Experiment Design for Computer Sciences (0AL0400)

Topic 01 - What is an experiment?

Claus Aranha
caranha@cs.tsukuba.ac.jp

University of Tsukuba, Department of Computer Sciences

April 14, 2022

Version 2021.1

Lecture Outline

In this lecture, we talk about the key concepts of this course:

1. What is a scientific experiment?
2. What is science?

Part I: What is Science – A discussion based on Examples.

What is Science?

At the end of your 2-year course, you will receive a diploma that says:

Master in Computer Science

What does that word mean?

Give me **your** answer on manaba!

What is Science?

At the end of your 2-year course, you will receive a diploma that says:

Master in Computer Science

What does that word mean?

Some answers from students in past years

- Science is a method to learn about the world;
- Science is a method to reach the truth;
- Science is useful when it contributes to society;
- Science is how we develop new technologies;

Give me **your** answer on manaba!

What is Science

(Not) answering the question

- All of the answers in the last slide are correct;
- "Science" may mean different things to people, depending on their background;
- But even the different answers have some common characteristics:
  - The discovery of new knowledge;
  - Understanding the natural world;
  - Focus on correctness and methodology;
- There are also some characteristics that are not often discussed
  - Science as a **community**
  - Science as a **continuous process**
  - The relationship between **science and society**

What is Science

I know it when I see it

One way that I like to use to understand science, is to look at people who are doing it, and think about what they do.

I think it is important for us to inspire ourselves on the work of other scientists. It is good to have heroes!

So let's talk about Marie Curie

Marie Curie

Fact Sheet

- From Poland, born in 1867, died in 1934.
- Physicist and Chemist
- Pioneer of radioactivity
- First woman to win the Nobel Prize
- Only woman to win the Nobel Prize Twice
- Only person to win the Nobel Prize
  In two different fields
  (Physics and Chemistry)

Marie Curie

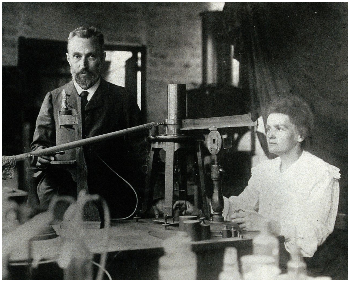
Humble Beginnings

Marie Curie and her sister

- Born in Poland
- Could not enroll at the local university;
  (only accepted men at the time)
- Got educated at the clandestine "Flying University"
- Worked as a tutor and home teacher to sustain herself;

Marie Curie

Moving to Paris



- Moved to Paris and earned a Physics Degree;
- As a researcher, worked in a small shed;
- Had difficulty finding research money (funding);

Marie Curie

Research in Radiology

- In Marie Curie's time, there was a lot of interest in radioactive materials;
  - Why did some materials emit radiation?
  - What was radiation?
  - What could we use it for?
- One of her significant discoveries was that the quantity of radiation depends only on the amount of material;
  - This meant that radiation was an innate property of radioactive material, not something that was acquired.
- Marie Curie did not patent the techniques she discovered to study radioactive materials, so that other scientists could also improve their work, and science could progress even faster.

Marie Curie

Applications of her research

- Observed that tumour cells died more quickly to radiation than healthy cells;
- Developed mobile X-Ray units to be used for surgery during World War I ("little curies")
- Developed "Radium Needles" for sterilizing tissue;

Marie Curie

Legacy

Unfortunately, Marie Curie died early from radiation damage, like many scientists of the time. Her research notebooks are still radioactive, and must be held in special containers!

What can you learn from the history of Marie Curie?

What are the scientists that you know? or that inspire you. What can you learn from their history?

Part II – The Cycle of Science

What is Science?

Scientific Discoveries

Science comes in many different forms.

Let's discuss two interesting and very different scientific discoveries:

- The cosmic background radiation;
- Citrus fruits and scurvy;

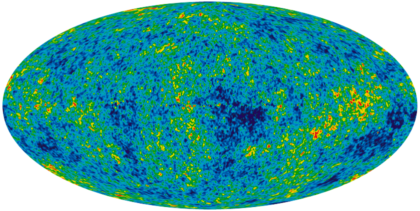
Example 1: The Origins of the Universe

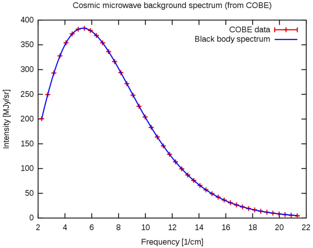
- Physics is a discipline interested in explaining how the universe works. One question of particular interest to physicists is how the universe began.
- One of the most well supported theories for the beginning of the universe is the "inflation theory". According to inflation, the universe started as very dense plasma, and then there was a period of very quick expansion and cooling.
- Why is the "inflation theory" considered to be well supported? How can we know what happened in the beginning of the universe?

Example 1: The Origins of the Universe

What evidence supports "inflation"?

- "Inflation Theory" predicted was that for a very short moment, the universe had expanded enough to be transparent, but was still hot enough to glow.
- The theory allowed a calculation of the duration and intensity of this glow, and that it would be observed from all directions at once.
- Many years later, astronomical radiations actually found evidence for this glow, by accident, which confirmed the theory.





Example 2: Citrus fruits and scurvy

In the 18th century, the British Empire had a very large fleet of ships plundering the entire world.

