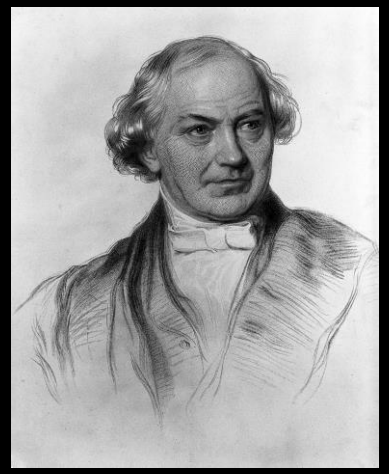
Paul Thagard,
“The Best Explanation: Criteria for Theory Choice”

Recall: Inference to the Best Explanation (IBE), also called "abduction," is classified by Norton as a type of *Hypothetical Induction*

Principle	Ability to entail the evidence is a mark of truth.
Weakness	Too indiscriminate. Frivolous conjunction: A&B entails A; so A confirms B, for any B.

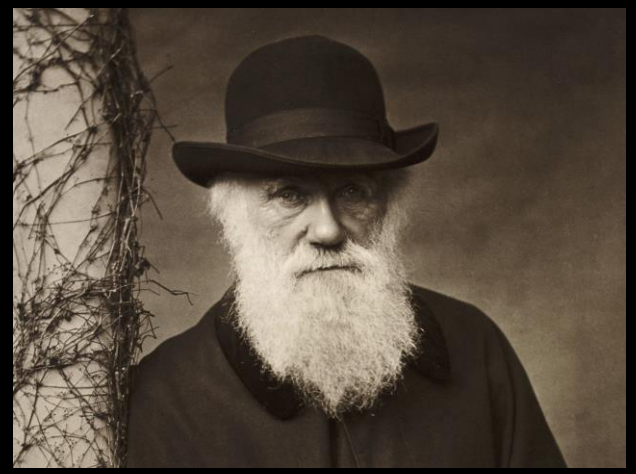
Elaboration	**E confirms H if H (and auxiliaries) entail E AND…**
Abduction: Inference to the best explanation (Pierce, Harman, Lipton)	…H is the best explanation.

The illustrious history of IBE:



Whewell

Pierce



Darwin

Lavoisier

Huygens

and many more …

Sounds good, but how do we decide which hypothesis is the *best* explanation of the evidence?

Thagard promises a set of **criteria** for deciding.

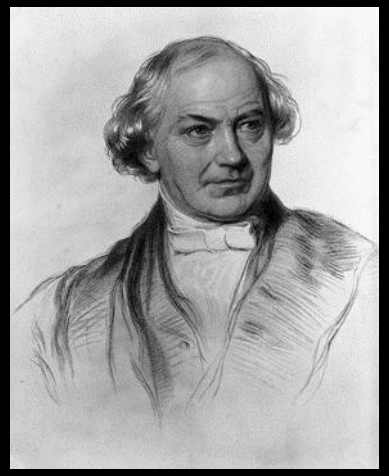
"…the criteria furnish a **comprehensive account of the justification of scientific theories**"

Thagard's **criteria**:

- Consilience
- Simplicity
- Analogy

Consilience

“one theory is more consilient than another if it explains *more* **classes of facts** than the other does”



William Whewell

Consilience

"one theory is more consilient than another if it explains *more* **classes of facts** than the other does"

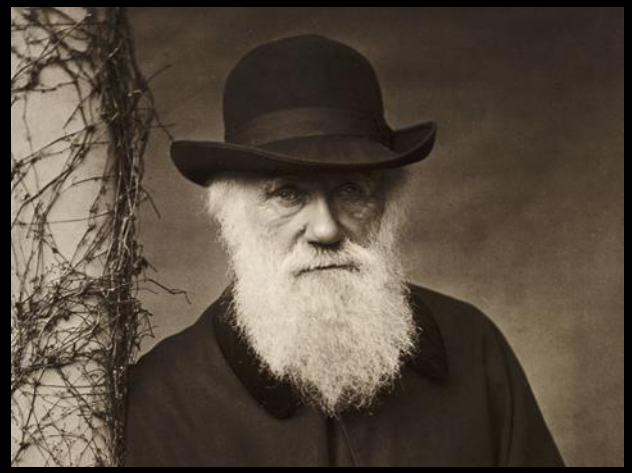
Examples:

1. Lavoisier's **oxygen theory of combustion** explained phenomena that phlogiston theory couldn't (increase in mass of combusted matter)
2. Darwin's **theory of evolution** explained anatomical homologies, vestigial organs, and biogeographical patterns that special creation couldn't
3. Huygen's **wave theory of light** explained linear propagation, reflection and refraction of light better than Newton's corpuscular theory (esp. after additions by Young, Fresnel in 19th C)
4. **Newtonian mechanics** explained motion of planets, comets, tides, etc.
5. **Quantum mechanics** explains spectral frequencies of atoms, magnetism, solid state of matter, photoelectric effect, etc.

