Knowledge-First Inferential Evidence: A Response to Dunn

Timothy Williamson

[invited for special issue of *The Monist* on infallibilism, edited by Christos Kyriacou and Gregory Stoutenburg]

In his carefully argued paper 'Inferential Evidence', Jeffrey Dunn aims to show that the account of evidence in *Knowledge and its Limits* is internally inconsistent (Dunn 2014, Williamson 2000). More generally, he argues that Bayesian epistemology excludes evidence gained via inductive inference, while my account both shares the relevant features of Bayesianism and accepts inductive inferential evidence. He considers various ways of modifying my view to avoid the problem. I will explain how my view, unmodified, meets his challenge.

Here is the core of Dunn's argument (206).[1] An inference from new evidence E to a conclusion A is *inductive* when the prior conditional probability of A on E was less than one: $\text{prob}_{t0}(A \mid E) < 1$. Here t0 is a time immediately before you gain E (and no further evidence), and prob_{t0} is the probability distribution on your evidence at t0. According to Bayesianism, updating is by standard conditionalization, so the posterior unconditional probability of A is its prior probability conditional on the new evidence, and so is also less than one: $\text{prob}_{t1}(A) = \text{prob}_{t0}(A \mid E) < 1$. But if A itself becomes part of your evidence via this inference, then A is entailed by your evidence at t1, so its posterior probability on your evidence at t1 must be one: $\text{prob}_{t1}(A) = 1$. That is a contradiction.

As Dunn says, my account of evidence was firmly within the target area of his argument (Williamson 2000: 184-237); it remains so (Williamson 202X). Given my equation of one's total evidence with the total content of one's knowledge (E = K), gaining evidence via inductive inference is equivalent to gaining knowledge via inductive inference. I insist that we often gain knowledge via what Dunn would describe as inductive inference. To use his example, 'When light filters through the curtains in the morning, I do not directly perceive that the sun has risen', but it is still natural to say that my evidence for thinking that it is morning is that the sun has risen (202). I come to know that the sun has risen; *ipso facto*, that the sun has risen becomes part of my evidence. Although I deny that probabilities on one's evidence update *only* by conditionalization, since knowledge and so evidence can be lost as well as gained, my view entails that when no knowledge is lost, probabilities on one's evidence update by conditionalization. That is enough for Dunn's argument, since no relevant knowledge is lost in many cases of updating, such as those he describes.

There are non-trivial issues as to just what an inference *is*. However, I am not accusing Dunn, and Dunn is not accusing me, of equivocating on the term 'inference', or misusing it. Our disagreement is not merely terminological. For present purposes, we can understand the term informally, and allow the complexities to emerge as we proceed.

Dunn distinguishes inductive from deductive inference according to whether the prior probability of the conclusion on the evidence is less than or equal to one. He also describes inductive inference as 'less than fully certain' and deductive inference as 'fully certain' (206). Those are unfortunate ways of drawing the distinction. One difficulty is that probability distributions over infinitely many possibilities often have to assign probability zero to possible events (in technical terms, they are non-regular). For instance, on plausible assumptions, the probability that a fair coin comes up heads on each of an infinite sequence of tosses is zero, even if infinitesimal probabilities are available; likewise for every other specific outcome of the tosses (Williamson 2007). Thus if E is 'the coin will be tossed infinitely many times' and A is 'the coin will come up tails at least once', the prior probability of A on E is one, but the probabilistic inference from E to A is not naturally described as 'deductive'. By demanding standards, it is not even 'fully certain', since the conjunction A & ¬E is compatible with everything one knows. For a given probability space, it would be more natural to describe the relation between E and A as deductive when every E-possibility in the space is an A-possibility (E strictly implies A), and as inductive when not every E-possibility is an A-possibility (E does not strictly imply A). A variant of Dunn's argument could still be run on that way of drawing the distinction, since updating by conditionalization is motivated by the elimination of all possibilities in which the new evidence does not hold. Thus, for the *reductio*, the argument that the posterior probability of A is less than one would be replaced by a corresponding argument that not all remaining possibilities are A-possibilities, and the argument that the posterior probability of A *is* one would be replaced by a corresponding argument that all remaining possibilities *are* A-possibilities.

The revised distinction would still not make 'deductive' equivalent to 'fully certain' in an epistemic sense, for even when E strictly implies A, the argument from E to A may depend on a long and tricky mathematical proof which has not yet been checked by experts. There can also be what is naturally described as inductive evidence for what is in fact a logical truth. Before Fermat's Last Theorem was proved, some people may have believed it on the inductive basis that it had been checked for very many combinations of numbers. Once human limitations are factored in, many inductive arguments provide more accessible certainty than many deductive ones.

For present purposes, however, we need not worry about those complications, for numerous cases do fit Dunn's initial description; we can simply concentrate on them. My account is inconsistent *if* it has inconsistent consequences for even one such case.

