Inference to the Best Explanation

The Good, the Bad, and the Ugly

- **Definition**

  An *explanation* is a hypothesis (or story) about what caused an object to exist, or an event to occur.

  An *inference to the best explanation* (IBE) means judging a hypothesis to be probably true, on the basis that it explains the available evidence *better* than any *competing* hypothesis.

The Mysterious Moving Log

- Fred has a large cedar log, weighing about 1200 pounds, on his front lawn. One morning Fred woke to find that his log had moved about 20 feet up a slight incline. He wondered what caused this.

Competing Hypotheses

- H_1: (A small child had been observed the previous evening riding her bike near the log.) This child ran into the log, knocking it.
- H_2: A troop of 20 trained baboons, with a strong rope, were passing through the neighbourhood. They tied the rope to the log and pulled it.
- H_3: A neighbour just bought a new truck, and wanted to test it. He tied one end of a rope around the log, the other end to the truck, and pulled it.


trek


FRONTIER

Best explanation?

- There could be more hypotheses, of course, perhaps thousands of them. (Limited only by one's imagination.)
- But suppose that these are the only three. Which of them is the best, and why?
- It would be nice to have some *criteria* for how good an explanation is.

Criteria for a Strong Explanation

1. Causation (or “Production”) Condition
   - A *cause* of the evidence must be provided.

2. Prediction (Empirical Adequacy) Condition
   - The proposed cause must *predict* the effect.

3. Prior Plausibility Condition
   - The proposed cause must be *plausible* in itself.

When is *H* a good explanation of *E*?

1. *Causation* Condition (or “production”)

   H makes a claim about something that *caused E*.

   (It may describe the nature of a known cause, or posit the existence of a previously unknown cause.)

- Which of the three hypotheses about the moving log provide a cause?

- All of them.

When is *H* a good explanation of *E*?

- ***Prediction* / Empirical Adequacy Condition**

- *E* can be predicted (or inferred) from *H*, to a high degree.
  – Assuming *H* to be true, one would *expect* to observe *E*.
  – If *E* is a very detailed and precise phenomenon, this criterion means that *H* has to be very detailed and precise as well.

Support requires prediction

- N.B. a theory is *supported* by empirical data only to the extent that it *predicts* the data.

  There is no support without prediction.

E.g. Is Adler's theory *supported* by this?

"The Freudian analysts emphasized that their theories were constantly verified by their "clinical observations." As for Adler, I was much impressed by a personal experience. Once, in 1919, I reported to him a case which to me did not seem particularly Adlerian, but which he found no difficulty in analyzing in terms of his theory of inferiority feelings, Although he had not even seen the child. Slightly shocked, I asked him how he could be so sure. "Because of my thousandfold experience," he replied; whereupon I could not help saying: "And with this new case, I suppose, your experience has become thousand-and-one-fold."

- Popper, "Science as Falsification"

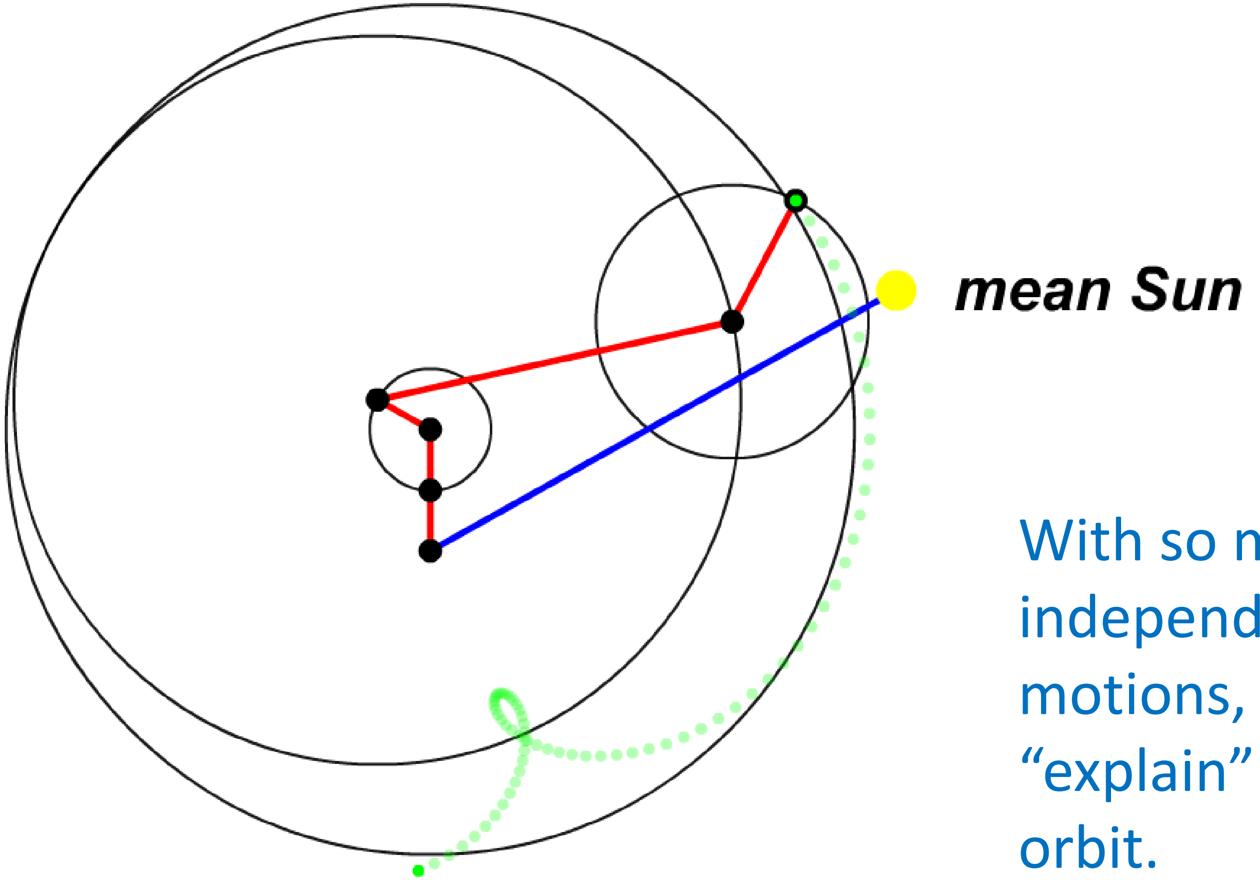
Predictive vs. Post Hoc “science”

