Notes for Week 3 of *Confirmation*

09/12/07

Branden Fitelson

1 Chapter 3 of Stroud's *Hume*: The Negative Phase

1.1 What kind of inductive skepticism are we talking about here?

I want to begin by distinguishing three kinds of inductive skepticism that might be attributed to Hume. First, a bit of background set-up for the discussion. Let E be a statement summarizing some body of inductive evidence possessed by an (epistemic) agent ϕ. Typically, E will report something about the nature of a sample (S) from some population ($\mathcal{P}$) that has been observed by ϕ, *e.g.*, "all emeralds ϕ has observed so far have been green". Let H be some claim about unobserved objects from the population $\mathcal{P}$, *e.g.*, "the next emerald ϕ observes will be green". This example would be an epistemic *singular predictive induction.* But, we could also talk about an epistemic *universal induction* to a claim like ($H^\star$) "*all* emeralds are green".

- **Inductive Support Skepticism (ISS)**. According to inductive support skepticism (ISS), E cannot provide *any degree of evidential support for* H, for any epistemic agent ϕ (in any epistemic context C).
- **Inductive Justified Belief Skepticism (IJBS)**. According to inductive justified belief skepticism (IJBS), no epistemic agent ϕ would be *justified in believing* H on the basis of E. [An evidentialist might gloss this by saying that E cannot provide *sufficiently strong evidential support* for H, so as to render ϕ's belief in H (on the basis of E) justified. On such a reading, (ISS) is logically stronger than (IJBS).]
- **Inductive Knowledge Skepticism (IKS)**. According to inductive knowlddge skepticism (IKS), no epistemic agent ϕ can ever *know* H on the basis of E. [Again, an evidentialist who requires evidential support both for justified belief and for knowledge would view (IKS) as the *weakest* of the three kinds of inductive skepticism. I will adopt an evidentialist stance, so as to *order* these in terms of *strength*.]

1.2 A "Popular" modern reading of Hume

According to Stroud, a "popular" modern reading of Hume's inductive skepticism traces its source to the *invalidity of certain inductive arguments* that may be said to "underlie" inductive *inferences*. Before getting to the "popular" *vs* Stroudian readings of Hume, I want to get one background issue out of the way. Stroud's discussion is a bit awkward (from our perspective), since it seems to be mainly about (observed *vs* unobserved) *events*, rather than *objects*. For instance, Stroud presents the "invalid arguments" in question (those which Hume discusses in connection with of our idea of causality), schematically, as follows (on page 53):

(PE) All events of type A (that have been observed so far) have been followed by events of type B.

(PI) An event of type A is observed now.

(FE) Therefore, an event of type B will occur.

Later, Stroud talks about inductions involving propositions about *objects*. To simplify things, I will just talk about the following *specific* argument ($\mathcal{A}$), involving propositions about observed and unobserved *objects*:

(E) All emeralds observed so far (by ϕ) have been green.

(H) Therefore, the next emerald observed (by ϕ) will be green.

We are to imagine an epistemic agent ϕ "*inferring*" H from E (where "inference" is to be broadly construed in a way that allows ISS, IJBS, and IKS to all involve "inference"). And, we will think of ($\mathcal{A}$) as "the argument that underlies ϕ's inference of H from E." Now, as Stroud explains, the proponent of the "popular" modern reading of Hume gives something like the following rendering of Hume's skeptical argument [for (ISS)]:

(1) The argument ($\mathcal{A}$) from E to H is invalid.

(2) Therefore, E cannot provide any evidential support for H, for any epistemic agent ϕ (*i.e.*, E alone cannot give ϕ *any* "reason to believe" H — as Stroud puts it on p. 57). Note: this is basically just (ISS).

(3) As such, to the extent that ($\mathcal{A}$) undergirds a reasonable inference to H (on the part of ϕ), it must be *enthymematic*. The "missing premise" is usually taken to be a *uniformity principle* $E^{\star}$, which entails (at least) the conditional $E \supset H$. Hence, H is *really* being supported (if at all) by E *and* $E^{\star}$, not just E.

(4) But, now our would-be inductivist (ϕ) faces a dilemma/regress/circularity. $E^{\star}$ is not knowable *a priori* (by ϕ). Thus, the only support (for ϕ) for $E^{\star}$ must itself come from an inductive inference by ϕ. This leads either to a circle or a regress in the grounding of the inductive inference at hand (from E to H). Since circular/regressive inferences don't lend support to their conclusions, H is not supported (or justified) for ϕ [(ISS)]. [We'll return to the presuppositions underlying this last premise/step, below.]

