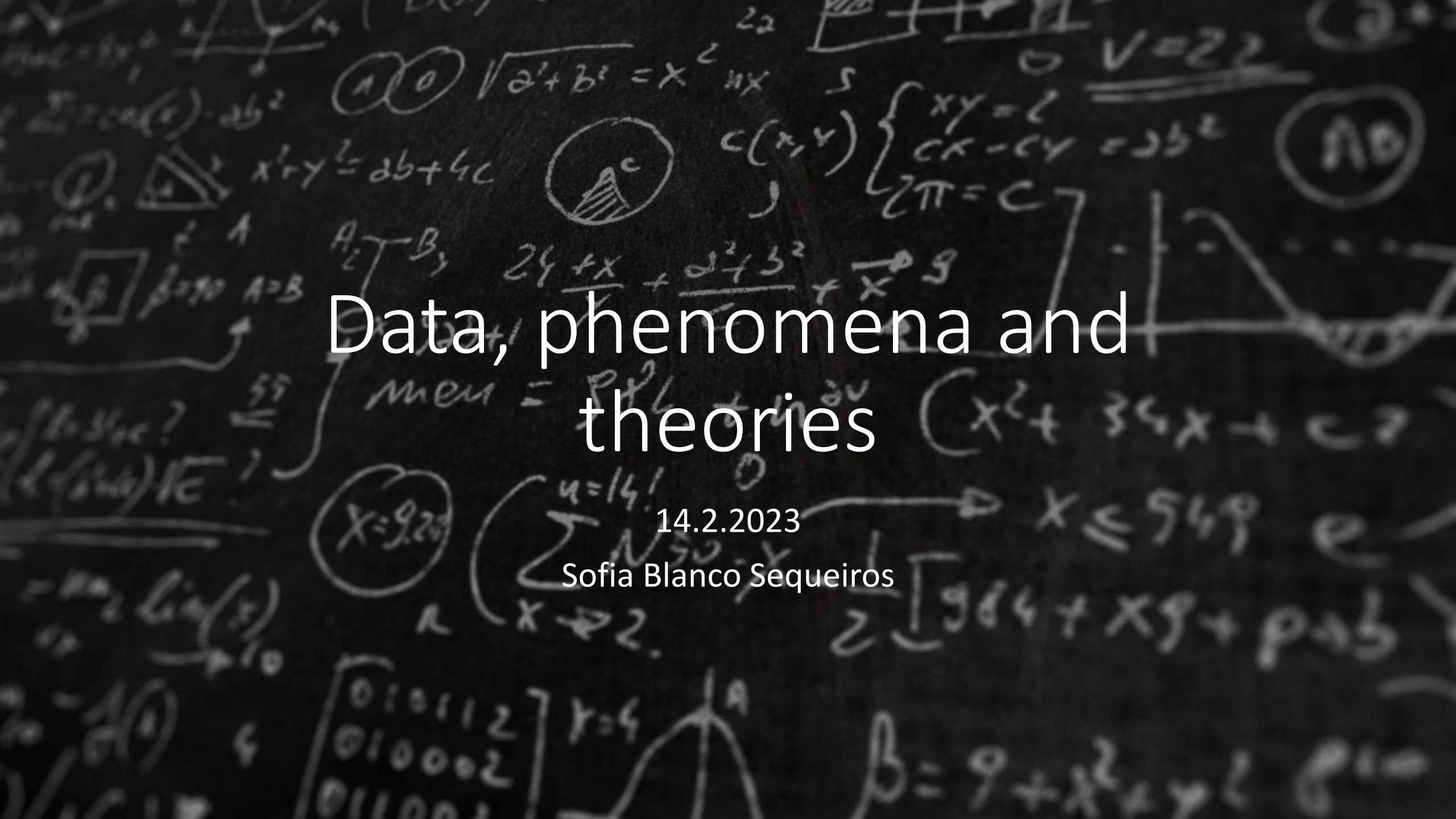

Data, phenomena and theories

14.2.2023

Sofia Blanco Sequeiros

Last time: Theories and Observations

- Theory and observation: two basic ”building blocks” of science
  - History of philosophy of science
- What are scientific theories?
  - Sets of sentences; sets of nonlinguistic (=mathematical or metamathematical) models; sentences, models, practices…
  - The role of theory in science and research
- What is observation?
  - Empirical observation (sensory observation) as the basis of scientific knowledge; critique of ”pure” empiricism and inductivism
- Theory-ladenness of observation

Theory-ladenness of observation

- A chemist will see different things in a chemical reaction than a non-chemist, thanks to their different theoretical backgrounds
- More than that: different scientists will see different things when they observe the same phenomenon
- When Priestley and Lavoisier watched the same experiment with fire, Lavoisier should (or would) have seen combustion and respiration are oxidation processes. Priestley's observations should (or would) have seen burning and respiration as processes of phlogiston release (Boyd, Stanford Encyclopedia of Philosophy)
- NB! Does not relativize all evidence production, but shows how evidence production is dependent on theoretical assumptions, experiences etc.
  - How can science overcome these dependencies? Should it?