Predictive science

Inflexible theory → predict, from the theory, what the data should be → Get data and compare with theory → If they agree, then the theory is confirmed

Post hoc "science"

flexible theoretical framework → get data → construct theory, within framework, to explain data → framework is "confirmed"



With so many independent circular motions, the theory can "explain" any observed orbit.

- Ptolemy's model of Mercury's orbit

- Assuming a child ran into the log, would you predict it to move 20 feet up hill?
  - No. Prediction condition *FAIL.*
- Assuming 20 baboons pulled the log, would you predict it to move 20 feet up hill?
  - Yes. Prediction condition PASS
- Assuming a truck pulled the log, would you predict it to move 20 feet up hill?
  - Yes. Prediction condition PASS

When is *H* a good explanation of *E*?

- ***Prior Plausibility* Condition**

- The cause proposed by *H* is plausible (it seems fairly likely) given our background knowledge.
  - N.B. This judgement of plausibility has **nothing to do with the evidence *E***.
  - How plausible is the theory *before* we learn about *E*?

How do we judge plausibility?

- A plausible hypothesis is one that agrees with what we already believe about the world.
  - It fits our “background knowledge”, or “paradigm”.
    - (This background knowledge may include empirical data other than *E*.)
  - A *simple* hypothesis is more plausible than a complex one, other things being equal. (Ockham’s Razor.)
  - A hypothesis involving known causes is more plausible than one involving unknown causes.

- Is it plausible that a child rode into the log?
  - The child was seen near the log. It's a known cause. Prior plausibility PASS

- Is it plausible that baboons dragged the log?
  - Baboons are known to exist, but not to work in teams, dragging logs using ropes. And they rarely visit south Burnaby. Prior plausibility *FAIL*

- Is it plausible that a truck dragged the log?
  - Trucks are known to exist, and to be present in the neighbourhood. Prior plausibility PASS

Overall, which is best?

	Cause proposed?	Cause is plausible?	Cause predicts E?
H_1 (child)	Yes	Yes	No
H_2 (baboons)	Yes	No	Yes
H_3 (truck)	Yes	Yes	Yes

H_1 is *weak* because it fails to predict the evidence.

H_2 is *weak* because it is implausible.

H_3 is *strong* because it is plausible and predicts the evidence.

$\Rightarrow$ H_3 is the best explanation.

	Cause proposed?	Cause is plausible?	Cause predicts E?
H_1 (child)	Yes	Yes	No
H_2 (baboons)	Yes	No	Yes
H_3 (truck)	Yes	Yes	Yes

Which of H_1 and H_2 is stronger, do you think?

If they were the only options, which would *you* believe?

Stick with your priors, or be persuaded by evidence?

Degrees of Plausibility

- In reality, the plausibility of a hypothesis isn't a yes/no matter. There are *degrees* of plausibility.
- The degree to which a hypothesis is plausible, prior to the (new) evidence, is called its *prior probability*.
- Relative to background knowledge *K*, we can write the prior probability of H as $P_K(H)$.

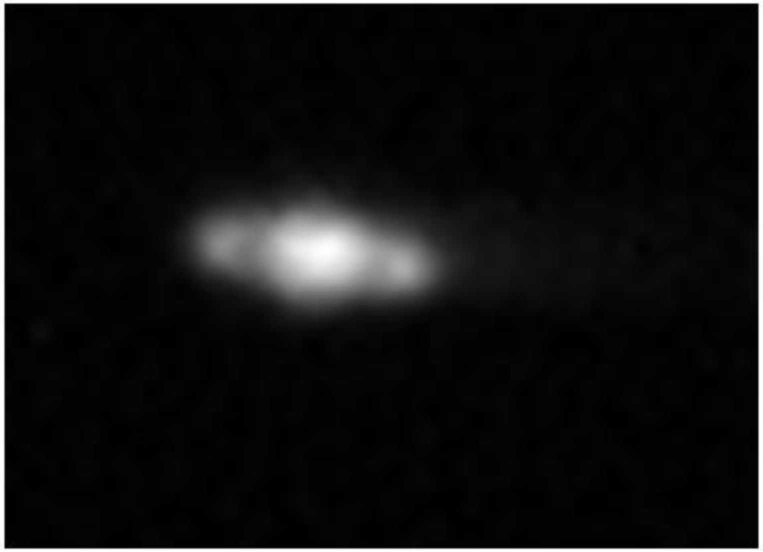
A Theory of Saturn

On 30 July 1610 Galileo he wrote to his Medici patron:

"I discovered another very strange wonder . . . . the star of Saturn is not a single star, but is a composite of three, which almost touch each other, never change or move relative to each other, and are arranged in a row along the zodiac, the middle one being three times larger than the lateral ones, and they are situated in this form: oOo."