To this reading of Hume's argument, the modern inductivist can respond simply by pointing out (to their "deductivist Humean") that the argument step from (1) to (2) is *itself* invalid. As such, (1) alone cannot provide such a deductivist Humean with any *reason to believe* (2) — *by their own lights*. In order for such a deductivist to safely *infer* (2), they must add a *bridge principle* that entails, at least, the following conditional:

(1′) (1) ⊃ (2).

And, of course, the modern inductivist will wonder why anyone should believe *that*. Anyhow, the inductivist will simply *deny* (1′), and this is how they will avoid their deductivist reconstruction of Hume's skeptical argument. Both of these arguments seem to rest on an assumption that ϕ's belief in a certain conditional cannot be grounded in a certain way — *i.e.*, that ϕ cannot know such conditionals [($E^{\star}$)/(1′)] *a priori*. It seems quite implausible (to me) that $E^{\star}$ could be known *a priori* in the salient contexts (although perhaps some "dogmatic inductivists" do believe this, or something akin to it, concerning "inductive probabilities" — more on this over the next several weeks). On the other hand, (1′) is more difficult to assess. Some authors (like Burnyeat's Stoics from last week's Milton reading) seem to suppose that (1′) is a truth of *logic/epistemology*, which *may* be knowable *a priori*. It seems to me that (1′) is false, but that's because I have a certain (weakly Harmanian) modern conception of deductive logic and its relation to epistemology. I won't dwell on this right now, but these issues are quite important, and we'll be thinking about them throughout the semester.

1.3 Stroud's alternative "internalist-epistemic" reading of Hume

Stroud is not completely satisfied with this reconstruction/diagnosis of Hume's argument. He thinks there is something missing. Stroud offers an alternative *internalist-epistemic* rendering of Hume's argument. I will re-cast Stroud's reconstruction into our terminology and notation, but I will try to remain faithful to the gist of his approach. Stroud thinks that it *isn't merely* the *invalidity of the argument* from E to H that is worrying Hume. Rather, Stroud suggests we should reconstruct Hume's argument in the following, alternative way:

(i) If ϕ believes H on the basis of E, then ϕ's belief in H is reasonable (*viz.*, ϕ is justified in believing H on the basis of E) only if E gives ϕ reason to believe H (*i.e.*, only if E *evidentially supports* H for ϕ, in the context in question). [This presupposes something like a variety of *internalist evidentialism*.]

(ii) E gives ϕ reason to believe H only if ϕ is justified in believing that E gives ϕ reason to believe H. [This is a *second-order*, *internalist* epistemic principle. It is closely related to the so-called "JJ principle".]

(iii) ϕ is not justified in believing that E gives ϕ reason to believe H. [more on this key premise, below]

(iv) Therefore, ϕ is not justified in believing H on the basis of E. [Note: this is (IJBS), not (ISS).]

Stroud's reading is thus *doubly* (internalist) epistemic. Not only doesn't Stroud think that the problem can be addressed by adding a *uniformity principle* $E^{\star}$ to the "enthymematic" $\mathcal{A}$, he doesn't even think adding an *epistemic* rendition of uniformity — that ϕ is *justified in believing* $E^{\star}$ — would address Hume's problem. Stroud thinks Hume is implicitly imposing a second-order epistemic requirement here: that ϕ must be justified in believing *that* E gives ϕ reason to believe H, in order for ϕ's belief in H (on the basis of E) to be justified. This has Hume making a rather deeply (and second-order) internalist epistemic complaint about the kinds of inductive inferences in question. This is an interesting reading of Hume, which surely differs from the "popular" reading. Stroud provides textual (and intuitive) support for it as follows (63):

> ... Hume sometimes expresses the additional requirement for a reasonable inference from the observed to the unobserved by saying that it requires the principle that 'instances of which we have had no experience, *must* resemble those, of which we have had experience', or that the past is a '*rule* for the future' ... that comes close to the claim that one must reasonably believe that what is and has been observed ... gives one good reason to believe certain things about the unobserved, and not just that the observed *is* actually like

> the unobserved. …it is plausible to argue that no one who has observed a constant conjunction between *A*s and *B*s and is currently observing an *A* will reasonably believe on that basis that a *B* will occur unless he also reasonably believes that what he has experienced is good reason to believe that a *B* will occur.

Be that as it may, the key premise in this reading of Hume's argument is clearly (iii). Stroud (pp. 66–67) offers an argument for (iii), which parallels Hume's "dilemma" in the *Enquiry*. Here's my reconstruction:

(a) ($\star$) "ϕ's evidence consists of E (*i.e.*, certain observed (by-ϕ) instances of green emeralds)" *does not entail* (†) "E gives ϕ reason to believe H" (or "ϕ's evidence E *supports* H").