As a psychological process, inference takes time. Especially just after waking, one may take several seconds to get from the fact that light is filtering through the curtains to the conclusion that the sun has risen. For an enumerative induction, the time lag between gathering the data and completing the inductive inference may be much longer. Thus the time t1 at which you acquire the new evidence E must be distinguished from the later time t2 at which you make the inductive inference to A and thereby come to know A. At t1, you do not know A; you may not even believe A (cases where you already believe A for some other reason will not help Dunn). Consequently, given E = K, at t1 A is not yet part of your evidence. Thus there is no inconsistency in the assumption that $prob_{t1}(A) < 1$. By contrast, at t2 you start knowing A, so your evidence starts including A, so $prob_{t2}(A) = 1$. The change in evidential probability reflects the change in evidence: your evidence has expanded from E to E ∧ A (ignoring irrelevant evidence). So what is the problem?

Dunn considers a somewhat similar story. In his example, the probabilities are these:

$prob_{t0}(A \mid E) = 0.9$ $prob_{t0}(A) = 0.5$ $prob_{t0}(E) = 0.5$

$prob_{t1}(A \mid E) = 0.9$ $prob_{t1}(A) = 0.9$ $prob_{t1}(E) = 1$

$prob_{t2}(A \mid E) = 1$ $prob_{t2}(A) = 1$ $prob_{t2}(E) = 1$

We can use these numbers for the sake of argument. However, Dunn envisages the inductive inference as having been made at t1, rather than t2, and so objects thus:

> The problem with this response is that the change from t1 to t2 is now completely mysterious. According to this view, the initial inference from E gets me to 0.9 credence in A, and then, subsequently, with no outside input, this credence gets bumped to 1. I see how one can say that a change in the credence to A from 0.5 to 0.9 is the result of an inference from E. But I do not see how one can say that this extra bump from 0.9 to 1 is in any way the result of an inference from E. By t2 A may be evidence, but not in virtue of any kind of inference from E. Thus, this response fails. (207)

Dunn phrases his objection in terms of credences, belief-based probabilities, whereas my primary concern is with knowledge-based probabilities on one's evidence. What has happened at t1 may simply be the automatic updating by conditionalization of one's evidential probabilities in virtue of an expansion of one's evidence, which requires no psychological process of its own, although the change of evidence may cause various further changes in one's psychological states. At t1, one may not even have *considered* A. The key inference, concluding with A, happens only at t2. That is when one's evidence expands by A, making the change in probabilities from t1 to t2 non-mysterious.

For Dunn, one's posterior credence in A 'is the result of an inference from one's full credence in E' only if one's posterior unconditional credence in A is one's prior credence in A conditional on E (205). Since that necessary condition is not met, Dunn denies that the change in one's credence in A between t1 and t2 is the result of an inference (inductive or deductive). He does not intend this to be merely a taxonomic point about what to count as an 'inference'. Rather, he thinks that reclassifying knowledge of A as non-inferential would only move the underlying epistemological problem somewhere else (207-8). The picture seems to be that once probabilities have been conditionalized on E, E has done all it can for A, and any further legitimate boost to the probability of A must come from elsewhere.

Dunn's elaboration 'with no outside input' in the quoted paragraph is odd, since new evidence need not come from outside. The most obvious examples is knowledge of one's inner states, for example of pleasure or pain, but the content may also be directed outwards. For example, I see someone who looks vaguely familiar, but I cannot place her. Some minutes after leaving, I realize that it was John's sister, whom I have not met for several years. I come to know that John's sister was there, and my evidence expands accordingly, with no outside input when it expands. Although I might also gain knowledge of my own inner states in the course of such a process, it is normally inessential to the knowledge of the external world.

Such examples suggest a more general moral. We have many recognitional capacities, for both individuals and kinds. When triggered, they typically produce knowledge that a given object is a particular individual or belongs to a specific kind. As we shall see, they can produce knowledge of other sorts too. Recognitional capacities are often triggered by evidence. But for evidence E to trigger recognitional knowledge of A, it is by no means necessary for the prior probability of A conditional on E to have been one. The connection between E and A may have been cognitively inaccessible to the subject until the recognitional capacity was triggered.

Couldn't you have known the conditional 'If E, A' in advance?[2] You might *suppose* E, prior to having gained E as evidence. The mere supposition might be enough to trigger the recognitional capacity offline, and thereby enable you to learn A on the supposition E, from which you could come to know 'If E, A' by the usual process for assessing conditionals in natural language, with supposing E and learning A conditional on E as distinct but causally related processes (Williamson 2020). Then your prior evidence would include 'If E, A', in which case your prior evidential probability for A conditional on E would indeed be one. However, that route to knowing 'If E, A' is by no means guaranteed to be open. You might not be capable of so much as supposing E in the relevant format. Perceptual evidence is often extremely rich; your imagination might not be capable of simulating E in adequate detail: for instance, face recognition depends on very subtle cues of which one is not consciously aware. The recognitional capacity too might be less reliable offline, and fail to deliver the required knowledge. In general, a rational agent's capacity to know A having learnt E does not depend on prior knowledge of 'If E, A', or even on the possibility of such prior knowledge.

