Philosophy of Science: Applications in Linguistics

Austin Blodgett, Georgetown University
ajb341@georgetown.edu

Abstract

This tutorial will present a brief history of how different scholars answer the questions “What makes science science?” or “What procedures are scientific?”. We will review a brief timeline of the history of science, and discuss topics such as falsifiability, empiricism, rationalism, the statistical revolution, and the trend toward making data mathematical.

The tutorial will include exercises intended to prompt discussions about open problems in linguistics and relate them to topics in philosophy of science. These discussions are intended to be completely open-ended. Topics to be discussed include the DP-hypothesis, Optimality Theory, Modular vs. Interconnected Theories of Language Cognition, and others.

The tutorial is primarily designed to be fun and interesting, but it will also give participants tools for analyzing linguistic topics, critiquing or defending linguistic claims, and understanding their relationship with science more broadly.

Philosophy of Science: Outline

- A Concise Timeline
- Characteristics of Scientific **Theories**
  - Popper: Falsifiability and Predictiveness
  - Copernicus: What makes one theory better than another?
  - Exercises 1 & 2: DP-Hypothesis and Optimality Theory
- Characteristics of Scientific **Procedures**
  - Empiricism vs. Rationalism
  - Hume's Problem: Why believe inductive arguments?
  - Statistical Revolution and Modern Perspectives
  - Exercises 3: Language Acquisition of Subjects
- **Data** in the Sciences
  - Dalton on Atoms: How to reason about unobservables
  - Noise and Measurement Error
  - Exercises 4 & 5: Universal Coverage Problem and Modular vs. Interconnected Cognition

Our Approach

In philosophy of science, the **problem of demarcation** is the problem of finding a set of necessary and sufficient conditions that define science.

We will not:

- Solve the problem of demarcation

We will:

- Cover a broad selection of perspectives on what makes science science.

Our Approach

Skills for dissecting a scientific theory:

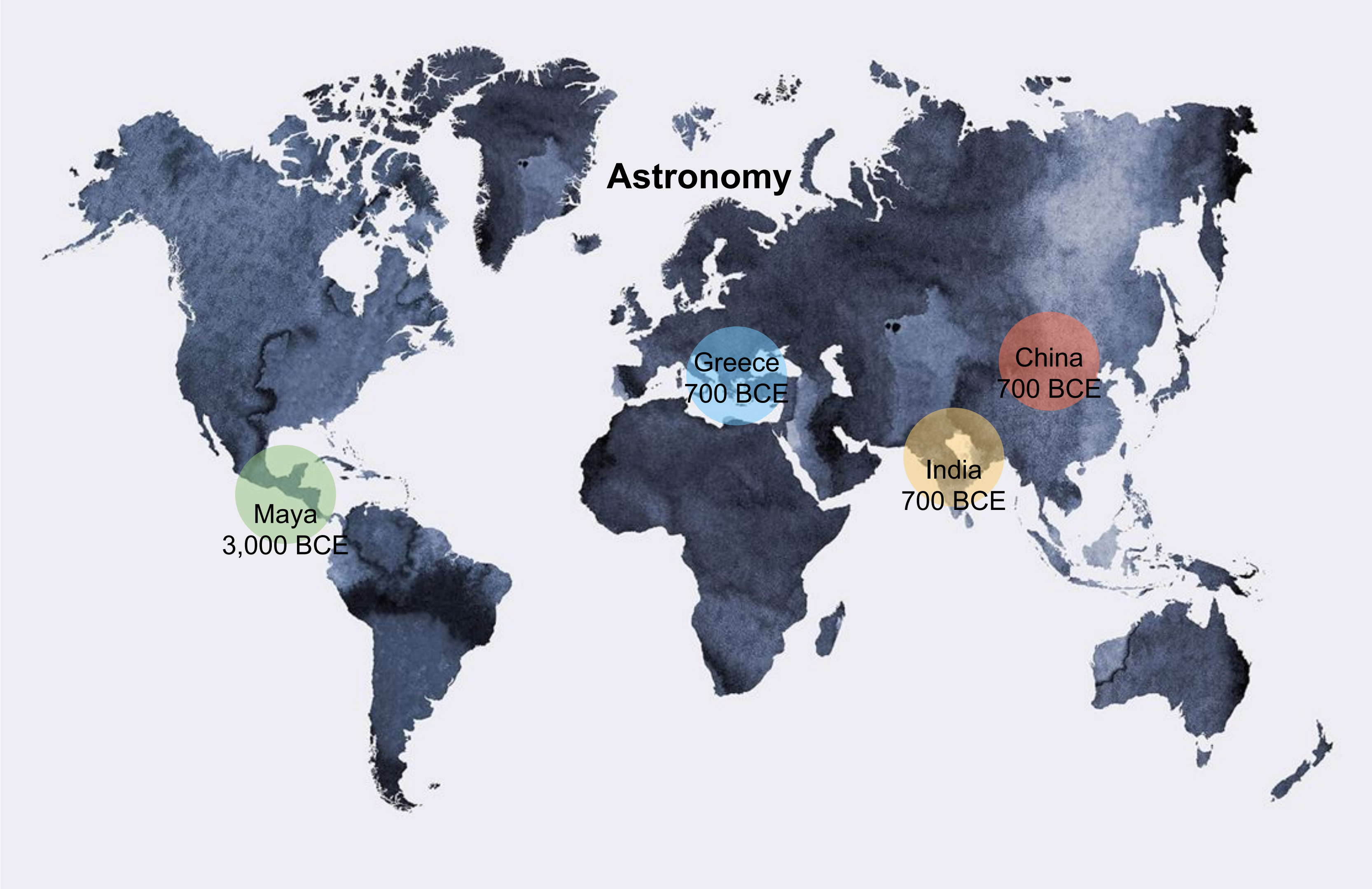
- How did the scientist(s) build on previous work?