Consilience

"one theory is more consilient than another if it explains *more* **classes of facts** than the other does"

Thagard notes that Darwin used the "**classes of facts**" terminology:



"It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large **classes of facts** above specified."

Consilience

"one theory is more consilient than another if it explains *more* **classes of facts** than the other does"

Problem: How do we count "classes of facts"?

Thagard's response: The problem is just a pragmatic one; the answer will depend on the historical context, namely, how the scientific corpus is organized at a given time.

Examples:

For Newton and Huygens, reflection and refraction phenomena were different classes of facts

For Darwin, biogeography was one class of facts and comparative anatomy was another

Consilience

"one theory is more consilient than another if it explains *more* **classes of facts** than the other does"

Problem: What if H1 explains more overall than H2, but H2 explains some things that H1 doesn't?

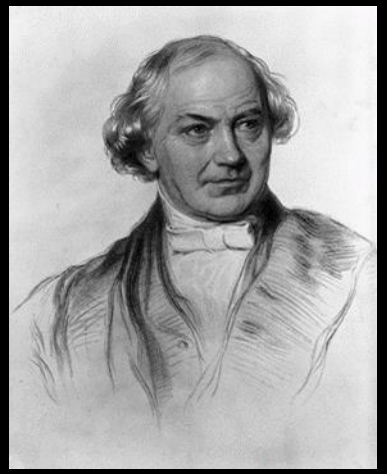
Thagard's suggestion: "decisions concerning the best explanation must be made according to what theory explains the most important facts, or on the basis of other criteria discussed below"

Thagard points out that consilience is more applicable to historical examples than Hempel's syntactic, logical account of the "systemic power" of theories.

Deductive consequences of a theory are not enough; we need to judge the **variety** of the consequences of a theory, and variety doesn't lend itself to a neat formal characterization.

Whewell's notion of consilience is **dynamic**: a theory is consilient when it explains facts it **wasn't invented to explain**.

Thagard: When theories are expanded to explain more facts, but the expansions explain only the facts they were introduced to explain, they are *ad hoc*.



A final remark about consilience:

The most consilient theory would be one that explains *everything*. But isn't it problematic when a theory can explain everything? (recall Popper)

Thagard suggests that an increase in consilience should be bounded by tradeoffs in other desirable criteria, such as precision and **simplicity** …

Simplicity

Theories explain evidence in conjunction with *auxiliary hypotheses*.

“An auxiliary hypothesis is a statement, not part of the original theory, which is assumed in order to help explain” the evidence.

Theories which explain more (classes of) facts with fewer auxiliary hypotheses are **simpler** (and thus more explanatory) than others.

Simplicity

Examples:

1. Newton’s corpuscular theory of light needed separate auxiliary assumptions for each phenomenon it explained, but the wave theory uses the same assumptions for each application
2. Phlogiston theory needed multiple (inconsistent) assumptions to explain phenomena the oxygen theory more easily explained

Simplicity

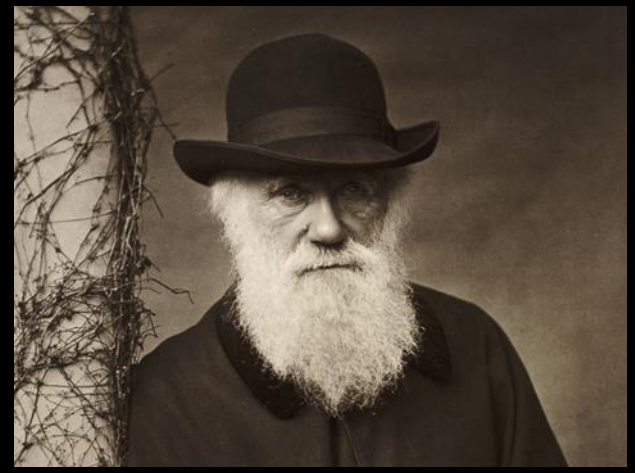
“These examples show how simplicity puts a constraint on consilience: a simple consilient theory not only must explain a range of facts; it must explain those facts without making a host of assumptions with narrow application.”