- But why a composite of three spheres? Why not a *giant soup tureen*?

Does this hypothesis not predict the data?

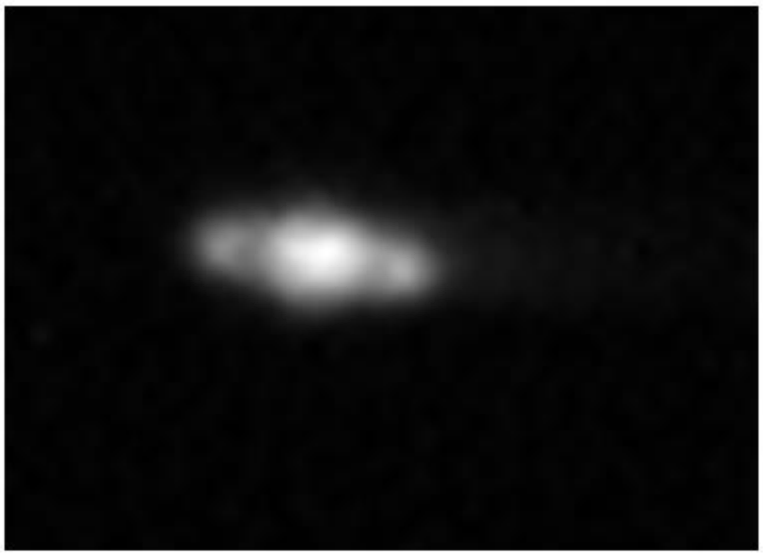


data

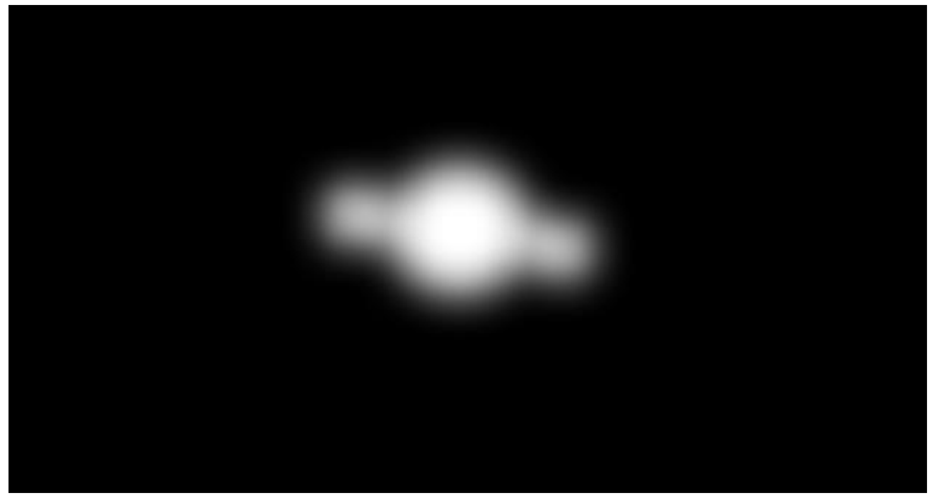
prediction

(I just need to fine-tune those handles a little.)

Does Galileo's theory predict the data?



Data



Galileo's prediction

Somewhat, but not too great.

A scientist's sense of plausibility is fallible …

"In 1825, Mr. McEnery, of Torquay, discovered worked flints along with the remains of extinct animals in the celebrated Kent's Hole Cavern, but his account of his discoveries was simply laughed at.

In 1840, one of our first geologists, Mr. Godwin Austin, brought this matter before the Geological Society, and Mr. Vivian, of Torquay, sent in a paper fully confirming Mr. McEnery's discoveries, but it was thought too improbable to be published.

Fourteen years later, the Torquay Natural History Society made further observations, entirely confirming the previous ones, and sent an account of them to the Geological Society of London, but the paper was rejected as too improbable for publication. "

- (From A. R. Wallace, 1870)

Degrees of Prediction

- There are in degrees to which a hypothesis predicts a piece of evidence.
  - (Some hypotheses predict the evidence more strongly than others.)

- The degree to which a hypothesis predicts the evidence is called the *likelihood* of the evidence, under that hypothesis.

- The likelihood of E under H can be written $P_K(E \mid H)$.

Strength of an Explanation

- The overall strength of an a hypothesis H, as an explanation of E, relative to background knowledge K, is:

$$\text{Strength(H)} = P_K(\text{H}) \times P_K(\text{E} \mid \text{H}).$$

$$= \textit{plausibility} \times \textit{empirical adequacy}$$

In other words, a strong hypothesis has to be plausible *and* predict (be supported by) the evidence.

