Scientific Explanation & Scientific Realism: Day 4

- Administrative:
  - Two hour meetings from now on (mercifully)
  - I'll present stuff tonight, but then it's student presentations
  - We need a volunteer for next week (Causal – Lewis &/$\vee$ Salmon)
- Brief Review & Finishing-up From Last Time
  - Overview of Accounts of Probabilistic Explanation, So Far
    - I–S, I–N (my fictional Hempel), S–R, and Railton's D–N–P
- Harman's Account of Pr Explanation (emphasis on "quality")
  - What makes one Explanation "better than" another?
  - Greater Probabilistic Relevance & Greater Conditional Probability?
- Preemption, Probability Raising, and Causal/Explanatory Relevance
- Lewis *vs* Davidson on Causal *vs* Explanatory Relevance

Overview of "Probabilistic Explanation" 1

- A Hempelian I–S explanation of Ga appeals to the fact that a is F, and the statistical frequency r with which G occurs among the F's (known to be K's) to explain why a (known to be a K) is a G. Big r means good explanation.
  - I–S: a (known to be a K) is a G because it is an F and $\Pr(Gx \mid Fx \,\&\, K) = r$.
    - The K must satisfy the *requirement of maximal specificity* (RMS).
    - I–S Explanation is *epistemically relative* (K is supposed to capture "everything we know" about a, *modulo* K's satisfaction of the (RMS))
    - This stems from Hempel's view that "Inductive Logic" is not logical
    - I–S Explanation ignores *probabilistic relevance* of F to G (in K)
    - Bigger r $\Rightarrow$ stronger argument $\Rightarrow$ better I–S explanation (more below)
- An I–N Explanation of Ga-at-t (relative to $\mathcal{B}$-at-t' and probability model $\mathcal{M}$) appeals to the fact that a is F at $t' < t$, and to either the conditional probability *in* $\mathcal{M}$ (Hempel): $\Pr_{\mathcal{M}}(Ga\text{-at-}t \mid Fa\text{-at-}t' \,\&\, \mathcal{B}\text{-at-}t')$, or the likelihood-ratio *in* $\mathcal{M}$ (me): $\dfrac{\Pr_{\mathcal{M}}(Fb\text{-at-}t' \mid Gb\text{-at-}t \,\&\, \mathcal{B}\text{-at-}t')}{\Pr_{\mathcal{M}}(Fb\text{-at-}t' \mid {\sim}Gb\text{-at-}t \,\&\, \mathcal{B}\text{-at-}t')}$ [relevant *vs* non-relevant renditions].

Overview of "Probabilistic Explanation" 2

- I–N Explanation fills a historical gap in the literature:
  - It is *not* epistemically relative, since it is *logical* (*i.e.*, inferential, *analytic*, and based on some probabilistic generalization of deductive entailment)
  - It is capable of capturing intuitions of *probabilistic relevance* (my version)
  - No reference-class or Simpson's problems [no (RMS) or "homogenaeity"]
  - It is not limited by any particular kind or interpretation of probability.
  - However, one still needs to say where the *explanatorily salient* probability models $\mathcal{M}$ come from. Here, we can look to science (a la Railton, below).
  - I–N assumes that science (or the appropriate source of $\mathcal{M}$) can give us *complete enough* probability models (*e.g.*, theories may not always yield $\Pr(F \mid {\sim}G)$, but see I.J. Good's discussion of radioactive decay, on website).
  - I–N is still an *inferential* (logical) approach, and it is based only on *probabilistic* (not causal) relations. Faces preemption cases (see below).
  - Like I–S, I–N suggests that "better arguments are better explanations".