  e.g., Einstein's special relativity

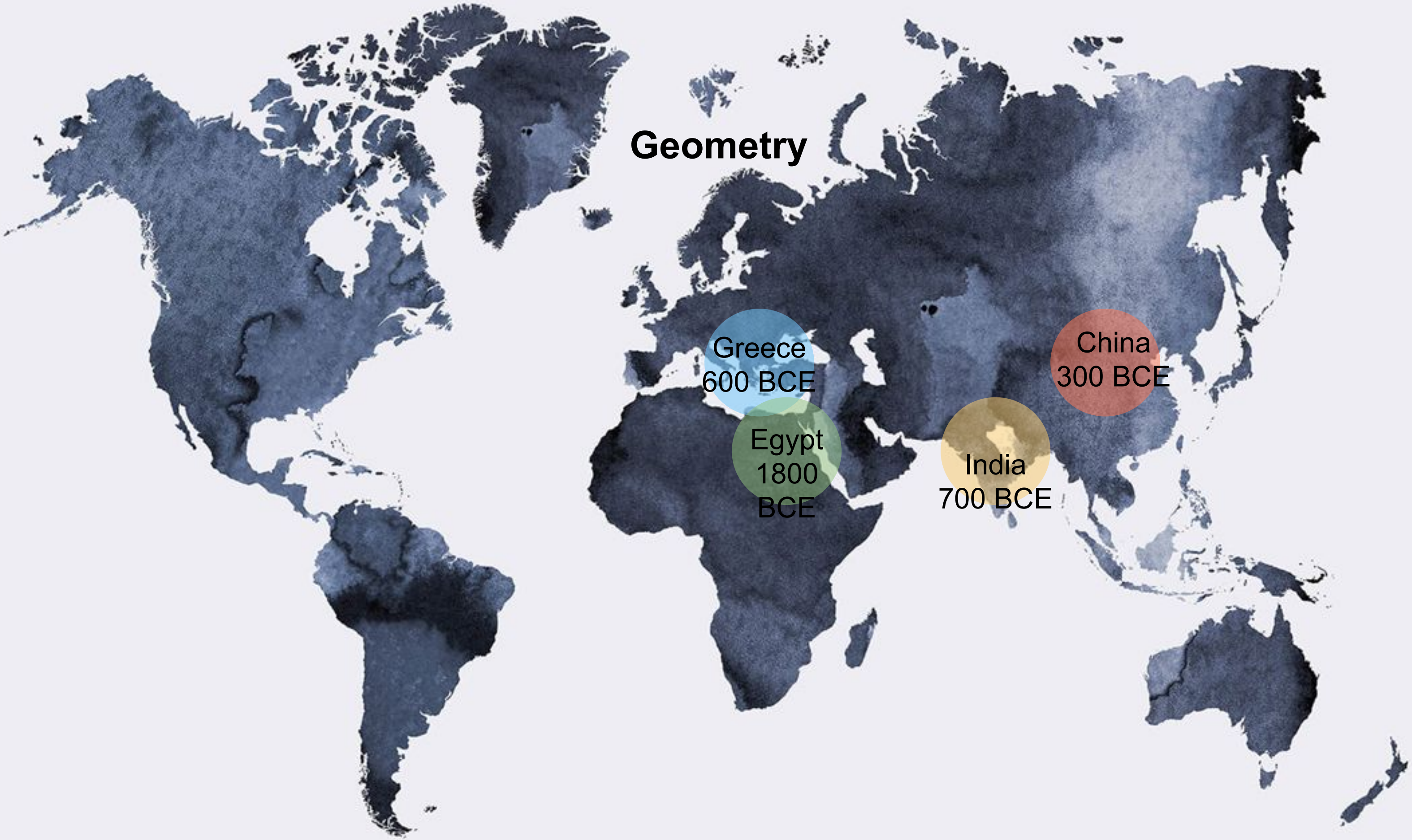
- Be able to break large theories into individual claims and contributions

  e.g., Copernicus' heliocentrism

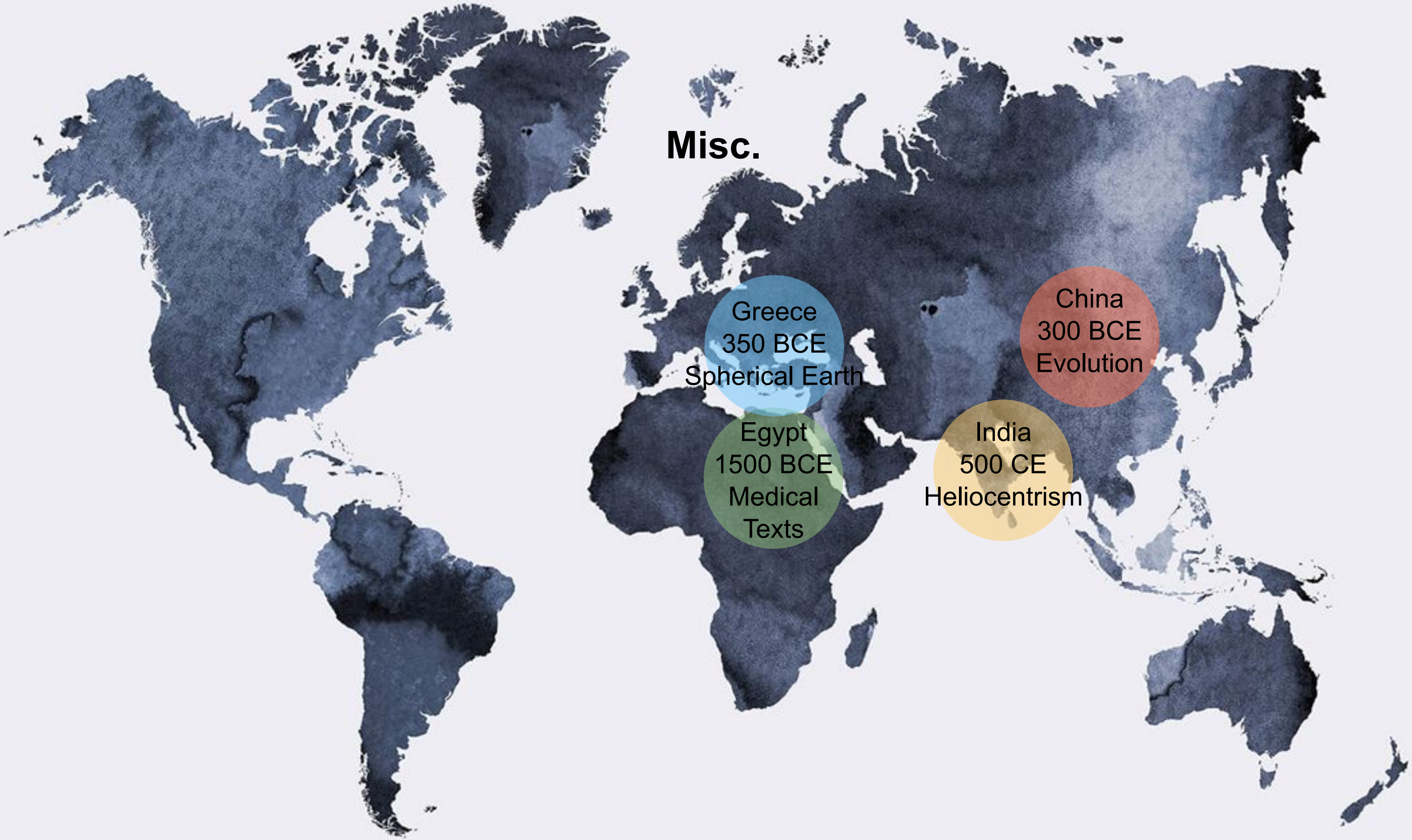
1. A Concise Timeline


Astronomy
Greece
700 BCE
China
700 BCE
India
700 BCE
Maya
3,000 BCE

Geometry



Misc.


