accurate than the other, since each theory answers different questions. In particular, the higher-level theory abstracts away from details described by the lower-level theory.

I cannot agree. Often, the only evidence we have for the "lower-level" theory is *provided by* the higher-level one. Most evidence we have that statistical mechanics adequately applies to any given kind of physical system, *e.g.*, is itself thermodynamical in character (in the sense that one must use thermodynamics both to formulate claims about it and to corroborate it experimentally), so they had better be "answering the same questions", at least in these cases.

6 The Role of the Investigator

Can a fact can have evidential value solely in virtue of "objective" relations it stands in to the relevant theory? (Whether it can be "derived as a prediction" from the theory, and so on.) Or does the activity of an actual epistemic agent sometimes, or always, play a role? What about the general epistemic state of the community?

1. Sometimes it seems as though an observed fact must stand in a particular relation to the past actions of a particular investigator for it to have evidential value. Example: Laura opined that I regularly feel crappy on Wednesdays, because Tuesdays are always long and hard days for me (preparing lecture, giving lecture, meetings, *etc.*); I awoke this Wednesday morning and, feeling crappy, remembered that; because I did not *predict* I would feel crappy, it can't count as evidence in favor of Laura's hypothesis, because of the possibility of observational bias (I take note of how I feel on Wednesdays only when I feel crappy, which, in fact, is not often). How can a relation to a particular action by a particular epistemic agent transform a fact into evidence, or preclude it from counting as such?

2. Often, if not always, the general state of the entire epistemic community also bears on whether a fact can count as evidence. What good standards of argumentation with regard to a particular theory, for example, constrains what can legitimately be thought of as a "prediction" of the theory, and that can change over time.

7 Approximation, Error and Tolerance

1. The identification of solutions with evidentiary data—I think essentially always—comes down in the end to the brute comparison, one by one, of the values of theoretically calculated quantities on the one side with the values of the analogues of those quantities in the structured data on the other, with the choice being made each time whether the values are close enough, given one's understanding of the errors and imprecisions in measurement, the introduction of skew in statistical reconstruction, one's pragmatically chosen error tolerance, and other such factors.[1]

8 The Evolution of Evidence

What once needed evidence now stands as evidence for other claims. This sometimes happens when what once had been taken as kinematic is now understood as dynamic (*e.g.*, understanding a kinematical constant such as shear viscosity in Navier-Stokes theory as a manifestation of the underlying molecular micro-dynamics)—such a shift in understanding requires, marks, in part

1. [*** cite Fillion on "standards of accuracy" (*e.g.*, backward-error analysis versus perturbative), needed over and above norm and threshold, to determine whether inexactitude of solution is acceptable, how subtle and complex this issue is; he explained it in talk given at EPSA 2019 Geneva ***].

constitutes one kind of maturity of the treatment of the relevant phenomena by theory and experimentation in tandem. The science has become, to use a term much *en vogue*, a data-driven science. The data it collects once stood in need of powerful evidence for its credibility; its collection now is a matter of course and provides credibility both for refinements of existing models for some class of phenomena (such as deeper and more extensive articulation of systematic regularities and dependencies) and for extension of experimental reach to new classes of phenomena.

9 Interpreting Evidence

See

1. Loeb, "Good data are not enough", *Nature*, `Desktop/phil/sci-evid-meth/loeb-good-data-not-enough.pdf`

10 Misleading Evidence

on the view that self-locating evidence is required for "distinguishing different large-population hypotheses" (Dorr and Arntzenius, "Self-locating Priors and Cosmological Measures"): seems to rule out the possibility that we could ever discover that "unlikely" hypotheses are actually true (when they are true)—think of the case in which the Born Rule works for everything *except* that every time we send a silver atom through a Stern-Gerlach device, it comes out spin-up. Smeenk: "Bayesianism does not (and should not be expected to!) guarantee that agents boost credence in true theory in all possible worlds." Yes, absolutely, but—I want to say: we should still believe in quantum theory, but we would just stop using silver atoms for Stern-Gerlach experiments. The *systematicity* and *specificity* of the failure is evidence here. (Need to define "systematicity" and "specificity" here.)

References

Harper, William L. 2011. *Isaac Newton's Scientific Method: Turning Data into Evidence about Gravity and Cosmology.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199570409.001.0001.

Morgan, M., and M. Morrison, editors. 1999. *Models as Mediators: Perspectives on Natural and Social Science.* Ideas in Context 52. Cambridge: Cambridge University Press.

Newton, Isaac. 1726. *Philosophiæ Naturalis Principia Mathematica.* Third. Volume I. The translation by A. Motte of the third edition (1726), originally produced in 1729, revised by F. Cajori and published in 1934. The first edition of the *Principia* was published in 1686, the second in 1713. Berkeley, CA: University of California Press.

Peirce, C. S. 1878. "Deduction, Induction, and Hypothesis". *Popular Science Monthly* 13:470–482.

Peirce, Charles Sanders. 1992. "Deduction, Induction, and Hypothesis". Chapter 12 in *The Essential Peirce: Selected Philosophical Writings,* edited by N. Houser and C. Kloesel, volume 1 (1867–1893), 186–199. Bloomington, IN: Indiana University Press. Originally published as Peirce (1878).

Popper, K. 1959. *The Logic of Scientific Discovery.* London: Hutchinson & Co.

Smith, George E. 2014. "Closing the Loop: Testing Newtonian Gravity Then and Now". Chapter 10 in *Newton and Empiricism,* edited by Zvi Biener and Eric Schliesser, 262–351. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199337095.003.0011.

Stein, Howard. 1967. "Newtonian Space-Time". *Texas Quarterly* 10:174–200.