Observation

- How do you observe your own object of research (whether you study art, chemical engineering, computer science…)
- What do you observe when you make observations?
- How are your observations theory-laden? Are they (answer: yes!)?

Traditional models of scientific inference

- Pure inductivism: science proceeds by gathering empirical observations, from which inductive generalizations are inferred
  - Inductive vs. deductive reasoning

- Hypothetico-deductivism: scientists formulate hypotheses and test them with observations or data
  - A generalization based on Popper’s idea of falsificationism
  - True theories ”survive” being tested with observations/data

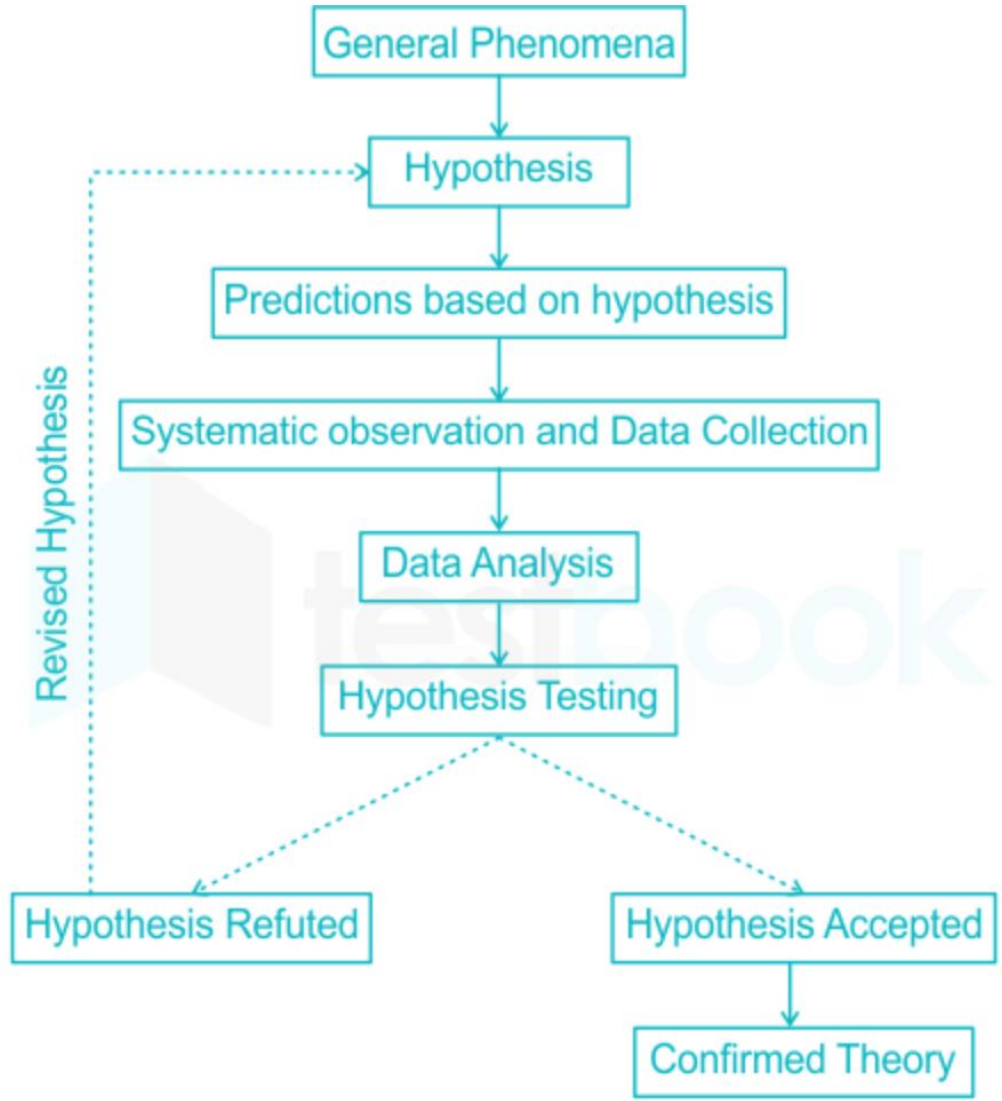

Figure: Hypothetico - Deductive Method

Popper's falsificationism

- Hypotheses can never be verified by observation
  - An infinite number of possible observations; what counts as "enough"?
  - How many white swans do we need to observe to conclude that the hypothesis "All swans are white" is verified?
- Instead, hypotheses should be falsified: they should be formulated in a way that makes them falsifiable, and scientists should look for data that falsifies those hypotheses
  - Still a highly idealized view of natural science, even in the natural sciences
  - Not all science can or should aspire to this ideal
  - Two theoretical problems: Duhem-Quine thesis and underdetermination of theory by data

Duhem-Quine-thesis

- Just because our experiment falsifies (or fails to falsify) a hypothesis, we do not know why it does so
- Is it the experimental setup? Or the auxiliary claims that support the hypothesis?
  - “All people will avoid taxes if they are given the opportunity”
  - If our experiment disproves this claim, we do not know if it is because: 1) The hypothesis is wrong, 2) The experimental setup is wrong, 3) The auxiliary claims that support the hypothesis – e.g. that tax evasion can be studied experimentally – are wrong
  - Should we abandon our hypothesis or not?
- Unambiguous falsification is impossible!

Underdetermination of theory by data

- A single (or even a set of observations) does not help us choose between competing hypotheses
- Me observing my cat eat a mouse multiple times does not help choose between different hypotheses:
  - The cat likes to eat mice
  - The cat does not like mice, but does not get enough food at home