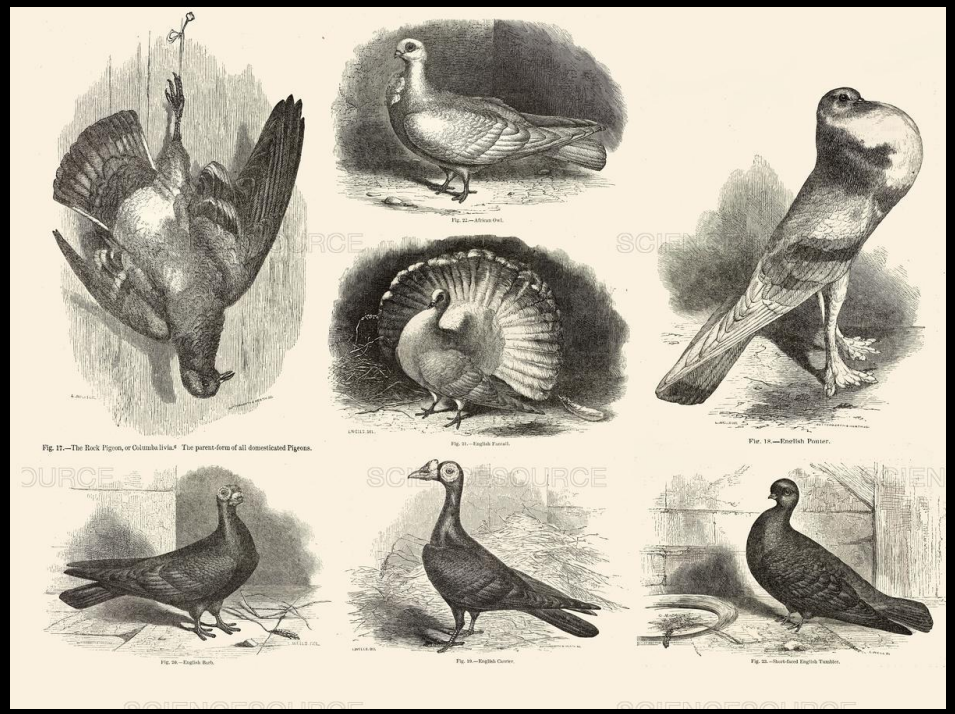
Another way to say this is that theories with fewer *ad hoc* hypotheses are simpler—and thus more explanatory—than those with more.