Overview of "Probabilistic Explanation" 3

- S–R provides an explanation of "why this F (a) is a G" consisting of the unconditional statistical probability of G (within F), a homogeneous relevant partition $\{F \,\&\, P_i\}$ of F with respect to G, the conditional statistical probability of G in each cell $F \,\&\, P_i$ of the partition, and a statement of the location of the individual a in a particular cell $F \,\&\, P_k$, *i.e.*, that $Fa \,\&\, P_k a$. More precisely:
  - $\Pr(Gx \mid Fx) = r$
  - $\{F \,\&\, P_i\}$ is a homogeneous relevant partition of F with respect to G
    - Relevant: $\Pr(G \mid F \,\&\, P_i) \neq \Pr(G \mid F \,\&\, P_j)$, for $i \neq j$ [need not be $>$]
    - Homogeneous: There is no *further relevant* partition of F, beyond $\{F \,\&\, P_i\}$.
  - $\Pr(Gx \mid Fx \,\&\, P_i x) = r_i$
  - $Fa \,\&\, P_k a$, for some cell k of the partition $\{F \,\&\, P_i\}$
- Note: S–R inherits the problems of statistical Pr from I–S. It "solves" the reference class problem by building the reference-class into the explanandum.
- So, no (RMS) is needed, but "homogenaeity" is (for Simpson's paradox).
- S–R requires probabilistic *relevance*, but it says nothing about "quality" [needn't even be *positive* relevance]. And, S–R is *not* an inferential account.

Overview of "Probabilistic Explanation" 4

- A Railtonian D–N–P explanation of Ga-at-t appeals to the fact that a is F at $t' < t$, and a Railtonian Pr-law of the form $(\forall x)(Fx\text{-at-}t' \supset \Pr(Gx\text{-at-}t) = r)$.
  - The law is *deduced* from an *indeterministic theory*, together with the salient facts of the case (experimental set-up, *etc.*). This deduction handles what otherwise would have been "reference class problems" and the like. [Note: I–N may say the same thing about its $\mathcal{M}$'s, *if* theories entail them!]
  - The Pr function is interpreted as objective, physical, single-case probability (or propensity). So, Ga is assumed to be indeterministic. [D–N–P differs here from I–S and I–N, which don't assume this.]
  - The "goodness" of the explanation does *not* vary with r. So, the "quality" of a D–N–P explanation does not depend at all on the probability that the explanandum occurs. [This is another key divergence from I–S/I–N.]
  - So, an explanation of Ga would have served "equally well" as an explanation of $\sim Ga$, since *the very same explanans*, and *the very same probability statement* $\ulcorner \Pr(Gx\text{-at-}t) = r \urcorner$ would have appeared in both.

Overview of "Probabilistic Explanation" 5

- On the question of the "quality" of D–N–P explanations, Railton seems to be equivocating. If his account is a D–N one, then the *explanandum* is the *conclusion* of the argument from the Probabilistic Law + The Facts.
- But, then, the explanandum is $\Pr(Ga\text{-at-}t) = r$, *not* Ga. It is true that $\Pr(Ga\text{-at-}t) = r$ is a fact about a, and not a general law (so this is not D–S). But, since this is a D–N, its "quality" is good *iff* it's a *valid argument*.
- Thus, he seems compelled to say that everything that fits the D–N–P pattern is an "equally good" explanation (they are all deductively valid arguments!).
- But, he also wants to say that what we are explaining by giving a D–N–P is Ga, not $\Pr(Ga\text{-at-}t) = r$, since it is the explanation of token events that people are talking about in this literature. Then, his claim about "quality" is dubious.

  …there is a great deal more we could say about [Ga]. Deliberately left out … are innumerable details about the experimental apparatus … and about the epistemic position of the scientific community at the time. These facts are omitted as *explanatorily irrelevant* [to Ga] because they are *causally irrelevant* to the physical [probability] for decay … and to whether or not that possibility was realized.

High Probability *Versus* Probabilistic Relevance

- Irrelevance cases seem to indicate that HP is *not sufficient* for explanation.
- Jeffrey: high probability is *not necessary* for explanation either:

  Consider a genuinely indeterministic coin which is biased strongly (p = 0.9) toward heads when tossed. Suppose that if it is not tossed the coin has probability of 0.5 of being in either the heads or tails position and that whether or not the coin is tossed is the only factor that is statistically relevant to whether it is heads or tails. According to the IS model, if the coin is tossed and comes up heads, we can explain this outcome by appealing to the fact that the coin was tossed (since under this condition the probability of heads is high) but if the coin is tossed and comes up tails we cannot explain this outcome, since its probability is low …The fact that the coin has been tossed is the only factor relevant to either outcome and that factor is common to both outcomes once we have cited the toss …we left nothing out that influences the outcome.
- Woodward: such arguments presuppose that "it is not possible for all of the information that is relevant to some [Ga] to be insufficient to explain it. …It is far from self-evident that this assumption is correct." What's Railton say?