Greece
350 BCE
Spherical Earth
China
300 BCE
Evolution
Egypt
1500 BCE
Medical
Texts
India
500 CE
Heliocentrism

Medieval Period & Islamic Golden Age

- 628 – Brahmagupta; Hindu-Arabic Number System (with Zero)
- 820 – Al-Khwarizmi; Algebra, first Algorithms
- 1020 – Ibn Al-Haytham; Scientific Method
- 1100 – Ibn Al-Nafis; Anatomy, Discovery of Cardiovascular System
- 1532 – Nicolaus Copernicus; Heliocentric Model

1600s: Foundations of Physics

- 1609 – Kepler; Laws of Planetary Motion
- 1610 – Galileo Galilei; Moons of Jupiter, Galilean Relativity, Free Fall
- 1620 – Francis Bacon; Work on the Scientific Method (Philosopher)
- 1644 – Rene Descartes; *Principles of Philosophy* (Philosopher)
- 1687 – Isaac Newton; *Mathematical Principles of Natural Philosophy*

Around the 1700s: Statistical Revolution

- 1654 – Blaise Pascal, Pierre de Fermat, Gerolamo Cardano; Probability Theory
- 1713 – Jacob Bernoulli; Law of Large Numbers
- 1739 – David Hume; *A Treatise of Human Nature* (Philosopher)
- 1763 – Thomas Bayes; Bayes' Theorem
- 1809 – Carl Friedrich Gauss; Gaussian Distribution
- 1814 – Pierre-Simon Laplace; Probability for Inductive Reasoning

1800s: Biology, Chemistry, & Medicine

- 1808 – John Dalton; First Evidence of Atoms
- 1843 – Ada Lovelace; First Computer Algorithm
- 1847 – Ignaz Semmelweis; Germ Theory of Disease
- 1859 – Charles Darwin, Alfred Russel Wallace; Evolution by Natural Selection
- 1866 – Gregor Mendel; Genes and a theory of Inheritance
- 1869 – Dmitri Mendeleev; First Periodic Table
- 1880 – Louis Pasteur; First Vaccine
- 1898 – Marie Curie; Discovery of Radioactivity

1900s: Modern Science

- 1900 – Max Planck; First Description of Energy Quanta
- 1905 – Albert Einstein; Special Relativity, Photoelectric effect
- 1927 – Georges Lemaître; Proposal of Big Bang Theory
- 1928 – Alexander Fleming; Discovery of Antibiotics
- 1935 – Karl Popper; *The Logic of Scientific Discovery* (Philosopher)
- 1938 – Guy Callendar; Climate Change & Greenhouse Effect
- 1953 – Rosalind Franklin, James Watson, Francis Crick; Structure of DNA
- 1957 – USSR; Sputnik 1, First Artificial Satellite
- 1962 – Thomas Kuhn; *The Structure of Scientific Revolutions* (Philosopher)
- 1969 – NASA; Apollo 11 Moon landing

Recent Developments

- 2003 – Human Genome Project; Completed Mapping the Human Genome
- 2012 – Large Hadron Collider; Discovery of the Higgs Boson
- 2015 – NASA; Flyby Images of Pluto
- 2019 – MRC Laboratory; Artificial Bacteria
- 2019 – Event Horizon Telescopes; First Image of a Black Hole
- 2020 – SpaceX; First Commercial Space Flight

2. Properties of Scientific Theories

Karl Popper (1934)

A Scientific theory should be:

Falsifiable - A scientific theory needs to make claims that could be proven wrong.

Predictive - A scientific theory needs to make surprising predictions that turn out to be true.

Karl Popper (1934)

Popper compares:

Einstein's general relativity - Makes a surprising prediction (light bends around massive objects) which was proven to be true

Freud's psychoanalysis - Makes claims about many types of behaviors but is consistent with all possible observations.

Karl Popper (1934)

Criticisms:

Poppers thesis doesn’t say anything about the procedures scientists use to make or test a theory. Most of the actual work of science in not discussed by Popper (though there was already plenty of scholarship on that).

Karl Popper (1934)

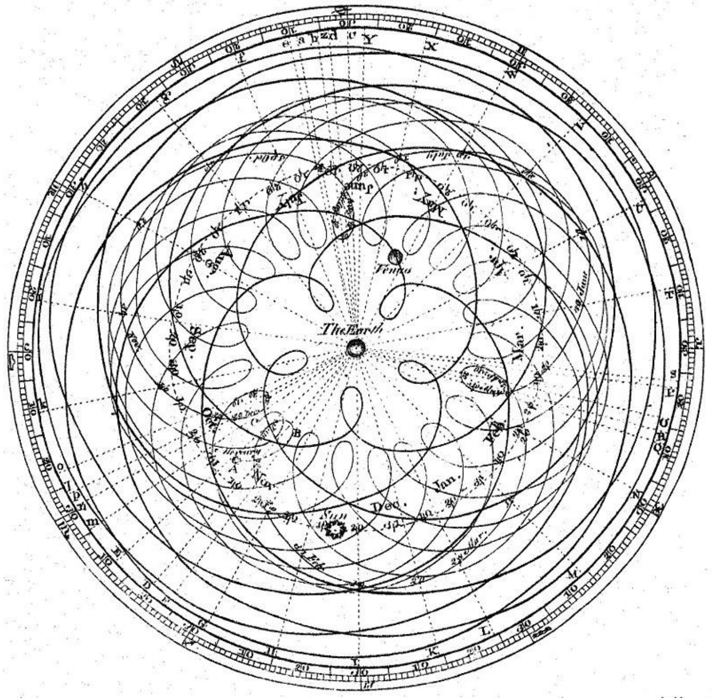
Criticisms:

Falsification often doesn’t describe how scientists behave in practice. Even after a theory has been falsified, scientists will often continue to rely on it until a better theory comes along.

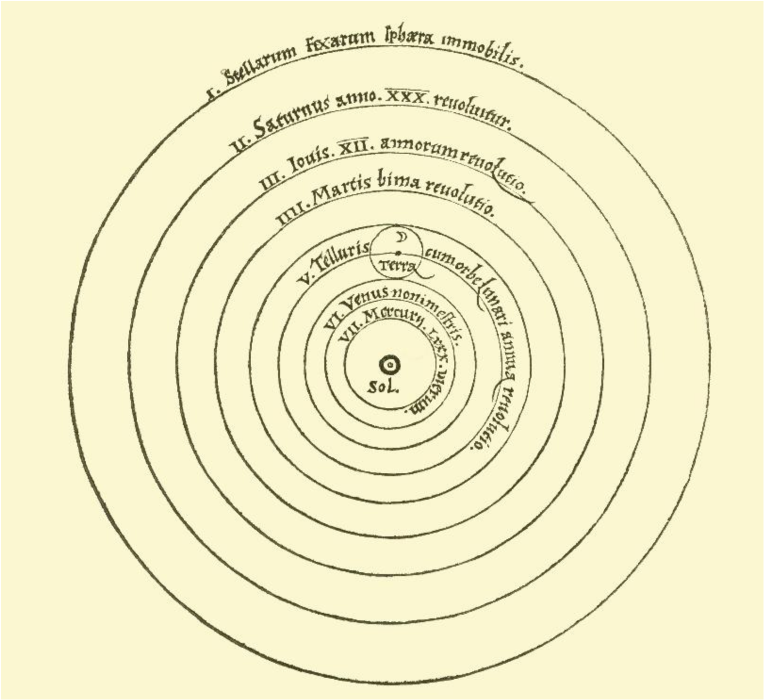
> “The discovery of an inconsistency—or of an anomaly—[need not] immediately stop the development of a programme: it may be rational to put the inconsistency into some temporary, *ad hoc* quarantine, and carry on with the positive heuristic of the programme.”
>
> - Imre Lakatos (1970)

Ptolemy's Geocentrism vs. Copernicus' Heliocentrism

Ptolemy's geocentrism



Copernicus' heliocentrism



What makes one theory better than another?