(b) Therefore, any support (from the point of view of ϕ) for (†) must come from a reasonable *inductive* inference (on the part of ϕ) from observed (by-ϕ) instances E' to the truth of "observed (by-ϕ) instances E of green emeralds give ϕ reason to believe H". [The idea here seems to be that because $(\star) \nvDash (\dagger)$, the support claim (†) cannot be known *a priori* by ϕ, *i.e.*, known *merely* on the basis of ($\star$) and ϕ's knowledge of the *meaning* of "reason" (or "evidential support"). And, if (†) cannot be known *a priori* by ϕ in this sense, then *the only other way* (what dilemma does this presuppose?) it can be known (by ϕ) is by a reasonable *inductive* inference (by ϕ) from observed instances E'. Note the sliding between talking about "ϕ's evidence *supporting* (†)" and "ϕ *knowing* (†)". Is this kosher? Also, one may wonder what the *logical* fact $(\star) \nvDash (\dagger)$ has to do with the possibility of ϕ having *a priori knowledge* of (†). Similar issues also arise in connection with premise (4) of the "popular" reconstruction of Hume.]

(c) But, *every* inference from the observed to the unobserved is such that it is reasonable or justified (for ϕ) only if ϕ has reason to believe that (certain) observed instances provide reason to believe a certain statement about unobserved instances.

(d) ∴ In particular, the inference from observed instances E' to the conclusion "observed instances E give ϕ reason to believe H" is reasonable or justified (for ϕ) only if ϕ has reason to believe that (certain) observed instances provide reason to believe a certain conclusion about unobserved instances.

(e) But, that would be "evidently going in a circle, and taking that for granted, which is the very point in question". [Note: Stroud's Humean "circularity" worry is different than the "popular" rendition, above, in (4). Stroud locates the circularity not in any inductive inference to the *non*-epistemic $E^\star$, but rather in an inductive inference to the *epistemic* "observed instances E give ϕ reason to believe H".]

(f) Circular inferences of this kind do not support (or, at least, do not justify) their conclusions.

(iii) Therefore, ϕ is not justified in believing that E gives ϕ reason to believe H.

I think all of this adds-up to a pretty clever reading of Hume's skeptical argument. In Hume's time, internalism in epistemology was dominant. And, Hume's argument does sound (to my ear) like an internalist-style epistemic regress or circularity argument. Moreover, the second-order principle Stroud relies on here does seem plausible from such an internalist perspective. Of course, if you're an externalist, you'll probably just deny premise (ii). But, I don't think that would make much contact with Hume's discussion. A few questions:

- Stroud reconstructs Hume's argument as an argument for (IJBS). But, he presents the "popular" argument as an argument for (ISS), which is a stronger conclusion (although, both entail (IKS), from an internalist perspective). Would his reading be as plausible if its conclusion were (ISS) instead of (IJBS)?
- Compare and contrast *Stroud's* use of *invalidity* in (a)–(b) [pp. 64–66], and the "popular" use of it in (1).
- Note that (IKS) is a much weaker kind of skepticism than either (ISS) or (IJBS). [Recall that Milton attributes (ISS) to Hume, but he attributes *merely* (IKS) to the pre-Humean inductive "skeptics". This increase in strength was distinctive of *Humean* inductive skepticism.] There is a nice connection here to Gettier problems. It has been argued (by Zagzebski) that (internalist) XTB theories of knowledge are open to Gettier problems, *unless X alone entails T*. Such theories of knowledge are called *infallibilist*. What's the relationship between infallibilism (in this sense) and the "popular" reading of Hume?

2 Chapter 4 of Stroud's *Hume*: The Positive Phase

This chapter largely leaves logic and epistemology behind, in favor of the *psychology* of inductive inference. Chapter 4 presupposes Humean inductive skepticism. That is, it presupposes that *epistemic* agents ϕ do *not* engage in "drawing inductive inferences", where this has its usual multiple ambiguities. Most of the chapter seems to be about the psychology of singular predictive induction, along the lines of the inference from E to H we've been discussing so far (although, Stroud sometimes slips into talking about universal induction). In any case, we're now focusing on *how actual* (cognitive) agents ψ *actually do* "draw inductive inferences".