Jeffrey, Railton, "Probabilistic Relevance", and "Relevance to the Probability"

- Jeffrey gives a probability model $\mathcal{M}$ in which $\Pr_{\mathcal{M}}(Ga \mid Fa) > \Pr_{\mathcal{M}}(Ga \mid \sim Fa)$, hence Fa is positively relevant to Ga in $\mathcal{M}$. Note: this implies that Fa is negatively relevant to $\sim Ga$ in $\mathcal{M}$. So, this *can't* be what Railton is doing.
- He is *not* talking about Fa-at-t' being probabilistically relevant to Ga-at-t in some model $\mathcal{M}$. For Railton, the "law" $(\forall x)(Fx\text{-at-}t' \supset \Pr(Gx\text{-at-}t) = r)$ captures Fa's "relevance to the probability of" Ga. This is much different.
- Note the scope ambiguity, yielding two senses of "X is relevant to Y":

  (*i*) X is probabilistically relevant to Y (Jeffrey, Salmon, I–N, Harman)

  (*ii*) X is relevant to the probability of Y (Railton)
- The crucial difference here is that, according to (*i*), if X is positively relevant to Y, then X is not positively relevant to $\sim Y$. But, this is not so for (*ii*), since $(\forall x)(Fx \supset \Pr(Gx) = r) \not\Rightarrow (\forall x)(Fx \not\supset \Pr(Gx) \neq r)$ [*e.g.*, assume nothing is F].
- So, when Railton says "nothing else [but F] is relevant" to G, he just means that nothing else [but F] appears in the antecedent of our derived Pr-law.

Harman on “Better” and “Best” Explanation 1

- Harman seems to be trying to explicate a comparative relation of the form “H_1 is a better explanation of E than H_2 is (in a context K)”. This relation is to be defined in terms of probabilities over H_1, H_2, and E. He gives 5 examples.
- These examples are supposed to illustrate the kinds of probabilistic relations that undergird Harman’s four-place relation $BE(H_1, E, H_2, K)$. There are various important questions we should ask ourselves about Harman’s paper:
  - Is Harman’s account inferential (like I–S, I–N, and D–N–P) or not?
  - What *kind* of probabilities are involved in Harman’s BE?
  - Is Harman endorsing some sort of Prediction/Explanation Symmetry?
  - What does Harman mean when he says he/BE is “empiricist”.
  - Do Harman’s conclusions about his own examples make sense?
  - What similarities and differences can we spot wrt other approaches?
  - How are we to combine the potentially conflicting considerations going into BE? Specifically, must there be a “best” explanation of E (in K)? Can explanations be incommensurable wrt “goodness” in Harman’s sense(s)?

Harman on “Better” and “Best” Explanation 2

- Example 1 (page 168): E = John has the new \$100 bill, H_1 = John won the (fair) coin toss with Sam for the \$100 bill, H_2 = John entered a 200 million to 1 lottery in which the prize was the \$100 bill. And, I’ll assume for simplicity that H_1 and H_2 are mutually exclusive and exhaustive (MEE).
- Harman claims (and wants) that $\Pr(H_1 \mid E) > Pr(H_2 \mid E)$. But, all we are told is that $\Pr(E \mid H_1) = 1/2 \gg Pr(E \mid H_2) = 1/200$ million. In order the get the desired inequality between posteriors, Harman must be assuming something about the priors $\Pr(H_1)$ and $\Pr(H_2)$. He doesn’t need to assume $\Pr(H_1) = \Pr(H_2) = 1/2$. Something considerably weaker will do the trick here.
- But, he does need to assume that $\Pr(H_1)$ is not astronomically less than $\Pr(H_2)$. In particular, he must assume that $\Pr(H_2) - \Pr(H_1) \not> 1 - 1/200$ million. This may seem like a “safe” assumption, but is it warranted?
- Note that there are two important requirements here: $\Pr(E \mid H_1) > Pr(E \mid H_2)$, *and* $\Pr(H_1 \mid E) > Pr(H_2 \mid E)$. This *combines* probability and likelihood-ratio. How are these to be *weighed* in judgments of the form $BE(H_1, E, H_2, \mathcal{M})$?