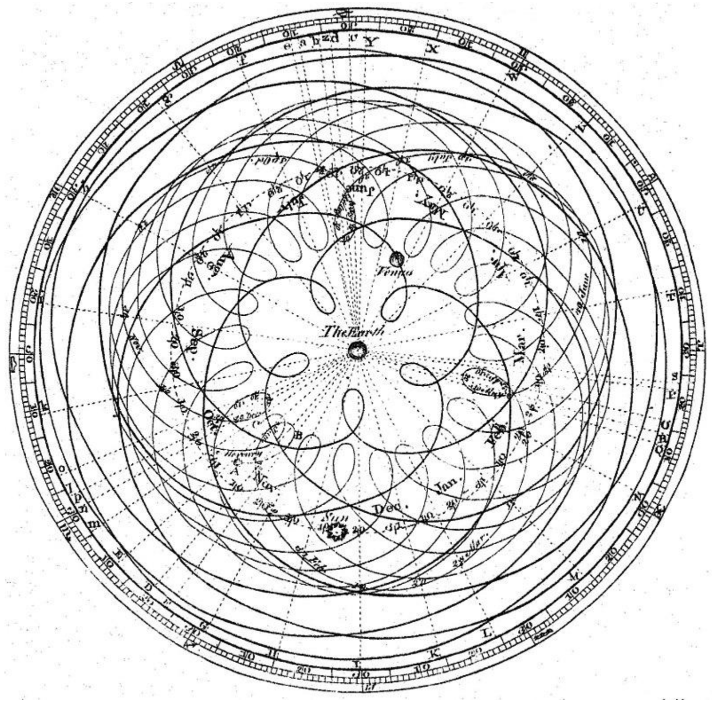
Observational Data

1. Apparent revolutions of sun (per day), moon (per month), and stars (per year)
2. Slightly different stars are visible from the northern regions than southern regions
3. Planets move across the sky independently of other heavenly bodies.
4. Apparent Retrograde Motion: Planets sometimes make ribbon loop motions in the night sky (appear to move forwards than backwards then forwards again)

What makes one theory better than another?

Ptolemy’s Geocentric theory

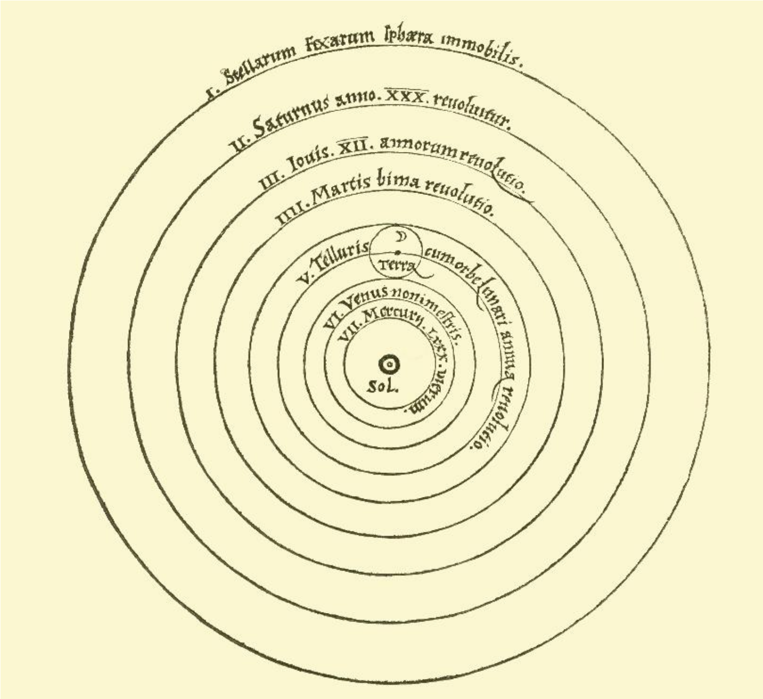
1. The earth is motionless.
2. The moon, sun, planets, and stars move around the earth.
3. Epicycles: The planets must follow an epicycle (ribbon loop) path to explain retrograde motion.



What makes one theory better than another?

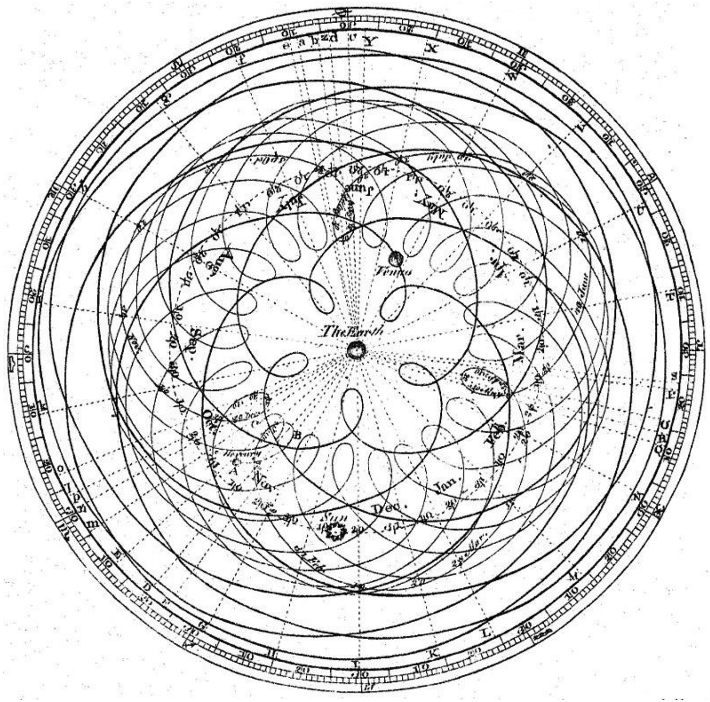
Copernicus’ Heliocentric theory (1543)

1. The sun is motionless
2. The earth is spherical and rotates on its axis.
3. Planets, including earth, move in concentric circles around the sun. (retrograde motion is an illusion caused by earth’s motion)
4. The moon moves in a circle around the earth.