Stroud reconstructs Hume's "positive" psychological approach in a very simple way. He reads Hume as asserting that cognitive agents ψ are prone to infer claims about unobvserved objects, on the basis of observations of objects with "constantly conjoined" properties. For instance, he has ψ's being disposed to make inferences like the (singular predictive) inference from E to H. Toward the end of the chapter, we get the clearest statement of Stroud's reading, in his discussion of Goodman's "grue" example. Stroud grants Hume that actual agents ψ *are* prone (presumably, when situated in some "normal epistemic context" C[1]) to infer (*e.g.*) H from E. He then considers the following alternative "psychologically recalcitrant" hypothesis:

(H') The next observed (by ψ) emerald will be *grue*.

Here, x is *grue* iff either x is green and x has been observed (by ψ) or x is non-green and x has not been observed (by ψ). This is an infamous example, due originally to Nelson Goodman. As we'll see later in the course, Goodman used this example for logical and epistemological purposes, but not for psychological purposes. Interestingly, Stroud puts it to *psychological* use here. His claim seems to be that while actual agents ψ *will* tend to infer H from E (in C), they will *not* tend to infer H' from E (in C). Of course, this is an *empirical* claim. And, one wonders why Stroud is confident that actual agents wouldn't tend to infer H' from E (or be *less* prone to infer H' from E than to infer H from E). I should think that this will depend on the *way* in which subjects are asked about "the inference from E to H'". For instance, *if all ψ is told is that*:

(E'') All emeralds that have been observed so far (by ψ) have been grue.

then wouldn't ψ be quite prone to infer H' from E''? Moreover, given our definition of "grue", E'' *is logically equivalent to* E (with ϕ replaced by ψ, of course). Thus, there is a sense in which this ***is*** *a way of finding out about ψ's tendency to draw* "the inference from E to H'." And, on *this way* of "surveying" subjects, I bet they would be just as prone to infer H' from E'' (which is equivalent to E) as they would be to infer H from E (which is equivalent to E''). On the other hand, if subjects are *also* told what "grue" *means*, then I suspect that they *might* start to become less prone to infer H' from E''/E. Of course, this is an empirical question that is not decidable *ex cathedra*. And, empirical questions like these become more and more subtle, the more information we pack into our "surveys". [We'll see this issue arise again later when we look at contemporary cognitive science experimentation concerning confirmation and induction.] One problem that inevitably arises in this setting is that subjects often presuppose that there must be a "correct" (in a *normative* sense) answer to the survey question(s). As a result, the more we "explain" to our subjects what "grue" means, *etc.*, the more likely we are to implicate to our subjects that it would somehow be "odd" or "incorrect", *etc.*, to infer H' from E/E''. If this happens, then our survey may actually be generating information about what our subjects believe about what a "rational" agent ϕ would (or should) say in response to the survey question, rather than what ψ *actually tends to infer*. As a result, what often happens in these cases is a "misfiring" of the experiment. This is one reason (among many) why this sort of psychological experiment is so tricky.

This trickiness is amplified in the case of "grue", since this is explicitly (and by design) a case in which the *description* of the evidence makes a big difference to how the evidence is perceived. If we describe the evidence in "grue" terms (E''), then we should (in some sense) be led to "predict" that the next emerald observed will be grue, since all of them have been grue so far. But, if we describe the evidence in "green" terms (E), then we should (in some sense) be led to "predict" that the next emerald observed will be *non*-grue (*viz.*, green), since all of them have been green so far. The fact that we are led to "inconsistent" predictions here is supposed to be "paradoxical".[2] But, such "inconsistency" is not as odd as it might sound. In fact, as Hempel emphasized (more on that later), it is rather commonplace in the context of inductive inference. Example: a card c is drawn at random from a standard deck. Let E assert that c is a black card, let H_1 be the claim that c is the ace of spades, and let H_2 be the claim that c is the jack of clubs. In this case, E *supports/confirms* both H_1 and H_2 — in a *probability-raising* sense — even though H_1 and H_2 are *inconsistent*. We'll discuss this and many other aspects of "grue" when get to Goodman in a few weeks.

[1] This turns out to be a *crucial* caveat. As we'll see in a few weeks, such expectations are also sensitive to *epistemic context*.

[2] Note that "the next emerald observed will be grue" and "the next emerald observed will be *non*-grue" are *inconsistent*, but they are ***not*** *logical opposites* (at least, if one assumes a *Russellian* theory of definite descriptions). This is a crucial distinction, from a *probabilistic-confirmation* point of view. Note, also, that "All emeralds are green" and "All emeralds are grue" are *not even logically inconsistent*. These and other fun facts about "grue" and its surrounding myths will be discussed in depth when we get to Goodman.