Four Decades of Scientific Explanation

- Administrative:
  - Try to turn in first drafts soon (few have not) ...
  - I will (try to!) return comments within a week of receipt.
  - Back to normal office hours this week ...
- The Deductive-Nomological (D-N) Account of Scientific Explanation — The First "Serious" Theory of Explanation
  - Review of the D–N Account
  - Some Problems/Issues Raised by D–N
  - A host of troublesome and important examples
- On to "Decade #2" — I–S Explanation
  - I–S as a generalization of D–N
  - Some special problems for I–S

The D–N Account of Scientific Explanation — Quick Summary

- $\langle T, C\rangle$ is a D-N explanation of E (a singular sentence) *iff*
  1. T is true and essentially general and C is true and singular,
  2. E is derivable from T and C jointly, but not from C alone.
  3. T is compatible (*i.e.*, logically consistent) with at least one class of basic sentences which has C but not E as a consequence.
- Alleged counterexample:
  - Let $T = (\forall x)Fx$ (*e.g.*, everyone is imperfect), and let $E = Ha$ (*e.g.*, Hempel is male). T is irrelevant to (*i.e.*, T can't explain) E.
  - $T \vDash T' = (\forall x)(\forall y)[Fx \vee (Gy \rightarrow Hy)]$. Let $C = (Fb \vee {\sim}Ga) \rightarrow Ha$.
  - Kaplan, *et al* show that $\langle T', C\rangle$ is a D–N explanation of E. They think this is counter-intuitive … Do you agree with this claim?
  - Next, I will say a bit more about this alleged counterexample, and one proposed way to avoid it (there's a good paper topic here).

The D–N Account of Scientific Explanation — "Counterexample" I

- To establish that $\langle T', C\rangle$ is a D–N explanation of E, we must show that:
  1. T' is true and essentially general and C is true and singular. T' says "every pair of people (x, y) is such that either x is imperfect or y is a non-philosopher or y is male." This is general and true [since $(\forall x)\ Fx$)]. C says "Oppenheim is perfect or Hempel is a philosopher or Hempel is male." This is singular and true (since Ha).
  2. $T' \mathbin{\&} C \vDash E$, but $C \nvDash E$. Salmon shows $T' \mathbin{\&} C \vdash E$ in a natural deduction system for first order logic. Can you give a simpler proof? It is clear that $C \nvDash E$, since "Oppenheim is perfect" $\vDash C$, but $\nvDash E$.
  3. T' is compatible (*i.e.*, logically consistent) with at least one class of basic sentences which has C but not E as a consequence. To see this, choose (as Salmon does) the set of basic sentences: $\{{\sim}Fb, Ga\}$. Can you see why T' is consistent with ${\sim}Fb \mathbin{\&} Ga$? And, can you see why ${\sim}Fb \mathbin{\&} Ga \vDash C$, but ${\sim}Fb \mathbin{\&} Ga \nvDash C$? See Salmon for the former.

The D–N Account of Scientific Explanation — "Counterexample" II

T'. Every pair of people (x, y) is such that either x is imperfect or y is a non-philosopher or y is male.

C. The pair (Oppenheim, Hempel) is such that either Oppenheim is perfect or Hempel is a philosopher or Hempel is male.

E. $\therefore$ Hempel is male.

- Why do we have doubts about thinking of this as an *explanation* of E?
- Kim suggests a "fix" — add the following condition to D–N definition:

  4. E must not entail any conjunct in the conjunctive normal form of C.

- In this case, the conjunctive normal form (CNF) of C is:

$$(\sim Fb \vee Ha) \,\&\, (Ga \vee Ha)$$

- So, E entails *both* conjuncts of the CNF of C. What are the *consequences* of adding Kim's (4)? **Hint**: use $\mathcal{L}$: $Axy \approx \sim Fx \vee Hy$, $Bx \approx Gx \vee Hx$, and $Dx \approx Fx \vee Hx$. What are C, E in $\mathcal{L}$? Does E entail any conjunct in the CNF of C in $\mathcal{L}$? Why not simply $E \nvDash C$?

Issues Raised by The D–N Account of Explanation

- Things needed to complete the D–N Account:
  1. Explications of model(s) of *probabilistic* or *statistical* explanation
  2. An adequate (D–N) account of the explanation of *laws*. On the current account, a derivative law L can be "explained" by the conjunction $L \,\&\, L'$, for *any* L', *no matter how irrelevant to* L' *may be to* L.
  3. A good explication of the concept of a *qualitative* predicate ("grue"?).
  4. A good explication of the concept of a law of nature.
- Potential problems within the underlying D–N framework:
  1. Are (*all*) explanations *arguments*, as H *&* O assume?
  2. Must all explanations make essential use of *law*(s) of nature?
  3. According to H *&* O, all (D–N) explanations are (potential) (H–D) predictions, and *vice versa*. Is this *symmetry thesis* correct?
  4. According to H *&* O, *causality* plays no essential role in the scientific explanation of particular, token events. Is this correct?
  5. Must the explanans of a good explanation be (*literally*) true?