Part 2

Some examples, and the Sherlock Holmes rule

“Last night my brother was giving me a ride home, and whenever we approached a red light, he pushed a button on his dashboard and it almost immediately turned green. It happened again and again – we never had to stop, not even once.”

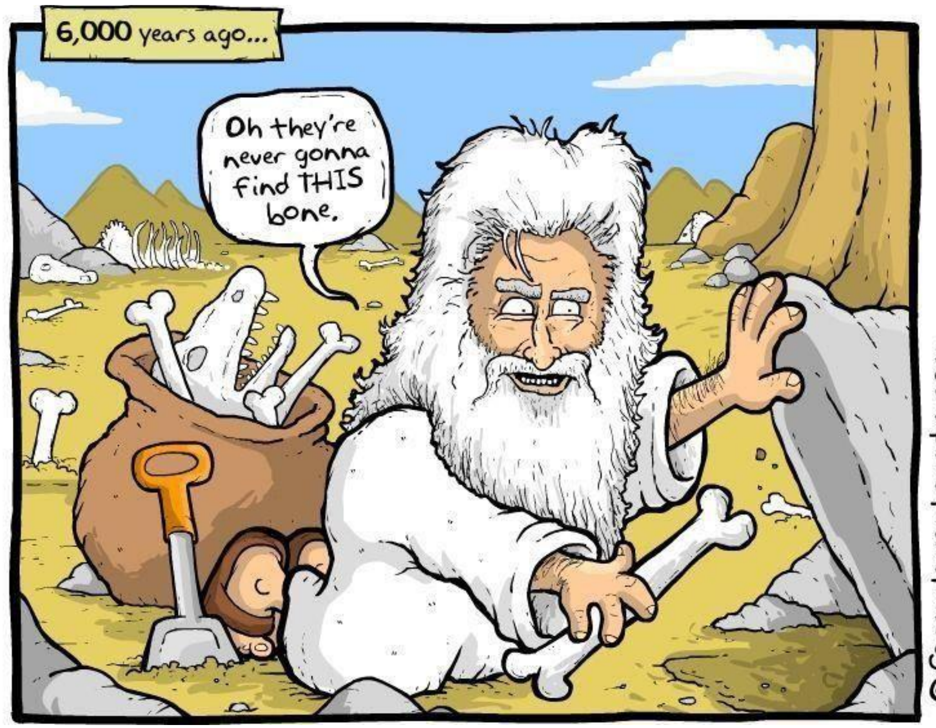
A. The button activates an electronic device that makes red traffic lights turn green.
B. You know that guy is a joker. The button isn’t connected to anything! He just knows how the lights are timed, as he drives that route a lot, and he adjusted his speed to make them all turn green on arrival.

“Ancient texts from around the world describe fantastic events, at around 1500 BC, such the plagues of Egypt, Athena springing from the head of Zeus, and the sun standing still while Joshua finished a battle.”

A. Velikovsky explained all this in *Worlds in Collision* (1950). The planet Venus was initially a comet ejected from Jupiter, that then passed close to earth, which temporarily stopped its rotation on its axis. Of course, the astronomers are too narrow-minded to accept it – they just complain that Velikovsky’s theories are inconsistent with Newtonian mechanics.

"We know that the earth is very old. It contains elements, in abundance, that seem to be the products of very slow nuclear reactions, requiring billions of years. We also see light from distant galaxies that would require billions of years to get here."

A: No. God created the universe about 10,000 years ago, but he made it *look* much older.



IBE is a *competition*

- The method of Inference to the Best Explanation says we should believe the explanation that is the best, or strongest.

- What if *all* the explanations are weak?
  - IBE says “Believe the one that’s a bit less weak than the others”

- What if there are two or three strong explanations?
  - IBE says “Believe the one that’s stronger than the others”

Inexplicable events

- Imagine you flip a coin 100 times, and get the outcome:

  hhhthhtttththhhhthttthtthhhthtthhhttttthhthttht
  hthhhhhhhtttththththhththttthhthhthththhthhttt
  hhthth.

- Is there a good explanation for this *exact* sequence?
  - No. It isn't predicted by *any* plausible theory.
  - Chance is probably the *best* explanation though.

E.g. the origin of life

- Living organisms are built out of proteins, which are *polymers*, i.e. long chain molecules made by connecting units together.
  - The units in these chains are amino acids, There are 20 different kinds of amino acid used in proteins.
- In order to function, a protein needs the right sequence of amino acids.
  - Problem: even if the right amino acids were available, how did they assemble themselves into the right sequence? Given the known laws of chemistry, it seems absurdly improbable.

- “Although a biologist, I must confess that I do not understand how life came about. . . . I consider that life only starts at the level of a functional cell. The most primitive cell may require at least several hundred different specific biological macro-molecules. How such already quite complex structures may have come together, remains a mystery to me.”

Werner Arber, microbiologist and Nobel laureate.

- “One has only to contemplate the magnitude of this task to conclude that the spontaneous generation of a living organism is impossible. Yet here we are -- as a result, I believe, of spontaneous generation.”

  George Wald, "The Origins of Life," in *The Physics and Chemistry of Life* (Simon & Schuster, 1955), p. 270.

  Isn’t it irrational to believe in an “impossible” theory?

What if every possible explanation is weak?

"When you have eliminated the impossible, whatever remains, however improbable, must be the truth" (Sherlock Holmes)

Bayesian version:

"When you have eliminated the absurdly weak explanations then whatever remains, even if it's rather weak, is probably true."