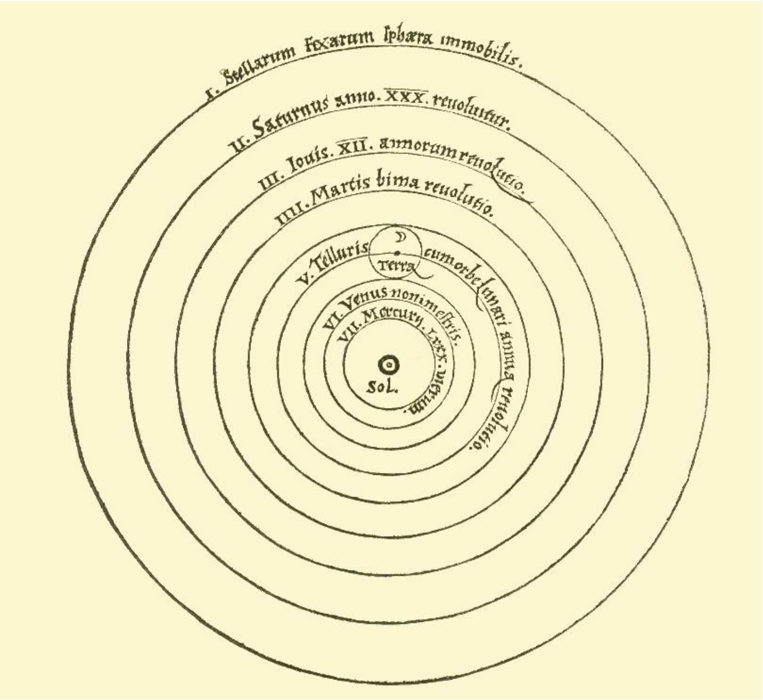


Theories: Difference between a model and a theory

Ptolemy’s geocentrism



Copernicus’ heliocentrism



Theories: Difference between a model and a theory

Model: a framework designed to make accurate predictions

Theory: a set of claims intended to accurately explain and describe some mechanism or phenomenon

Exercise 1: DP-Hypothesis

Exercise 1: DP-Hypothesis

The DP-Hypothesis is that noun phrases are always contained in a determiner phrase even in languages where no determiner is pronounced.

Exercise 1: DP-Hypothesis

Instructions: Choose one of two positions.

Is the DP-Hypothesis falsifiable?

- If yes, what is an example of evidence that would falsify it?
- If no, what would need to be added or changed about the hypothesis to make it falsifiable?

Exercise 2: Optimality Theory (OT)

Exercise 2

/kæt/ + /z/ → [kæts]						
		*Double Sibilant	Agree(Voice)	Max	Dep	Ident
a.	kætz		*!			
b. ☞	kæts					*
c.	kætɪz				*!	
d.	kætɪs				*!	*
e.	kæt			*!		

Exercise 2

Is Optimality Theory technically a theory or a model?

If a theory, what is the underlying mechanism that OT explains? (maybe elaborate on neurobiology or psychology)

If a model, what would need to be added to OT to convince you that it would then become a theory?

3. Scientific Procedures

Empiricism v. Rationalism

Empiricism

People: Isaac Newton, Francis Bacon, David Hume

Rationalism

People: Gottfried Leibniz, René Descartes

Empiricism v. Rationalism

Empiricism

People: Isaac Newton, Francis Bacon, David Hume

Emphasis: inductive reasoning, observation, and evidence

Rationalism

People: Gottfried Leibniz, René Descartes

Emphasis: deductive reasoning, reason, logic/mathematics

Empiricism v. Rationalism

Empiricism

People: Isaac Newton, Francis Bacon, David Hume

Emphasis: inductive reasoning, observation, and evidence

Claims: No a priori knowledge, theories have to be based on observation

Rationalism

People: Gottfried Leibniz, René Descartes

Emphasis: deductive reasoning, reason, logic/mathematics

Claims: A priori knowledge and deduction, inductive reasoning is not credible or reliable

History of the Scientific Method

The scientific method (or *inductive method*) has been proposed and refined by scholars in several centuries.

- 1020 Ibn Al-Haytham
- 1267 Roger Bacon
- 1620 Francis Bacon

One modern variant of the scientific method is the **hypothetico-deductive method**, which is a procedure of repeatedly forming and testing hypotheses until reaching a conclusion.

Hume's Problem

Hume's Problem of Induction: The only way to prove that inductive reasoning works is to use inductive reasoning.

Hume’s Problem

Hume’s Problem of Induction: The only way to prove that inductive reasoning works is to use inductive reasoning.

Example: Given a large number of observations that the sun has risen before, prove that the sun will rise tomorrow.

Hume's Problem

Hume's Problem of Induction: The only way to prove that inductive reasoning works is to use inductive reasoning.

Example: Given a large number of observations that the sun has risen before, prove that the sun will rise tomorrow. According to Hume, you can't do this without making assumptions that inductive reasoning is reliable or about uniformity of nature.

Modern Perspectives: Empiricism v. Rationalism

Empiricism	**Rationalism**
$\Downarrow$	$\Downarrow$
Statistical Methods	Mathematical Theories

Modern Perspectives: Empiricism v. Rationalism

Empiricism

$\Downarrow$

Statistical Methods

I.e., Empiricism becomes mathematically rigorous

Rationalism

$\Downarrow$

Mathematical Theories

E.g., Newton

$F = ma$

$F_1 = F_2 + F_3$

The Statistical Revolution

Pierre-Simon Laplace:

> “The theory of probabilities is at bottom nothing but common sense reduced to calculus; it enables us to appreciate with exactness that which accurate minds feel with a sort of instinct for which of times they are unable to account.”

"Théorie Analytique Des Probabilités". Book by Pierre-Simon Laplace, second edition, 1814.

The Statistical Revolution

Frequentist Statistics

Claim: probability measures a percentage of outcomes

Bayesian Statistics

Claim: probability measures belief or degrees of certainty

Modern Perspectives: Plank, Kuhn, and Lakatos

Max Planck:

> "A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die and a new generation grows up that is familiar with it"

Thomas Kuhn:

> "[When] paradigms change, there are usually significant shifts in the criteria determining the legitimacy of problems and of proposed procedures… Can it conceivable be by accident for example that astronomers first saw change in the previously immutable heavens during the half century after Copernicus...? "

Max Planck (1950). *Scientific autobiography*, p. 33, 97

Thomas S. Kuhn (1996) *The Structure of Scientific Revolutions*, 3rd ed. Chicago Press.

Modern Perspectives: Plank, Kuhn, and Lakatos

Imre Lakatos:

> "The discovery of an inconsistency—or of an anomaly—[need not] immediately stop the development of a programme: it may be rational to put the inconsistency into some temporary, ad hoc quarantine, and carry on with the positive heuristic of the programme."

Imre Lakatos (1970). *Falsification and the Methodology of Scientific Research Programmes*. P. 58.

Exercise 4: Language Acquisition of Subjects

Exercise 4: Language Acquisition of Subjects

Some linguists posit that humans are born with a “language acquisition device” which is a cognitive function that allows quick, adaptive acquisition of language from birth. Rather than debate this entire claim, we’ll focus on acquisition of one simple syntactic relation: the subject-predicate relation. Most if not all languages have a notion of a syntactic subject. But there is some ambiguity across languages of what counts as a subject.