  - The cat hates mice and wants to eradicate them from the face of the earth
- When trying to falsify hypotheses, the choice to falsify is always underdetermined by the available evidence

Traditional models of scientific inference

- According to the traditional view of science, scientists use theory to explain and predict observations
  - Copernicus' heliocentric model of the solar system explains and predicts the movements of planets and other heavenly bodies (=observations)
- What about modern science? Atoms colliding in CERN etc.?
- And what about the arts?

Data, phenomena and theories

- Jim Bogen and James Woodward: we should distinguish between **data**, **phenomena** and **theories**
  - Data: recorded observations, particular to their given contexts
  - Phenomena: stable regularities that persist over time
  - Theories: explain and predict (facts about) phenomena: “events, regularities, processes, etc. whose instances are uniform and uncomplicated enough to make them susceptible to systematic prediction and explanation” (Boyd, Stanford Encyclopedia of Philosophy; Bogen and Woodward 1988, 317)
- Using these concepts paints a more accurate picture of what scientists actually do, what scientific evidence is, and what scientific evidence is evidence about

Data, phenomena and theories

- Bogen and Woodward introduce two concepts with specific meanings: **data** and **phenomena**
  - More accurate: scientists do not tend to talk about ”observations” or ”observation reports”, instead they talk about data

- In science, **we do not observe ”phenomena”, we infer them from data**
  - Counterintuitive: in our daily lives, we do observe phenomena: poverty, ice melting, certain trends in graphic design or architecture, etc.
  - But we do not actually observe the melting point of water, we do not ”see” a thing called ”neobrutalist architecture”

Case example: Bruno Latour and the Soil Scientists

- Soil scientists wanted to figure out why the border between grassland and forest was moving
  - From a clump of soil to a graph on a journal page: a causal, material, step-by-step process

- The scientists started with observing a weird “phenomenon”: movement in the border between grassland and forest
- Collected samples, categorized and analyzed them = produced data
- Used that data to infer facts about the phenomenon

