

# Session 8. Inference to the best explanation

Dr. Neil Barton

n.a.barton@ifikk.uio.no

FIL4200: Epistemology and Philosophy of Science

13. February 2023

Last week we looked at different **theoretical virtues**.

This week we'll discuss:

**Inference to the best explanation** (or **IBE**) states that we should accept the **best explanation** for the given data (or when comparing theories, that one theory is **better justified** than another when it provides a **better explanation**).

## 1 A few examples

A example from Bas van Fraassen (who eventually comes to criticise IBE):

I hear scratching in the wall, the patter of little feet at midnight, my cheese disappears—and I infer that a mouse has come to live with me.

I haven't made any **repeated observations** as with induction.

But we also use abductive inference in a scientific context:

**Phlogiston vs. Oxygen.** The **phlogiston theory** states that combustible matter contains a substance **phlogiston**, that is emitted when it combusts. The **oxygen theory** states that there is a substance **oxygen**, naturally present in the air, that **combines** with combustible matter during combustion.

**Example.** The bell jar.

**Example.** Gas capture and mass change.

**Phlogiston theory can** incorporate gas capture, but only when augmented with a notion of **negative weight**.

The **oxygen theory** is now regarded as the **best explanation** for combustion.

IBE can be used even when we don't take there to be a

**Wave vs. particle theory of light.** Is light a **wave** or composed of **particles**?

Large debate about this pointing to wave-like aspects of light (e.g. diffraction effects).

Various experiments (e.g. double-slit experiment) now suggest the wave-particle duality of light.

IBE can now be used to argue that light exhibits both wave-like and particle-like behaviours.

**Question.** How much is IBE **actually** used? Or are there more subtle factors (e.g. psychology of proponent) often at play in theory choice (e.g. see White's paper)?

**Background Question.** There's an **enormous** literature on scientific explanation (e.g. Deductive-Nomological, Statistical Relevance, Bayesian methods, Casual Mechanical, explanation *as* unification, pragmatic accounts). To what extent should we fix the background account?

**Questions? Clarifications? Comments?**

## 2 Thagard on simplicity and consilience, and analogy

### 2.1 Consilience

**Consilience.** One theory is more consilient than another if it explains more classes of facts than the other does.

Consilience can also be *dynamic*:

**Dynamic consilience.**  $T_1$  is more dynamically consilient than  $T_2$  iff  $T_1$  has succeeded in adding more to its set of classes of facts explained than  $T_2$ .

**Question.** Should dynamic consilience be viewed as **aiming at the truth?** (e.g. as in the White paper)

**Note:** Different ways of explicating consilience: Cardinality vs. subsets.

**Note:** Cardinality can trivialise if we take a theory to explain infinitely many facts, so we should probably take a finitary notion of facts to be explained.

**Note:** Cardinality is also problematic depending on how we *carve up* the facts.

**Question.** What is a class of facts?

Thagard writes:

The problem is merely pragmatic, concerning the way in which, in particular historical contexts, the scientific corpus is organized

**Question.** How should they be carved up?

**Question.** What about classes of facts across different **disciplines?**

**Question.** What about across different **times?**

**Question.** Is there enough shared for this to be meaningful here?

**Question.** What about the "absolutist" vs. "relativist" picture of facts?

**Question.** Relationship between unification from last week?

### 2.2 Simplicity

**Problem:** We can achieve consilience by beefing up the auxiliary assumptions.

Here Thagard wants to appeal to simplicity as tempering consilience:

An ad hoc hypothesis is one that serves to explain no more phenomena than the narrow range it was introduced to explain. Hence a **simple theory** is one with few ad hoc hypotheses (p. 87)

Constrast Keas:

**Simplicity:**  $T$  explains the same facts as rivals, but with less theoretical content.

Does Thagard's notion of simplicity help elucidate the notion of theoretical content (we worried that it might be a little underspecified)?

Possible analysis from Sober:

Sober defines simplicity as informativeness, where a hypothesis  $H$  is more informative than  $H'$  with respect to a question  $Q$  if  $H$  requires less extra information than  $H'$  to answer  $Q$ .

An explanation is simpler the fewer the initial conditions required in the deduction of the explanandum from the hypothesis.

I don't think so: Both will require qualitative analysis (as Thagard notes—we can just conjoin the auxiliary assumptions).

**Question.** Are syntactic analyses of simplicity doomed as Thagard suggests?

In mathematics we have some good analyses of syntactic simplicity (e.g. via quantifier alternations).

**Ontological economy.** Thagard claims that ontological economy is only interesting insofar as it conforms with simplicity and consilience.

This contrasts with what we saw last week, where it was argued that sometimes we **can** have ontological economy playing a role. (e.g. Multiple unnecessary particles.)