For example:

- Erg-Abs languages use different morphology to mark the only object of an intransitive verb than Nom-Acc languages
- Topic-prominent languages often use specific word order for the topic (and the subject can be absent or not prominent)

Exercise 4: Language Acquisition of Subjects

Do humans have a language acquisition device for acquiring subjects? Alternatively, do young language learners learn that subjects exist? What evidence/arguments would you look for to test which of these hypotheses is correct?

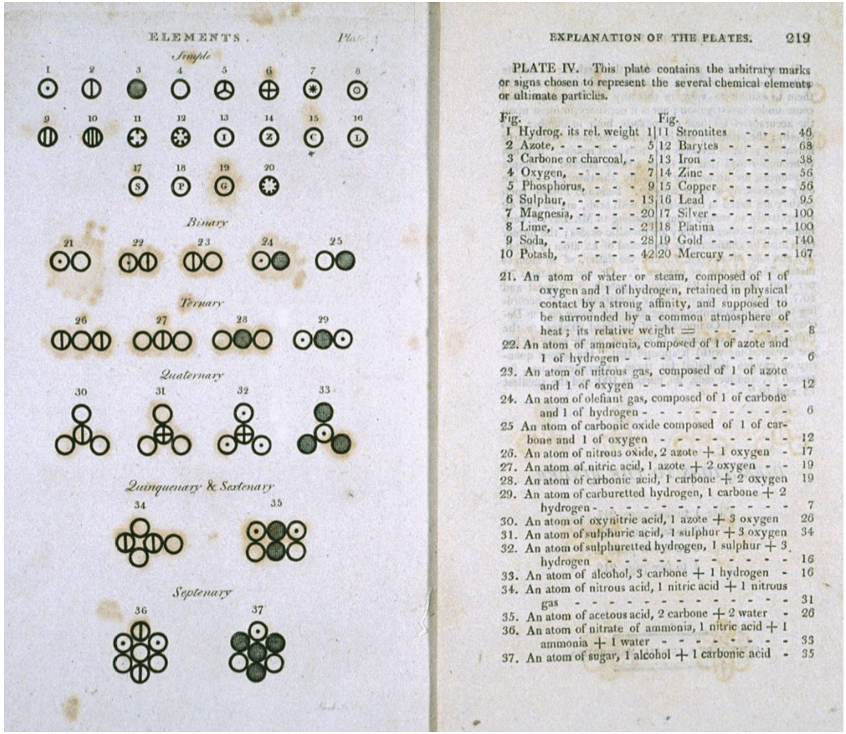
4. Properties of Data

Unknowns and Unobservables

A decent amount of scientific theory involves reasoning about objects and processes that can’t be observed directly. It’s worth looking at an example of how scientists reason about unobservables.

Interpreting Unobservables: John Dalton’s Proof of Atoms

How do we prove that atoms exist if we can’t observe them?



EXPLANATION OF THE PLATES. 219

PLATE IV. This plate contains the arbitrary marks or signs chosen to represent the several chemical elements or ultimate particles.

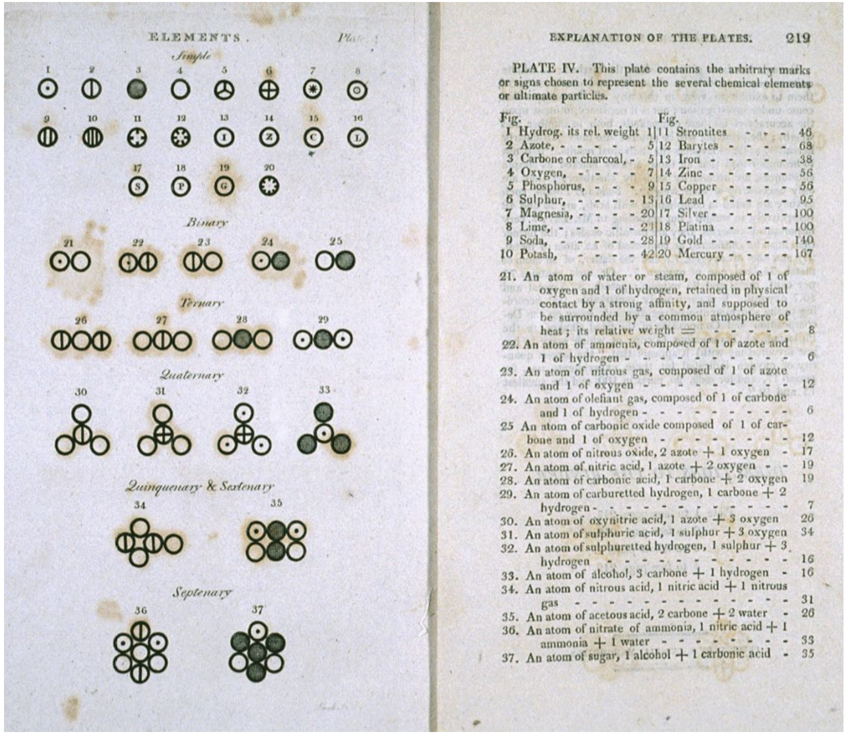
Fig.			Fig.		
1	Hydrog. its rel. weight	1	11	Strontites	46
2	Azote,	5	12	Barytes	68
3	Carbone or charcoal,	5	13	Iron	38
4	Oxygen,	7	14	Zinc	56
5	Phosphorus,	9	15	Copper	56
6	Sulphur,	13	16	Lead	95
7	Magnesia,	20	17	Silver	100
8	Lime,	23	18	Platina	100
9	Soda,	28	19	Gold	140
10	Potash,	42	20	Mercury	167

21. An atom of water or steam, composed of 1 of oxygen and 1 of hydrogen, retained in physical contact by a strong affinity, and supposed to be surrounded by a common atmosphere of heat; its relative weight = - - - - - 8
22. An atom of ammonia, composed of 1 of azote and 1 of hydrogen - - - - - - - - - - 6
23. An atom of nitrous gas, composed of 1 of azote and 1 of oxygen - - - - - - - - - 12
24. An atom of olefiant gas, composed of 1 of carbone and 1 of hydrogen - - - - - - - - - 6
25 An atom of carbonic oxide composed of 1 of carbone and 1 of oxygen - - - - - - - - 12
26. An atom of nitrous oxide, 2 azote + 1 oxygen - 17
27. An atom of nitric acid, 1 azote + 2 oxygen - - 19
28. An atom of carbonic acid, 1 carbone + 2 oxygen 19
29. An atom of carburetted hydrogen, 1 carbone + 2 hydrogen - - - - - - - - - - - - 7
30. An atom of oxynitric acid, 1 azote + 3 oxygen 26
31. An atom of sulphuric acid, 1 sulphur + 3 oxygen 34
32. An atom of sulphuretted hydrogen, 1 sulphur + 3 hydrogen - - - - - - - - - - - 16
33. An atom of alcohol, 3 carbone + 1 hydrogen - 16
34. An atom of nitrous acid, 1 nitric acid + 1 nitrous gas - - - - - - - - - - - - - 31
35. An atom of acetous acid, 2 carbone + 2 water - 26
36. An atom of nitrate of ammonia, 1 nitric acid + 1 ammonia + 1 water - - - - - - - - 33
37. An atom of sugar, 1 alcohol + 1 carbonic acid - 35