Harman on “Better” and “Best” Explanation 3

- Example 2 (page 169): E = Terry has the certificate, H_2 = Terry won a 1000 to 1 lottery for the certificate, H_1 = George gave Terry the certificate. We are told that if Terry hadn’t won the lottery, then George would have given him the certificate. So, this a case of *overdetermination* or *preemption* [see below].
- Again, we assume that H_1 and H_2 are (MEE). And, again, Harman claims that $\Pr(H_1 \mid E) > Pr(H_2 \mid E)$. This time, we’re told that $\Pr(E \mid H_1) = \Pr(E \mid H_2) = 1$, and $\Pr(H_2) = 1/1000$. So, here, we *do* know that $\Pr(H_1) > Pr(H_2)$, assuming H_1 and H_2 are (MEE), since all of this implies that $\Pr(H_1) = 999/1000$.
- So, $\Pr(H_1 \mid E) = 999/1000 \gg \Pr(H_2 \mid E) = 1/1000$. But, Harman says that this *shouldn’t* count as a case in which H_1 better explains E than H_2, since H_1 and H_2 are both *probabilistically irrelevant* to E (neither “makes a difference”).
- This drives home the dual aspect (non-relavance and relevance) of Harman’s account. He wants *both* greater probabilistic relevance of H_1 for E, *and* greater posterior probability of H_1 on E in order to have $BE(H_1, E, H_2, \mathcal{M})$.

Harman on “Better” and “Best” Explanation 4

- Example 3 (page 170). E = a sample of 10 in which 7 are defective, H_1 = the batch has 70% defective, H_2 = the batch has 1% defective. Again, H_1 and H_2 are (MEE). And, Harman again claims that $\Pr(H_1 \mid E) > \Pr(H_2 \mid E)$.
- All we really *know* here, I take it (as “empiricists”, and assuming a random sampling with replacement model!) is that $\Pr(E \mid H_1) \approx 0.0022$, and $\Pr(E \mid H_2) \approx 10^{-14}$ — in other words, we have a very large likelihood ratio, which indicates a high degree of relevance of H_1 to E, in this model.
- But, the claim about the posteriors is again dependent on some assumption about priors. This time, he explicitly assumes [$\Pr(H_1) = \Pr(H_2) = 1/2$], which entails this result. It would have sufficed to assume that $\Pr(H_1)$ is not astronomically small. But, once again, why should we assume that?
- This is much like example #1 above, but with the added twist of “random sampling with replacement”. Note how much work the impoverished space of alternative explanations is doing. What if we had 3 or more competitors?

Harman on "Better" and "Best" Explanation 5

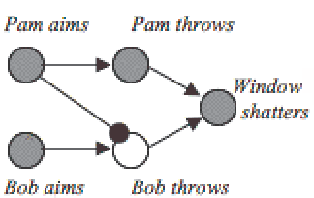
- Example 4 (page 170): E = coin comes up heads 3 times and tails 7 times, H_1 = coin is fair, H_2 = coin is biased 90% toward heads. Again, the H_1 and H_2 are (MEE). Harman claims that $\Pr(E \mid H_1)$ is considerably less than $1/2$. We know that – assuming a Bernoulli probability model of the experiment! – $\Pr(E \mid H_1) \approx 0.0009765$, and $\Pr(E \mid H_2) \approx 0.0004782$. [no model, no Pr's!]
- Again, this means H_1 is highly relevant to E in this model (large likelihood-ratio). But, as Harman points out, $\Pr(E \mid H_1)$ is quite small, and so one cannot expect explanations to make their explananda probable.
- But, after this example, Harman says something a little strange. He says that cases like this (”statistical explanation” cases, as he calls them) are cases in which we're not explaianing "why this happened rather than something else, since the same thing could have led to something else." But, surely, that is also the case when $\Pr(E \mid H_1) > 1/2$. So, I'm not sure I see this distinction.
- Example 5 (page 170): Very similar to #4 – so no need to go through it. But, the point is the same as before. His conditions don't imply $\Pr(E \mid H_1) > r$.