**Question.** What do we think of ontological economy in this light? (Can't there be interrelationships here?)

## 2.3 Analogy

Let's suppose that a  $A$  exhibits  $P, Q, R$ , and we **know** that  $P, Q$ , and  $R$  are **explained** by  $A$ 's having  $S$ .

Now we know that  $B$  exhibits  $P, Q$ , and  $R$ . Is  $B$  having  $S$  a **promising explanation**?

e.g. Example wave theory of sound vs. wave theory of light, both exhibit diffraction properties.

Thagard argues that there is a relationship with **modelling** here.

But the idea that modelling is analogy is **very** controversial.

**Question.** What about just having a view of there being an underpinning structural similarity?

(Contrast with Thagard's complaints about ontological economy, consilience, and simplicity.)

**Question.** Is there a psychological factor here? (Is such a factor unwelcome? e.g. pragmatic accounts of explanation...)

## 2.4 Weighting Thagard's virtues?

These virtues can be in tension with one another (e.g. simplicity and consilience wrt. auxiliary assumptions, analogy and a complex theory).

**Question.** Is there a way of weighting the virtues?

**Question.** Is there any priority on the virtues? (e.g. Is analogy weaker?)

## 3 Does truth explain?

One way of viewing IBE is as truth-directedness.

e.g.1. White provides an analysis of the epistemic advantage of prediction over accommodation via reliable aiming at the truth.

e.g.2. Putnam's no-miracles argument and IBE.

**Question.** Should we view these kinds of IBE as needing truth or just empirical adequacy? (Note: White is sensitive to this issue.)

**Challenge.** Provide anti-realist account of explanation that doesn't rely on truth.

A different problem for IBE and truth is identified by Nancy Cartwright.

**Ceteris paribus laws.** Cartwright points out that almost every law we have has ceteris paribus conditions (where these specify the range of application for the law).

e.g. Geraniums planted in warm soil, Snell's law for angles of incidence and refraction in isotropic media (refractive index same in all directions).

Ceteris paribus laws are *false* without the additional conditions, but can be explanatory.

e.g. Snell's law applying for *nearly* isotropic media.

**Question.** What about if we adopt the "patchwork" pluralistic model (that we'll look at later)?

**Question.** What about a notion of truth-like-ness or almost-structural-similarity?

**Note:** Latter idea is promising but hard to cash out.

## 4 Two kinds of epistemology

**Traditional epistemology.** Looks at concepts like **knowledge**, **belief**, and as we've seen **ignorance** and **suspension of judgement**.

**Bayesian epistemology.** Looks at **credences** and how we should **update** them given new evidence.

**Bayes Theorem.** States that the probability of the **hypothesis** given the **evidence** is equal to the **probability of the hypothesis times the probability of the evidence given the hypothesis** divided by the **probability of the evidence**.

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)}$$

**Observation.** Bayesian epistemology has been **empirically successful** (wide variety of applications, from treasure hunting, to scientific analysis, to warfare, to machine learning).

**Note:** Deep philosophical **question**—how are the two approaches to epistemologies related?

**Question.** Bayesian epistemology seems to make **no** mention of the best explanation. But might the theoretical virtues and IBE be related to Bayesian epistemology? See Lipton *Inference to the Best Explanation* (the 2004 version) here.

Thus conceived we can consider how different **theoretical virtues** might relate to the Bayesian approach epistemology.

## 5 Keas' theoretical virtues

**Evidential accuracy:** A theory T fits the empirical evidence well (regardless of causal claims).

**Causal adequacy:** T's causal factors plausibly produce the effects (evidence) in need of explanation.

**Explanatory depth:** T excels in causal history depth or in other depth measures such as the range of counterfactual questions that its law-like generalizations answer regarding the item being explained.

**Internal consistency:** T's components are not contradictory.

**Internal coherence:** T's components are coordinated into an intuitively plausible whole; T lacks ad hoc hypotheses-theoretical components merely tacked on to solve isolated problems.

**Universal coherence:** T sits well with (or is not obviously contrary to) other warranted belief.

**Durability:** T has survived testing by successful prediction or plausible accommodation of new data.

**Fruitfulness:** T has generated additional discovery by means such as successful novel prediction, unification, and non ad hoc theoretical elaboration.

**Applicability:** T has guided strategic action or control, such as in science-based technology.

**Beauty:** T evokes aesthetic pleasure in properly functioning and sufficiently informed persons.

**Unification:** T explains more kinds of facts than rivals with the same amount of theoretical content.

**Simplicity:** T explains the same facts as rivals, but with less theoretical content.

**Question.** Are there further relationships between what we've discussed this week and these theoretical virtues?

## 6 Housekeeping

**NO SEMINAR NEXT WEEK!** We meet next on 6th March.

**Presentation.** Would anyone like to present?