Interpreting Unobservables: John Dalton’s Proof of Atoms

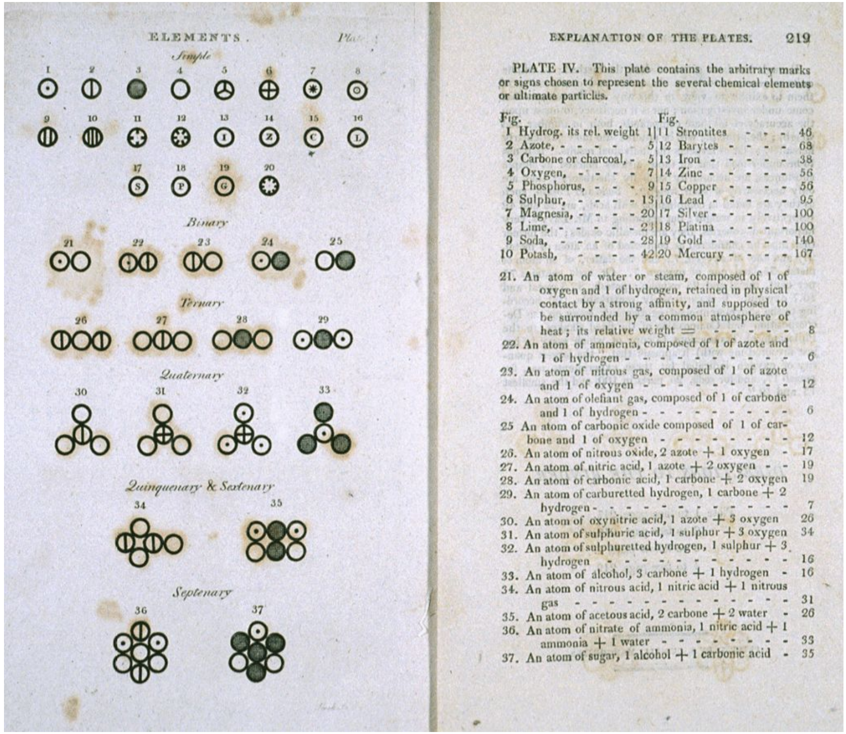
- *Law of Conservation of Mass*:
  mass remains the same before and after a reaction

- *Law of Multiple Proportions*:
  different combinations of elements always vary by whole number multiples.



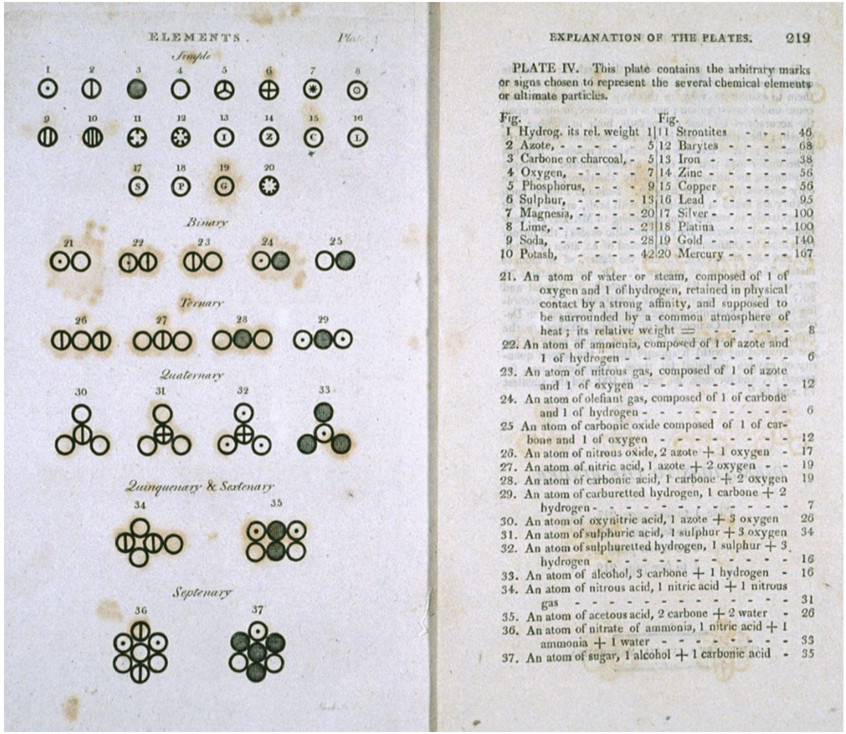
Interpreting Data: John Dalton's Evidence of Atoms

Example: 140g of nitrogen can combine with 160g of oxygen (to form *NO*) or 320g of oxygen (to form NO_2), but always some whole number ratio of oxygen.



Interpreting Data: John Dalton’s Evidence of Atoms

Dalton concluded that an atomic theory of chemistry could account for this by saying that a nitrogen atom combined with a different whole number of oxygen atoms, but theory of chemistry without couldn’t explain this fact.



EXPLANATION OF THE PLATES. 219

PLATE IV. This plate contains the arbitrary marks or signs chosen to represent the several chemical elements or ultimate particles.

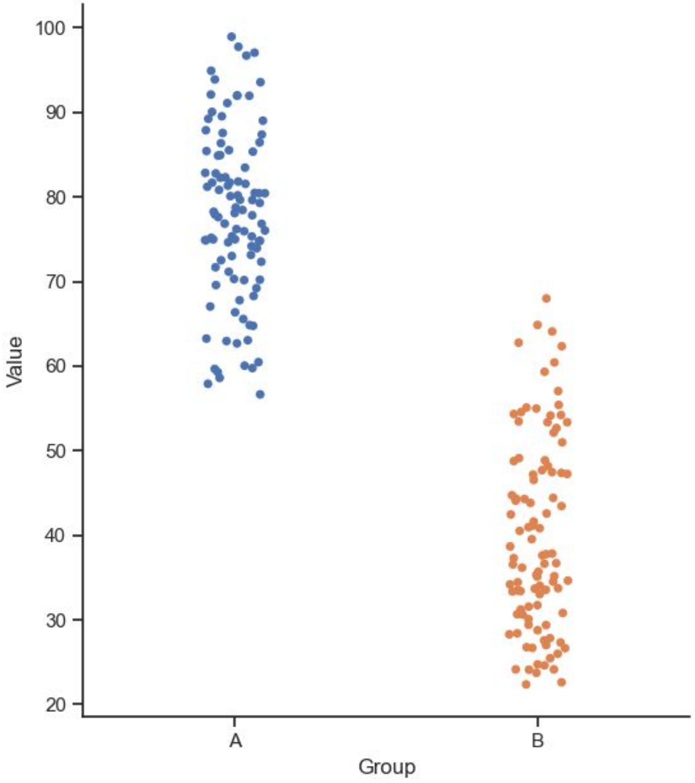
Fig.			Fig.		
1	Hydrog. its rel. weight	1	11	Strontites	46
2	Azote,	5	12	Barytes	68
3	Carbone or charcoal,	5	13	Iron	38
4	Oxygen,	7	14	Zinc	56
5	Phosphorus,	9	15	Copper	56
6	Sulphur,	13	16	Lead	95
7	Magnesia,	20	17	Silver	100
8	Lime,	23	18	Platina	100
9	Soda,	28	19	Gold	140
10	Potash,	42	20	Mercury	167