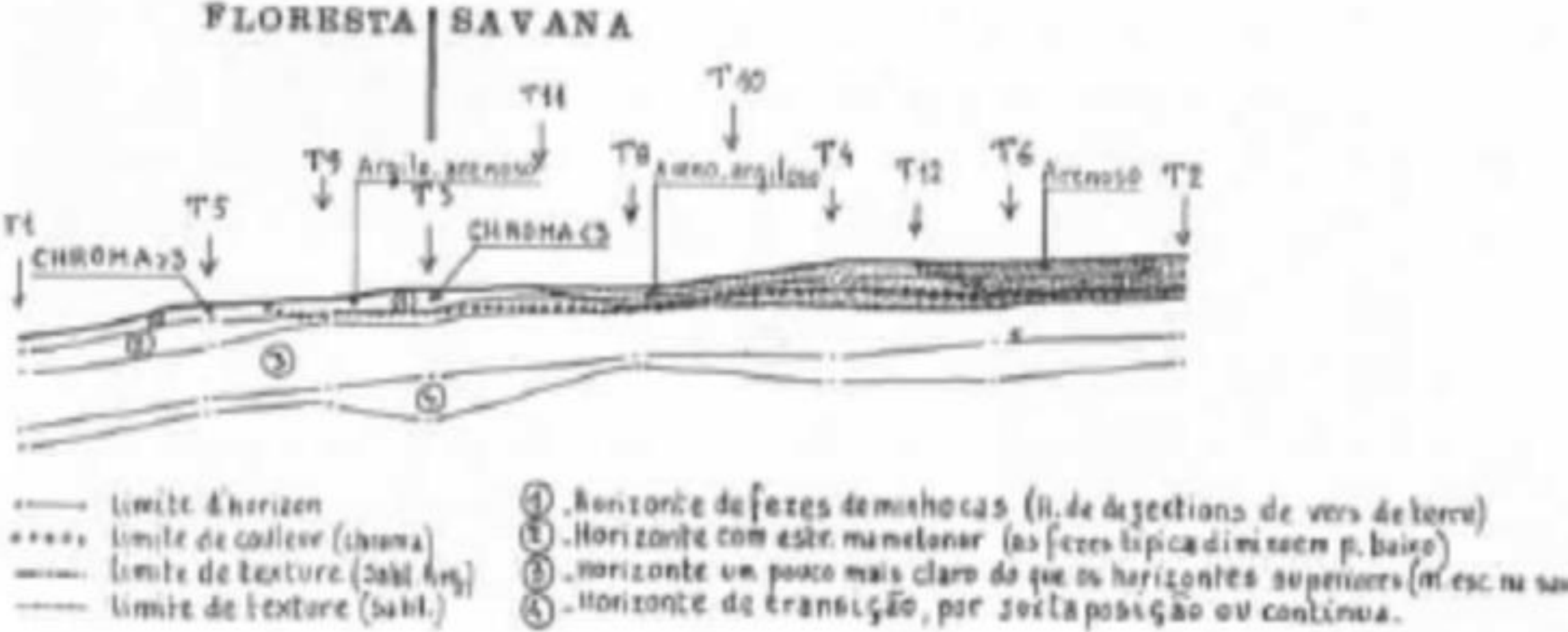

Figure 3. Coupe de transect 1

Figure 2.15

Questions

- What are the phenomena you study in your own field of research?
  - Everyday observations vs. scientifically inferred phenomena
- If the data-phenomena-theory -model does not describe your own field of research: why not? What is a better description?

Data

- The soil scientists **produced data** from ”crude” observations
- Data is particular to its given context: one location at the border produces one dataset, another location another dataset
- In the ideal case (from the scientists’ viewpoint), the different datasets will help the scientists infer a well-defined phenomenon: movement between forest and grassland that is explained by factors X, Y and Z

Nora Mills Boyd (Standfor Encyclopedia of Philosophy)

- “Empirical data are typically produced in ways that make it impossible to predict them from the generalizations they are used to test, or to derive instances of those generalizations from data and non ad hoc auxiliary hypotheses. Indeed, it is unusual for many members of a set of reasonably precise quantitative data to agree with one another, let alone with a quantitative prediction. That is because precise, publicly accessible data typically cannot be produced except through processes whose results reflect the influence of causal factors that are too numerous, too different in kind, and too irregular in behavior for any single theory to account for them."

Phenomena

- Phenomena are stable regularities in the world; something that we can infer with and through data
- Difference between everyday observation of phenomena and scientific observation of phenomena
  - We do not observe the melting point of water; scientists infer it from data
  - We do not observe the exact movement of the border between forest and grassland; scientists infer it from data
- Scientific theories explain (facts about) phenomena, not data
  - Data are recorded observations particular to a context, ”public records”
  - ”Theory T explains why Sofia’s ice block melted at +0 degrees Celcius on Tuesday 14.2. at 5PM” vs.
  - ”Theory T explains why ice melts at +0 degrees Celcius”

Data, phenomena and scientific theories

- Claims about data act as evidence for claims about phenomena
- Claims about phenomena act as evidence for scientific theories
- Data is not evidence for or against a theory by itself – at least when defined in the way that Bogen and Woodward define it
  - Data is evidence for or against a theory only when first used to support or disconfirm claims about phenomena