Do you agree?

What about *these* cases, Sherlock?

1. There are cases in the history of science where (for a long period) no one even **thought** of the true explanation. (e.g. quantum mechanics, Saturn's rings)

2. In other cases, the truth had been thought of, but dismissed as too implausible due to **mistaken background ideas**. (E.g. Kent's Hole Cavern, heliocentrism)

When the best explanation (we can think of) is very weak, should we regard it as probably true? Or should we guess that either (1) or (2) above applies?

Part 3

Fallacies involving IBE

A Priori Fallacy

- J. S. Mill identified a general fallacy of *a priori* reasoning: "the proposition … being embraced, not as proved, but as requiring no proof".

- As specific examples, Mill lists:
  - The simplest adequate explanation is probably true
  - There is no action at a distance
  - Everything in nature can be rationally understood, at least in principle
  - Causes resemble their effects

- Is it fallacious to assume such things?
  - I’d say **no**. For one thing, science cannot do without them.

- A fallacy arises, however, when someone:
  - holds onto one’s *a priori* notions **too firmly**
  - **pays too little attention to the empirical evidence**.

- “I always tell friends who proclaim **motorcycling** to be unsafe, that **it’s less dangerous than riding a bicycle around town**, and I truly believe this. I say this because cyclists are (typically) less visible than a motorcycle, are quieter, and travel at a speed much different than surrounding traffic. And surprisingly, the motorcycle itself often absorbs a large amount of the impact in a collision, depending on the type of crash. Bicycles aren’t the same, and you're also not wearing nearly the same type of head/face protection.”

- *What about the stats?*

Motorcycle:	200 deaths/ billion km
Bicycles	24 deaths/ billion km

"If we accept that the human brain is not designed to come into contact with concrete at any speed, tough enforcement of the bike helmet law makes sense."

(Jamie Graham, Chief Constable, Victoria Police Department)

"… the incremental contribution of [B.C.'s] provincial helmet legislation to reduce hospital admissions for head injuries seems to have been minimal."

Jessica Dennis *et al.*, *British Medical Journal*, May 2013.

Wishful thinking

- Similar to over-reliance on *a priori* thinking is **wishful thinking**.
- E.g. Harvard psychologist Steven Pinker reports:

"psychology … is sometimes driven by a utopian vision in which changes in child-rearing and education will ameliorate social pathologies and improve human welfare. … psychological theorists … argue, for example, that innatist theories open the door to inborn differences, which could foster racism, or that the theories imply that human traits are unchangeable, which could weaken support for social programs."

Wishful thinking

- B.C.'s Provincial Health Officer Perry Kendall (from 1999-2018) doesn't see mandatory helmets as much of an impediment to a bike-share program.

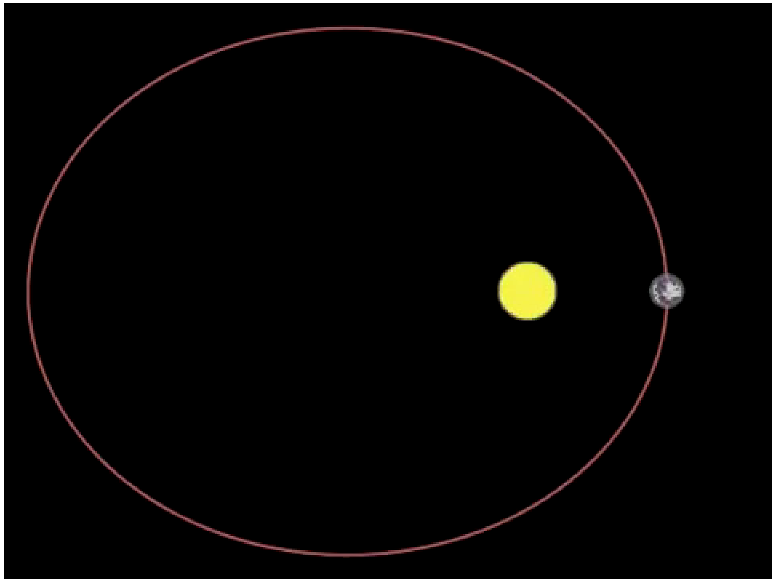
"I think that most people who use a bike-share program plan to use a bike and therefore are quite likely to stick a helmet in their backpack or [have one] strapped to the outside of their briefcase."

Failing to “do the math”

- The prediction condition is very important, of course. This is the only condition where the empirical evidence is considered.
- In many cases, however, it is *not known* whether or not a given hypothesis predicts the evidence.
  - What do we do then?

Example: Newton and the Moon

- After Newton formulated his laws of motion, and “inverse square” law of gravity, he tried to predict the motions of the planets from them.
  - He predicted all 3 of Kepler’s laws. Success!



- But as late as 1740, certain features of the lunar motion could not be predicted from Newton’s Laws.
  - In 1740, were Newton’s Laws a good explanation of the lunar orbit?
- Of course not (yet).
  - Someone has to “do the math”.

The collapse of WTC 1

- According to the BBC, "**Subsequent investigations made it clear that** the tower structures were weakened by the inferno from the planes and felled by the weight of collapsing floors. However even now some people refuse to believe this version of events."