Our recognitional capacities sometimes enable us to know A having learnt E with no further outside input, even when our prior evidential probability for A conditional on E was less than one. But are those cases of inductive inferential knowledge? The connection between learning E and learning A is not a merely causal one (Dunn raises a worry about merely causal connections at 208). The evidence E is knowledgeable input to the recognitional capacity, which triggers knowledgeable output A, and the process is highly sensitive to the content of both E and A. Embodied rational creatures need just such cognitive processes to find their way about their environment. Those processes may not exactly match traditional paradigms of inference, but nor does Dunn's model of inference as conditionalization of credences. I care little whether cognitive processes of the kind just described are classified as inductive inferences, so long as their epistemic legitimacy is recognized.

The formal three-stage structure also includes some cases not naturally described in terms of recognitional capacities. To borrow an example from Dunn (p.c.), let E be 'White smoke is coming from a chimney of the Sistine Chapel' and A 'A new Pope has been elected'. You know in advance that E is the conventional signal for A. Of course, E does not entail A, and before learning E your evidential probability for A conditional on E may have been slightly less than one. Nevertheless, by learning E, you may come to learn A too.

Do such cases vindicate Dunn's charge of inconsistency? The key to understanding them is that A's truth is not the only external condition underlying your knowledge of A. After all, if the white smoke had been let off by a disaffected Swiss Guard who thought that the election was still in progress, you would not have known A, though you would still have believed A. By contrast, we may assume, the actual case was not Gettierized; favourable external circumstances enabled you to know A. In accepting A outright, you took advantage of that opportunity, and gained knowledge. Since your learning E is the most salient factor in how you came to know A, the latter knowledge is naturally classified as inferential, but that should not blind us to the vital enabling role of the favourable external circumstances, which do not contribute to the prior probability of A conditional on E.

Other cases are statistical. For example, by enough sampling with replacement, you can sometimes come to know that all the balls in an urn are red, even though the prior evidential probability that they are all red conditional on the results of the sampling was less than one. Here the favourable conditions may concern the sampling: for instance, that you were not inadvertently taking the same ball each time out of the urn.

These examples suggest a more general schema for what may still be loosely described as inductively acquired evidence. At t0, you know neither E nor A; your evidential probability for A conditional on E is high, but less than one. Immediately after learning just E, at t1, you know E but not A, and your evidential probabilities are just the result of conditionalizing those at t0 on E, so the unconditional evidential probability of E is one while the unconditional evidential probability of A is high, but less than one. Learning E then prompts you to accept A outright; in favourable circumstances, at t2 you thereby come to know A, making A part of your evidence, so the unconditional evidential probability of A is trivially one. That three-stage process is fully consistent with what I say in *Knowledge and its Limits* about knowledge and evidence.

Sometimes, perhaps, accepting A is *simultaneous* with learning E: one learns both E and A. Then the middle stage t1 drops out, and stage t2 immediately follows stage t0. That telescoped process too is fully consistent with what I say in *Knowledge and its Limits*.

Contrary to Dunn's claims, my account of knowledge and evidence is quite consistent, and quite capable of characterizing without strain the phenomena on which his critique focusses.

Although I have concentrated on the challenge to my own account, analogous defences may also be available to traditional forms of subjective Bayesianism, which anyway enjoy more flexibility in what to count as evidence, given their traditional evasiveness on that score. Their schematic and normatively impoverished framework may give them comparatively little to say about the normative status of the cognitive processes at issue, but they are all the less likely to be caught out in a contradiction.

Despite my scepticism about Dunn's central argument, I will end by emphasizing that he has put his finger, if not exactly on, then at least very close to an excellent question about the fine structure of updating, one to which every properly developed account of evidence needs to have a good answer.[3]

Notes

1 All parenthetical numbers are page references to Dunn (2014). For ease of comparison, I mostly follow Dunn's notation.

2 An example of an epistemologist insisting that we must know (or be in a position to know) the conditional in advance is Bonjour 2013: 181-2.

3 Many thanks to Jeffrey Dunn, Christos Kyriacou, Sebastian Liu, and an anonymous referee for detailed comments on earlier drafts of this paper.

References

Bonjour, Laurence. 2013: ‘In defense of the a priori’, in Matthias Steup, John Turri, and Ernest Sosa (eds.), *Contemporary Debates in Epistemology,* 2nd ed., 177-184. Oxford: Wiley-Blackwell.

Dunn, Jeffrey. 2014: ‘Inferential evidence’, *American Philosophical Quarterly*, 51: 203-213.

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

Williamson, Timothy. 2007: ‘How probable is an infinite sequence of heads?’, *Analysis*, 67: 173-180.

Williamson, Timothy. 2020: *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford: Oxford University Press.

Williamson, Timothy. 202X. ‘E = K, but what about R?’ To appear in Maria Lasonen-Aarnio and Clayton Littlejohn (eds.), *The Routledge Handbook of Evidence*. London: Routledge.