Harman on "Better" and "Best" Explanation 6

- Extracting some "lessons learned" from Harman's paper, we have:
  - $BE(H_1, E, H_2, K)$ *only if both* [query: how are these two to be *weighed*?]:
    - $\Pr(H_1 \mid E) > \Pr(H_2 \mid E)$, *and*
    - $\Pr(E \mid H_1) > \Pr(E \mid H_2)$. [this clause is ambiguous – see below]
  - $BE(H_1, E, H_2, K)$ does *not* entail *either*:
    - $\Pr(E \mid H_1) > r$, for *any* $r > 0$ [divergence from I–S], *or*
    - H_1 explains "why E happened rather than something else" [huh?].
  - Harman tries to make it sound like $BE(H_1, E, H_2, K)$ does not depend sensitively on the prior probabilities of H_1 and H_2. But, this is misleading:
    - In his examples, he makes the likelihood ratios SO large, that the effect of the priors is effectively "swamped". What happens with smaller LRs?
    - His examples involve only 2 (MEE) alternative explanations. When there are 3 or more, the effects of the priors can be much less intuitive.
  - Using only 2 (MEE) alternatives also conflates $\Pr(E \mid {\sim}H_1)$ and $\Pr(E \mid H_2)$. There are really 2 "likelihood ratios" in here. Which one does he intend?

Preemption, Probabilistic Relevance, and Causal & Explanatory Relevance

Suppose that Pam and Bob each aim a brick at a window. Pam throws and shatters the window, while Bob holds his throw on seeing Pam in action (*i.e.*, because she aims). It seems Pam's throw caused the window to shatter — her brick is what crashes through the glass. But it does not need to be the case that Pam's throw raised the probability of the shattering — if Bob is a more reliable vandal, then Pam's throw might even have made the shattering *less* likely.



- The filling of a circle (not to be confused with the *highlighting* of a circle) represents the event occurring.
- The link represents a *prevention*.

- Responses: (1) Hold fixed *intermediary causal factors* (Lewis, Eells, Yablo). If one *holds fixed* the fact that Bob holds his throw, then Pam's throwing *does* raise the probability of the window shattering. (2) Require *precision* in the effect (Lewis, Paul). Then, Pam's throwing will be the cause of the window shattering *in a precise way* (different from the *precise* effect of *Fred's* throw).

Lewis *vs* Davidson on Causal *vs* Explanatory Relevance

- In "Causal Relations", Davidson offers an account of causal claims which has the consequence that it is impossible for the following claim to be true:
  - It was the bolt's giving way suddenly, and *not* the bolt's giving way (*simpliciter*) that caused the bridge to collapse.
- Davidson says that, if we think this is true, we are conflating explanation and causation. According to Davidson, the bolt's giving way suddenly can be explanatorily relevant to the collapse even if the bolt's giving way (simpliciter) is not, but the same cannot be said about their causal relevance.
- On the other hand, Lewis (2000) offers an account of causation which implies that very minute alterations in a cause will eliminate a great many of its causal powers, and create a great many new ones. He argues that intuitions to the contrary conflate causation and explanation. Minute changes may not change *explanatory* relevance, but they will inevitably change *causal* relevance.
- Lewis uses this distinction to rule-*in* certain *un*intuitive causes as "causally (but not explanatorily) relevant", while Davidson uses it to rule-*out* certain *in*tuitive causes as "explanatorily (but not causally) relevant". Interesting.