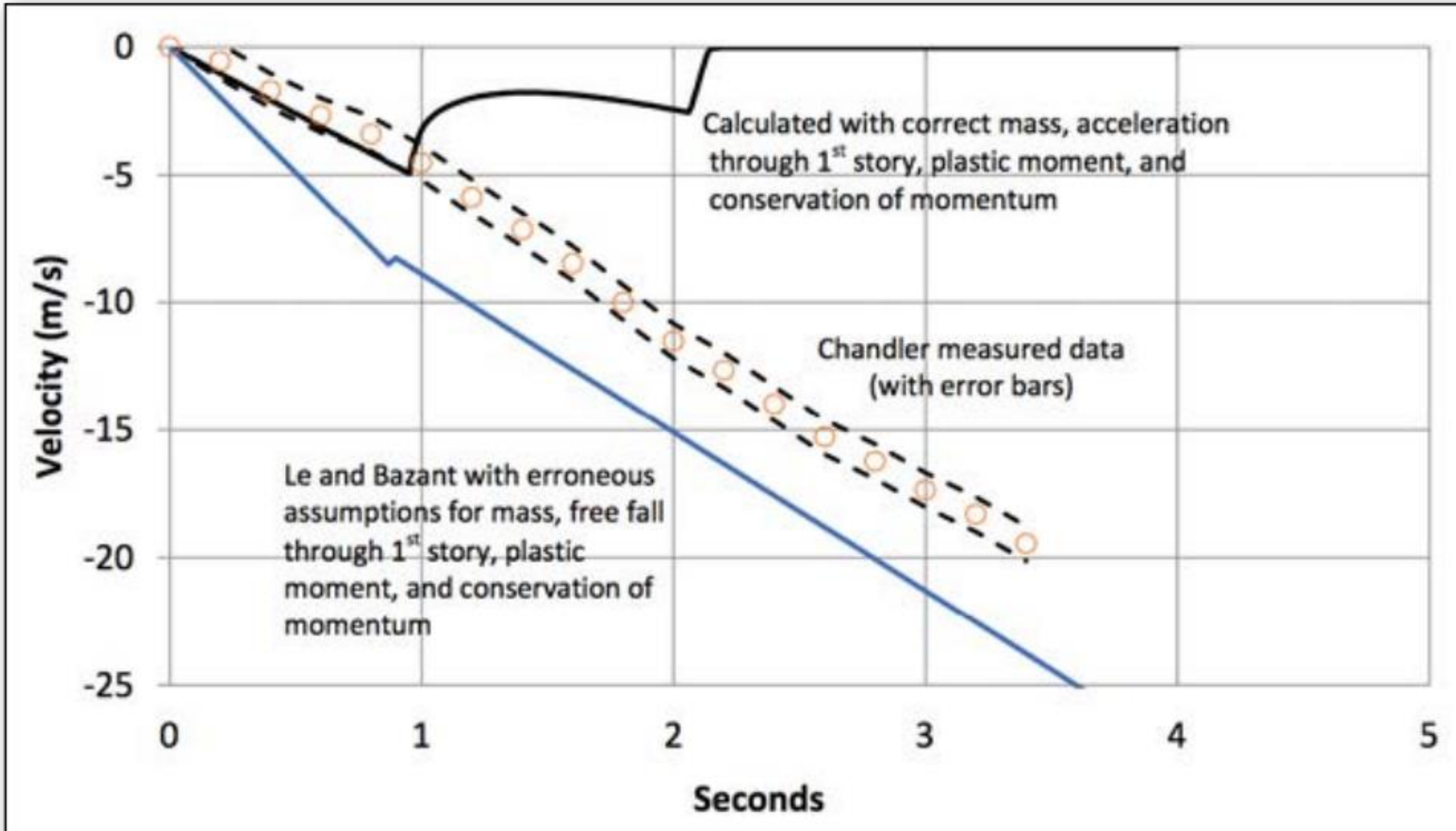
- **Does this hypothesis predict the evidence**, however? How do we know?
  - *Someone needs to do the math*.

Why didn’t NIST model the collapses?

“… At this point, because of the magnitude of the deflections and the number of failures occurring, **the computer models are not able to converge on a solution**. … we are unable to provide a full explanation of the total collapse.” (NIST letter responding to critics)

The situation is too complicated to do a full computer model, representing every bit of steel.

But can’t you do a simplified analysis?



▲ **FIG. 4:** The above graph [10] compares David Chandler's measurement [9] of the velocity of the roofline of WTC 1 with Bažant's erroneous calculation [11] and with Szamboti and Johns' calculation using corrected input values for mass, acceleration through the first story, conservation of momentum, and plastic moment (the maximum bending moment a structural section can withstand). The calculations show that—in the absence of explosives—the upper section of WTC 1 would have arrested after falling for two stories (Source: Ref. [10]).