Nine Famous Problematic Cases for D–N (*&* other theories!)

1. **The Eclipse**: One can D–N-explain a current total eclipse, using (say) Newton's laws of motion, together with past positions of the earth, sun, moon. But, one can also D–N-explain a current eclipse by appeal to NL plus *future* positions! Should this count as an *explanation*?

2. **The Flagpole**: We may D–N-explain the length of a shadow cast by a flagpole using certain laws of optics/geometry, together with the position of the sun in the sky, etc. But, we can also D–N-explain the height of the flagpole using the same laws, together with the length of the shadow and the position of the sun! Is this an *explanation*?

3. **The Barometer**: A falling barometer (together with the appropriate meteorological laws) can H–D-predict an approaching cold front. Thus, one could also D–N-explain the approach of the cold front using the barometer's falling, together with these same meteorological laws.

4. **The Moon and the Tides**: The (general and lawlike) *correlation* between the moon's position and the tides was well known for centuries before Newton's gravitational theory was known. So, H–D-predictions, and D–N-explanations of the tides were constructible by these ancestors of Newton. But, arguably, until the *causal story* behind the tides was told, no *legitimate* explanation was really available.

5. **Syphilis and Paresis**: Only people who have had syphilis can contract paresis. But, only a small fraction (around 25%) of syphilis patients contract paresis. It seems quite *explanatory* to say that a person got paresis *because* they had syphilis. But, this cannot be said on a D–N account (which requires *deduction* of each token case). Similar examples arise surrounding quantum-mechanical phenomena.

6. **The Hexed Salt**: Why did this sample of table salt dissolve in this cup of water? Because a person wearing a funny hat mumbled some non-sense syllables and waived a wand over it. That is, the table salt dissolved because it was hexed. And, it is a *law* that all hexed table salt dissolves when placed in water. This fits the D–N pattern …

7. **Fred Fox on the Pill**: Fred Fox (a male) has not become pregnant during the past year because he has faithfully consumed his wife's birth control pills. And, any male who faithfully takes birth-control pills will avoid becoming pregnant. This also fits the D–N pattern.

8. **Joint effects of a common cause**: Consider two TVs receiving a common broadcast (with one TV farther from the source). We can use the closer TV to *predict* what will be seen on the farther TV. But, does this *explain why* this is seen on the farther TV? What does this example suggest about the explanation/prediction *symmetry thesis*?

9. **Explanation by false/idealized theories**: We use Newton's theory all the time to explain various phenomena. But, we know Newton's theory is *false*. Moreover, for all we know, all of our current scientific theories are also false (in some subtle and as yet unseen way). Does this mean none of our current scientific explanations are good ones?

Think about the two directions of the explanation/prediction symmetry thesis [(potential) explanation $\Leftrightarrow$ (potential) prediction]. Is either correct? How does this depend on choice of theories of explanation/prediction?

Decade #2 — The Birth of Statistical Explanation

- H *&* O were aware that their D–N account (as originally stated) left nor room for the explanation of either (1) statistical laws, or (2) token events which cannot be derived from any theory (but on which some theory + auxiliaries/initial conditions may *confer a probability*).
- Hempel's first alteration (a very minor one) was to expand the explanation of *laws* (by more general laws) in D–N to the case of *statistical laws.* This led to the *Deductive-Statistical* (D–S) model.
- On the D–S model, a statistical law may be explained by appeal to more general laws (which may be statistical or universal).
- **Example**: we may derive the half-life of uranium-238 from the basic laws of quantum mechanics (together with the height of the potential barrier surrounding the nucleus and the kinetic energies of the alpha particles within the nucleus). D–S as a mere *variant* of D–N.
- Same problems as D–N (*plus* statistical laws & qualitative predicates!)

Inductive–Statistical Explanation I

- Hempel offered an inductive generalization of his D–N model of explanation. He called this model the Inductive–Statistical (I–S) model.
- Hempel updated his four high-level desiderata accordingly:
  1. An explanation is an argument having correct logical form (either deductive or inductive — Skyrmsian "strength" idea).
  2. The explanans must contain,. essentially, at least one general law (either universal or statistical).
  3. The general law must have empirical content.
  4. The statements in the explanans must be true.
- Hempel quickly realized that a fifth adequacy condition must be added, because of the *non-monotonicity* of "inductively strong" arguments:
  5. (RMS) The requirement of *maximal specificity*.