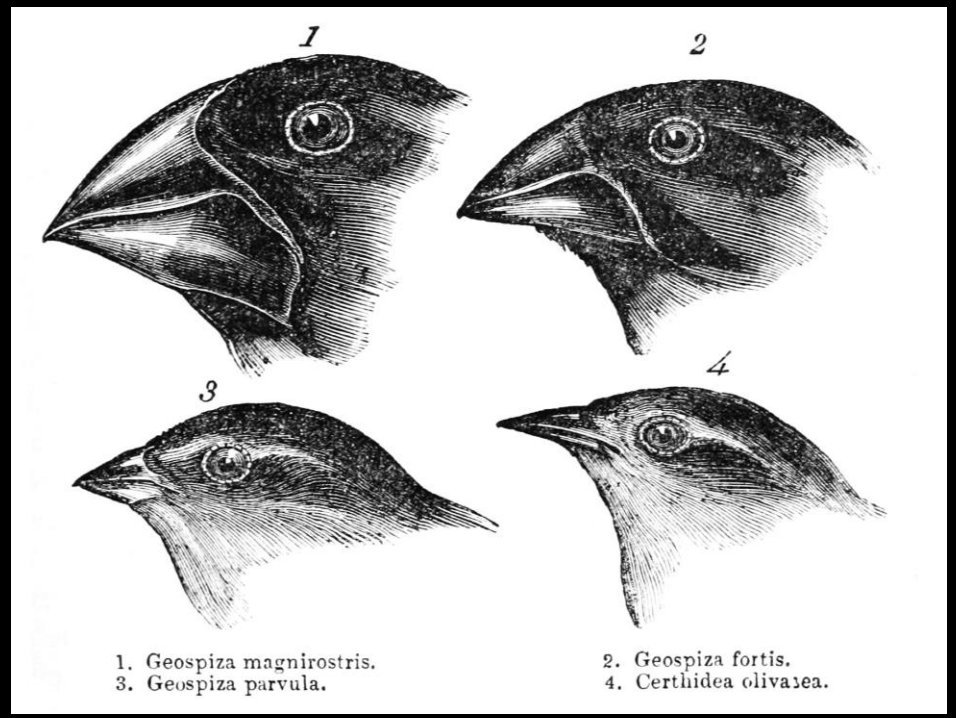
Analogy

Examples:

- Darwin’s analogy between selective breeding and natural selection



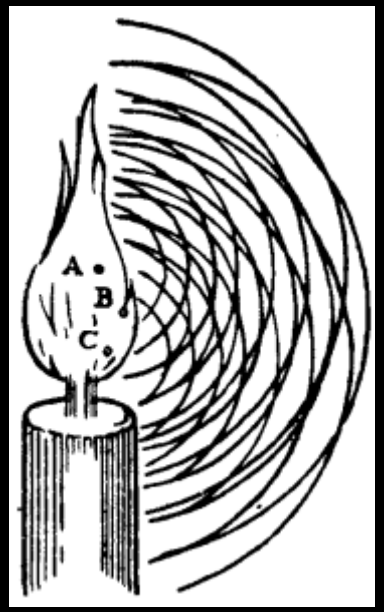




Analogy

Examples:

- Huygens' (and Young's and Fresnel's) analogy between sound and light



Analogy

At first, it isn’t obvious what analogy has to do with explanation.

A simple representation of analogical argument:

(AA) *A* is *P*, *Q*, *R*, *S*.
B is *P*, *Q*, *R*.
∴ *B* is *S*.

Analogy

At first, it isn't obvious what analogy has to do with explanation.

A simple representation of analogical argument:

(AA) *A* is *P*, *Q*, *R*, *S*.
B is *P*, *Q*, *R*.
∴ *B* is *S*.

The problem with (AA) is the presence of *disanalogies*, properties not shared by the source and target. How do we weigh analogies and disanalogies against each other?

Analogy

At first, it isn't obvious what analogy has to do with explanation.

A simple representation of analogical argument:

(AA) *A* is *P*, *Q*, *R*, *S*.
B is *P*, *Q*, *R*.
∴ *B* is *S*.

Thagard suggests that explanation can play a crucial role here. If *A*'s having properties *P*, *Q*, and *R* is *explained* by its having property *S*, then we may conclude that *B*'s having *S* would explain *B*'s having *P*, *Q*, and *R*.

Analogy

But it isn’t just that explanatory relations support analogies. There is a connection in the opposite direction, too: analogy supports explanation.

“We get increased understanding of one set of phenomena if the kind of explanation used—the kind of model—is similar to ones already used.”

Analogy

Thagard cites Hanson's notion of a "logic of discovery," distinct from a logic of justification. Thagard thinks there is no true distinction because analogy plays a role in both:

"Analogy may be used either to direct inquiry toward certain kinds of hypotheses or to support hypotheses already discovered."

In sum, Thagard has argued that scientific inference is inference to the best explanation, and that explanations are evaluated using three criteria:

1. Consilience, which is a measure of *how much* a hypothesis explains.

   Consilience is constrained by …
2. Simplicity, which is a measure of the number of *ad hoc* auxiliary

   hypotheses employed by a theory in order to explain more.
3. Analogy, a type of reasoning that is supported by explanatory

   considerations and which itself supports explanatory understanding.

Gems

Clear setup: here’s the problem with IBE, and I’ll show how to solve it

Consilience and *ad hoc*-ness are *dynamic*

Elegant segue between consilience and simplicity

Questions

- Did Thagard persuade us that his criteria “furnish a comprehensive account of the justification of scientific theories”?

- The criteria are ways to decide when an explanation is best, but Thagard assumes we know what explanation *is* to begin with. Do we?

- The “logic of discovery” versus “logic of justification” talk gets me thinking. What aspects of Thagard’s project—or the project of clarifying inductive inference generally—are merely descriptive, and what aspects are normative? Is Thagard giving scientists a road map for making inductive inferences, or is he just giving philosophers an account of what scientists already do? Or both?