europhysicsnews

August 2016

Latest articles | Most read articles

7 days | 30 days | 90 days

43919 Free Access

15 years later: on the physics of high-rise building collapses

Steven Jones, Robert Korol, Anthony Szamboti and Ted Walter

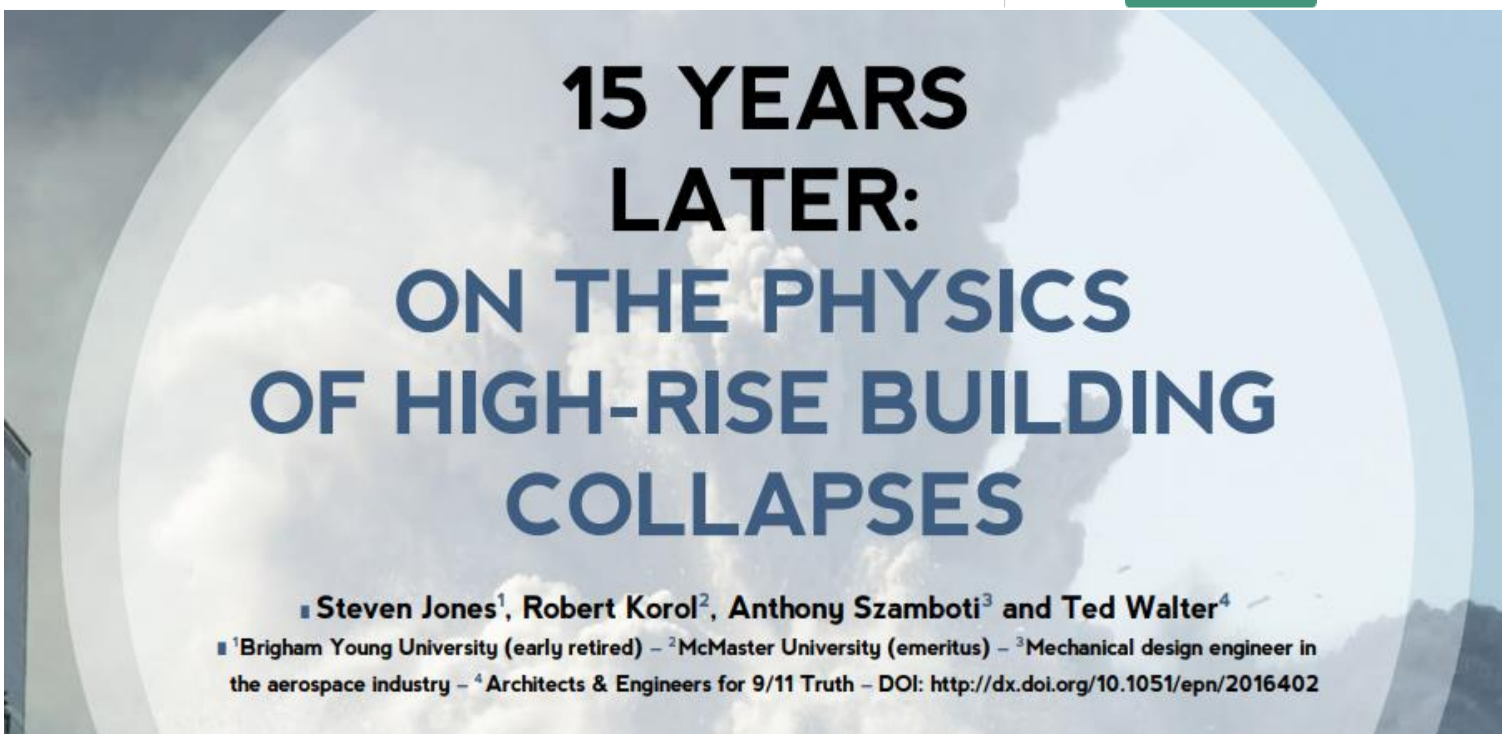
Europhysics News, 47 4 (2016) 21-26

Published online: 24 August 2016

DOI: https://doi.org/10.1051/epn/2016402

PDF (596.9 KB)