21. An atom of water or steam, composed of 1 of oxygen and 1 of hydrogen, retained in physical contact by a strong affinity, and supposed to be surrounded by a common atmosphere of heat; its relative weight = - - - - - - 8
22. An atom of ammonia, composed of 1 of azote and 1 of hydrogen - - - - - - - - - - - 6
23. An atom of nitrous gas, composed of 1 of azote and 1 of oxygen - - - - - - - - - - 12
24. An atom of olefiant gas, composed of 1 of carbone and 1 of hydrogen - - - - - - - - - 6
25 An atom of carbonic oxide composed of 1 of carbone and 1 of oxygen - - - - - - - - 12
26. An atom of nitrous oxide, 2 azote + 1 oxygen - 17
27. An atom of nitric acid, 1 azote + 2 oxygen - - 19
28. An atom of carbonic acid, 1 carbone + 2 oxygen 19
29. An atom of carburetted hydrogen, 1 carbone + 2 hydrogen - - - - - - - - - - - - - 7
30. An atom of oxynitric acid, 1 azote + 3 oxygen 26
31. An atom of sulphuric acid, 1 sulphur + 3 oxygen 34
32. An atom of sulphuretted hydrogen, 1 sulphur + 3 hydrogen - - - - - - - - - - - - 16
33. An atom of alcohol, 3 carbone + 1 hydrogen - 16
34. An atom of nitrous acid, 1 nitric acid + 1 nitrous gas - - - - - - - - - - - - - - 31
35. An atom of acetous acid, 2 carbone + 2 water - 26
36. An atom of nitrate of ammonia, 1 nitric acid + 1 ammonia + 1 water - - - - - - - - - 33
37. An atom of sugar, 1 alcohol + 1 carbonic acid - 35

Noise and Measurement Error

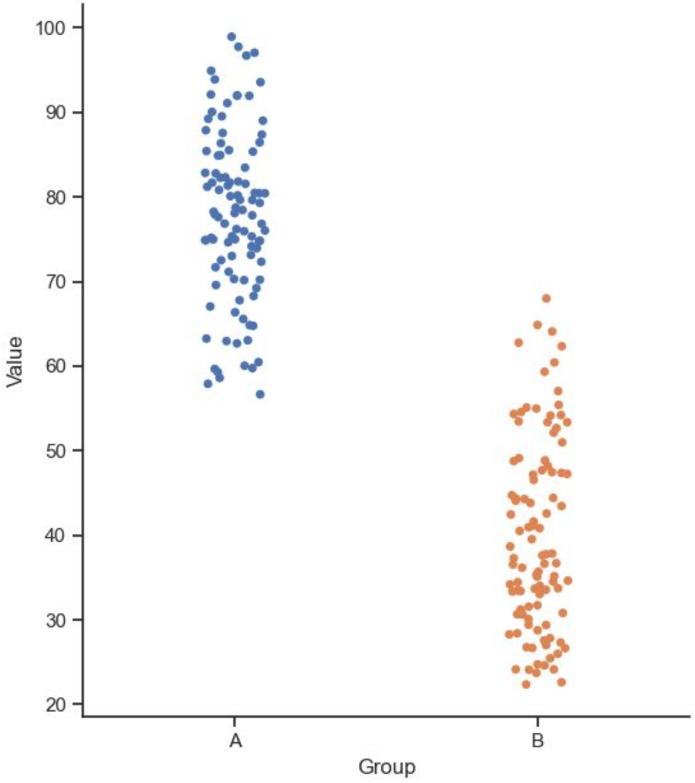
Assume we are behavioral geneticists conducting an experiment. We measure some value from participants and find our two groups have overlapping data.

Example: enjoyment of banana bread



Noise and Measurement Error

Assume we are behavioral geneticists conducting an experiment. We measure some value from participants and find our two groups have overlapping data.



Causes: social/cultural factors, other genes, psychological effects

Noise and Measurement Error

Position 1

Position 2

Noise implies that the theory is incomplete. For a theory to really be predictive we need to include social factors, psychological effects, genetic variation, etc. in the formulation of our theories.

Noise and Measurement Error

Position 1

Noise implies that the theory is incomplete. For a theory to really be predictive we need to include social factors, psychological effects, genetic variation, etc. in the formulation of our theories.

Position 2

Noise implies our measurement is imprecise. As scientists, we are interested in answering particular questions. The measurements we use are often not optimal at answering these questions.

Exercise 4: Universal Coverage Problem

Exercise 4

A common issue when developing formal grammars is the tendency for a grammar to produce the correct predictions* for most sentences (say 95%) but fail on a minority of sentences (say 5%).

*In the case of syntax, “correct” usually means making the correct grammaticality judgement.

Exercise 4

Suppose we have a theory of syntax such as Minimalism, HPSG, LFG, etc. with 95% sentence coverage on a diverse corpus in terms of languages and genres. Suppose this theory of syntax has the best coverage of any known theory of syntax on this corpus.

Does the last 5% of sentences disprove the theory?

If yes, what, in your view, is missing from that theory?

If no, why does the theory fail to predict syntax of last 5% of sentences?

Exercise 5: Modular vs. Interconnected Theories of Language Cognition

Exercise 5

Several schools of thought in linguistics differ in whether they consider language to be **modular** cognitive faculty—in which language functions autonomously from other areas of cognition—or an **interconnected** cognitive faculty—where language cognition is on a spectrum and can be contributed to or subsumed by other areas of cognition.

Exercise 5

Choose one of modular language cognition or interconnected language cognition as a hypothesis. Given that the neurological processes of language are unobservable, what types of evidence would you look for to test that hypothesis?

Exercise 5

Choose one of modular language cognition or interconnected language cognition as a hypothesis. Given that the neurological processes of language are unobservable, what types of evidence would you look for to test that hypothesis?

An example:

Foreground/Background (Figure/Ground) distinction:

> "The duck is in the field" vs. "The field surrounds the duck"

Thank You!

Suggested Readings

McGrew, T., Alspector-Kelly, M., & Allhoff, F. (2009). *Philosophy of Science: An Historical Anthology* (Vol. 14). John Wiley & Sons

Okasha, S. (2016). *Philosophy of Science: Very Short Introduction*. Oxford University Press..

DeKosky, R., & Allchin, D. (1999). *An Introduction to the History of Science in Non-Western Traditions*. History of Science Society.