Inductive–Statistical Explanation II

- The simplest schema for an I–S explanation would be:

$$\begin{array}{l} \Pr(Gx \mid Fx) = r \\ Fb \\ \hline\hline Gb \end{array} \qquad [r]$$

- Here, "$\Pr(Gx \mid Fx) = r$" is a statistical law which says that the relative frequency of Gs among Fs is r. And, "$[r]$" indicates the "inductive strength" (in a Skyrmsian sense) of the argument.
- **Example**: John Jones (b) recovers quickly (Gb) from Strep (Fb). Most strep infections (Fx) clear up quickly (Gx) when treated with penicillin (Hx). Thus, we have the following I–S explanation of the fact that Gb:

$$\begin{array}{l} \Pr(Gx \mid Fx \,\&\, Hx) = r \approx 1 \\ Fb \,\&\, Gb \\ \hline\hline Gb \end{array} \qquad [r \approx 1]$$

Inductive–Statistical Explanation III

- But, what if we were to subsequently learn that John Jones was infected with a *penicillin-resistant* strain of Strep (Jb)? Plausibly, this would lead to a "strong" I–S explanation of $\sim Gb$, as follows:

$$\begin{array}{l} \Pr(\sim Gx \mid Fx \,\&\, Hx \,\&\, Jx) = r_1 \approx 1 \\ Fb \,\&\, Gb \,\&\, Jb \\ \hline\hline \sim Gb \end{array} \quad [r_1 \approx 1]$$

- This illustrates the (well-known to you all by now!) *non-monotonicity* of inductive inferences. We now have two strong arguments whose conclusions are logically incompatible! Such a thing is *unheard of* in the realm of deductive inference (and D–N explanation)!

- In inductive *logic*, we usually add the requirement of *total evidence*. That is, we usually add the requirement that *no additional evidence that would change the degree of support is available at the time*. This will *not* work here, since the *explanandum* is *known*! What shall we do?

Inductive–Statistical Explanation IV

- Hempel adds the *requirement of maximal specificity* (RMS). Here, we assume that $\mathbf{P}$ is the conjunction of all of the premises of the I–S explanation, and K is the available background knowledge.

(RMS) If $\mathbf{P} \,\&\, K$ implies that b belongs to a class F_1 and that F_1 is a subclass of F, then $\mathbf{P} \,\&\, K$ must also imply a statement specifying the statistical probability of G in F_1, say

$$\Pr(G \mid F_1) = r_1.$$

Here, $r_1 = r$ unless the probability statement in question is simply a theorem of probability theory (proper).

- The "unless" clause in (RMS) is there to block the trivialization which might otherwise arise from the fact that we know from K that b is both F and G. But, of course, probability theory (proper) implies that $Pr(Gx \mid Fx \,\&\, Gx) = 1$. But, this should not undermine the explanation.

Inductive–Statistical Explanation V

- Note, deductive arguments *automatically* satisfy (RMS). *Why*?
- According to Hempel, the *non-monotonicity* of inductive inferences leads, inevitably, to the *epistemic relativity of statistical explanation*:

  > "The concept of statistical explanation for particular events is essentially relative to a given knowledge situation as represented by a class K of accepted statements."

- Skyrms seems to take a similar attitude in this regard. I can see clearly that inductive inferences are *contextual* (or *indexical*), but why does that force them to be inherently *epistemic*? We don't take the same attitude toward special relativity (*versus* Newtonian theory). We would say that things like "simultaneity" and "velocity" are *contextual* (*i.e.*, that they are relative to a *frame of reference*), but we *don't* seem to think they depend on *what we know*. Why the different attitude when it comes to inductive inference (*versus* deductive inference)?

Inductive–Statistical Explanation VI

- You can already guess what some of the problems with I–S explanation are going to be. As we saw with Skyrms' discussion of "strength", the most obvious difficulties involve the issue of *relevance*.
- "$\Pr(X \mid Y)$ is high" is neither necessary nor sufficient for "Y is relevant to X". The Fred Fox example, and the paresis/syphilis example (above) show that *relevance* is an important aspect of explanations.
- Both the D–N and the I–S accounts suffer from the relevance problem. The hexed salt example shows that D–N is also vulnerable here.
- While high probability is neither necessary nor sufficient for high explanatory power, it may still be plausible that (*other things being equal* — e.g., assuming that there *is* relevance, etc.) high*er* probabilities do lead to *better* explanations. Our own Michael Strevens discusses this point nicely in his recent paper "Do Large Probabilities Explain Better?". He will be giving us a guest lecture on T 11/20.