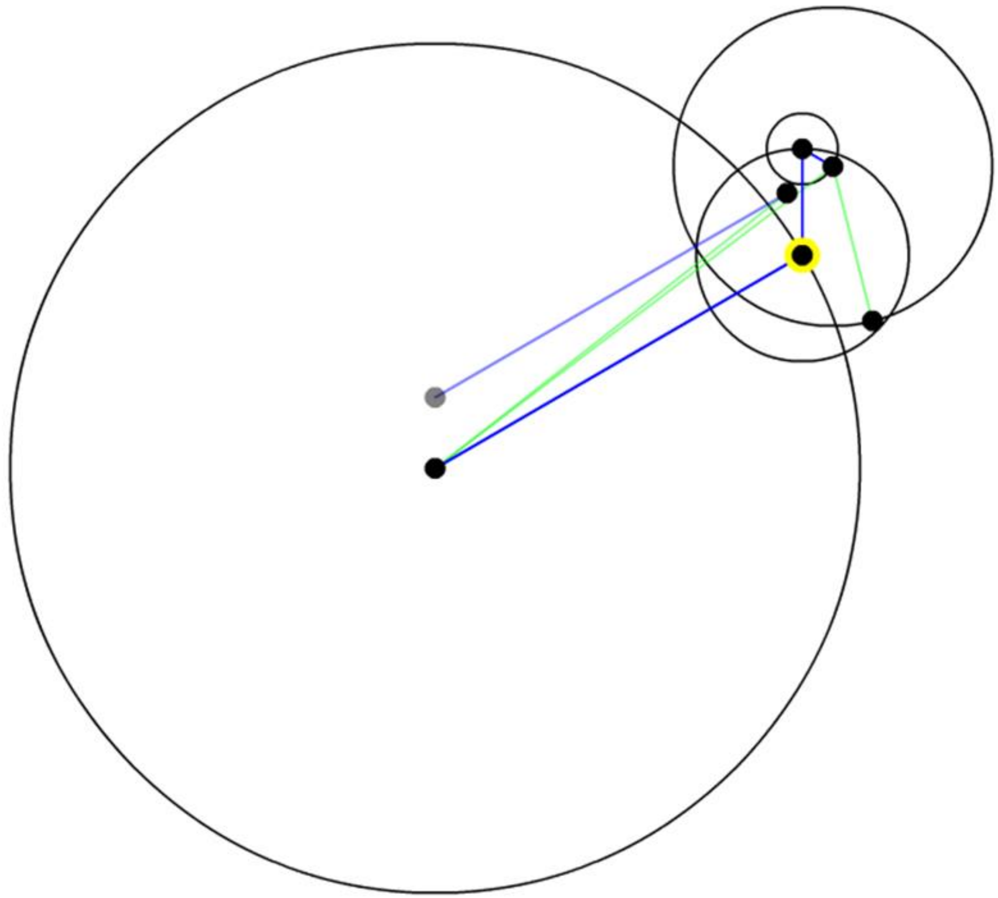
2783 Free Access



Flexible explanations

- Some explanations (or theories) have many "free parameters", making them *flexible*, or adjustable.
- If the data are already known, then *the explanation can be adjusted to fit the data*.
- (Is such flexibility a good or a bad thing?)

E.g. planetary orbits as epicycles

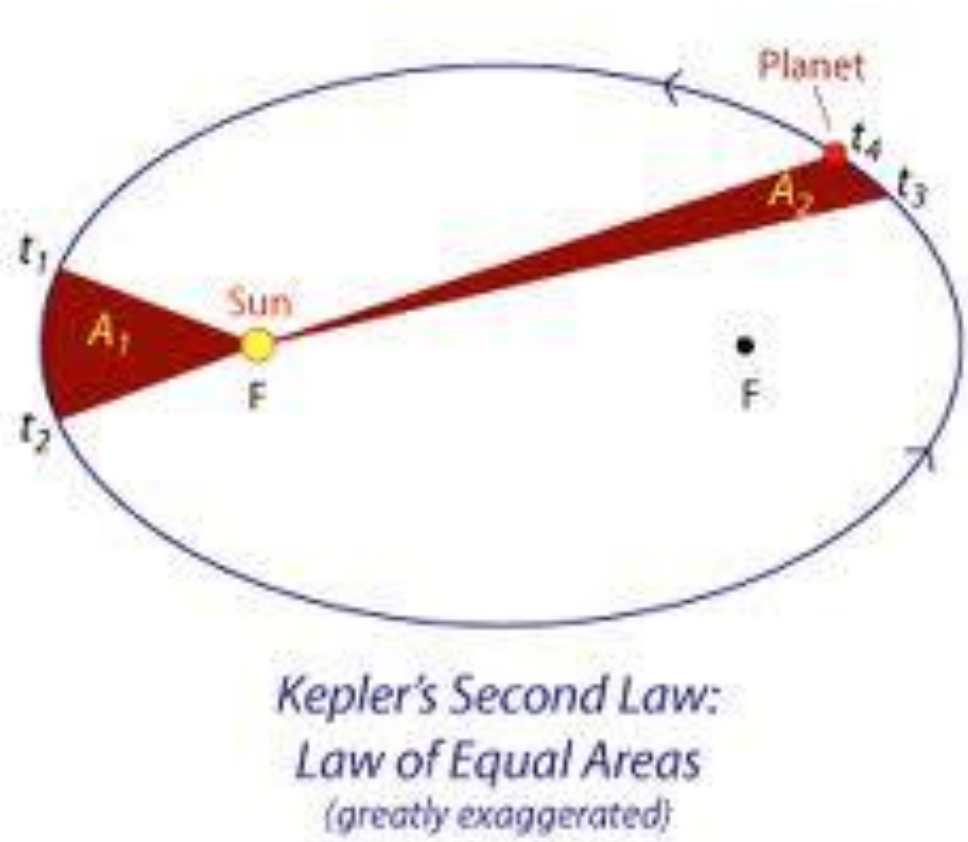


Ibn Al-Shatir's model of Mercury's orbit

Flexibility and plausibility

- Flexible explanations might appear to be strong, as they can often do well on the *prediction* criterion.
- But since the parameter values are chosen to fit the data, those values have no prior plausibility.
  - Scientists are well aware of this issue, and avoid such flexible theories (where possible).

Kepler's model – much less flexible



Example - psychoanalysis

- Unconscious mental states aren't observable by anyone.
- Hence Freud, Adler, etc. is are free to postulate any belief-desire combination, as necessary, to predict the data. The method is *extremely* flexible.

Example: Evolutionary theory

"… evolutionary theorizing … seems to possess a disquieting amount of elasticity or flexibility with regard to explaining organic phenomena. Anything and everything in the empirical biological world seems to be compatible with evolutionary explanations. … In evolutionary explanations **the theorist simply assumes everything he needs to make the explanation work**."

ARTHUR CAPLAN, "TESTABILITY, DISREPUTABILITY, AND THE STRUCTURE OF THE MODERN SYNTHETIC THEORY OF EVOLUTION", *Erkenntnis* 13 (1978) 261-278.