Maintaining the health of sailors during these long trips was an important issue.

James Lind, a scottish doctor, pioneered several ideas to improve naval hygiene. Among those, he conducted a trial to discover how to prevent **scurvy** among sailors.

Example 2: Citrus fruits prevents scurvy

James Lind (1747):

- Observation: scurvy in sailors;
- Conjecture: Caused by the body rottenning;
- Idea: attempt to avoid/reverse effects with acidic substances;

Separation of a group of 12 affected sailors in six groups with identical diets, except for the addition of a supplement:

Group 1
Cider.

Group 2
Vitriol.

Group 3
Vinegar.

Group 4
Sea water.

Group 5
Oranges and lemons.

Group 6
Tea.

A Framework for Science

From these examples, we see that the scientific process has some common characteristics.

Many of you might have learned the following description in high school or university:

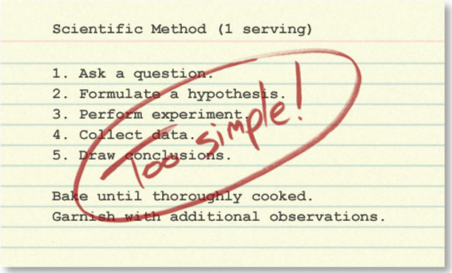
The Scientific Method

1. Observe a Phenomenom
2. Propose a Hypothesis
3. Perform an Experiment
4. Draw conclusions

Is this a good description of "What is Science?"

"The Scientific Method" – Too simple!

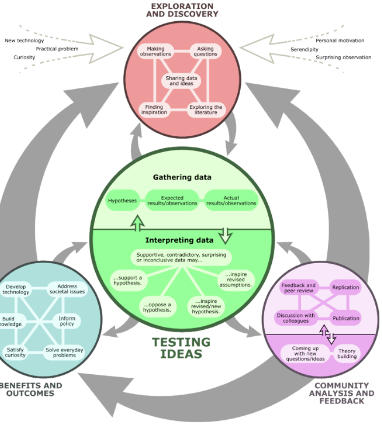
This description of the scientific method has a kernel of truth. But it has some limitations when compared to how science is actually done:



- It does not explain where questions or hypothesis come from.
- What happens when two scientists disagree about the data?
- It assumes that the scientific process ends after the data is collected.
  (How does this new knowledge reach society?)
  (How does this new knowledge influence science?)
- It assumes that old hypothesis or data never gets reviewed.
- etc...

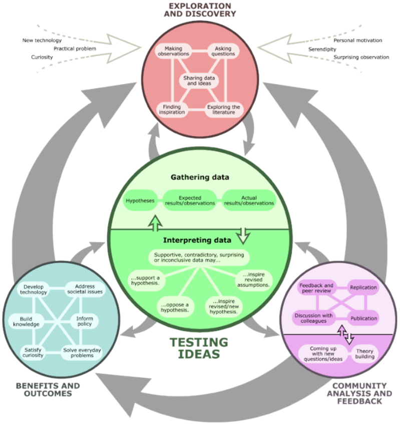
Science as an Interactive Process

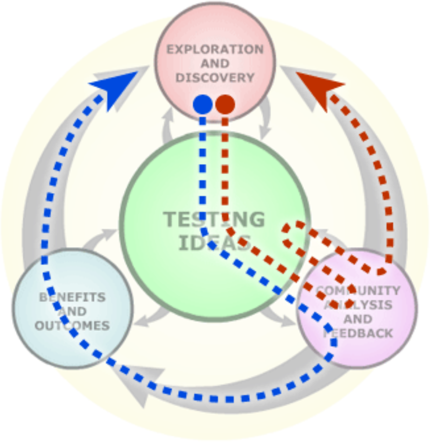
A more complete view of the process of science involves ideas, data, the scientific community, and society, in a continuous feedback circle.



Science as an Interactive Process

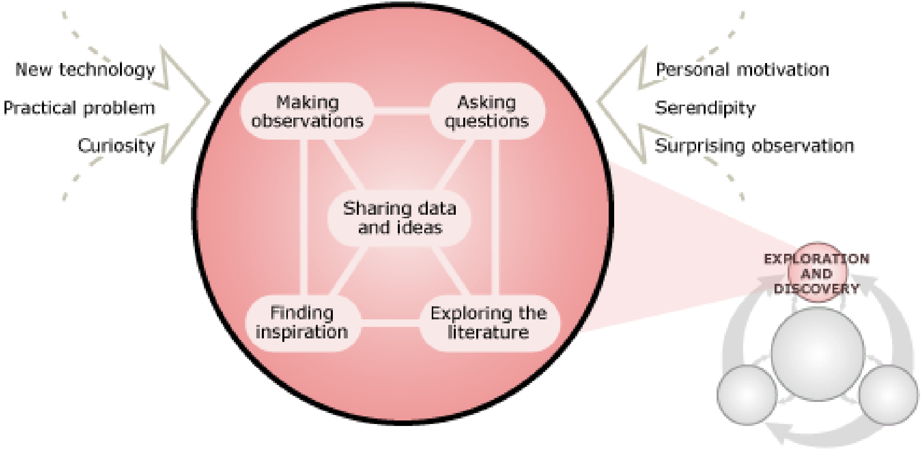
This understanding helps us see that there are many different paths to scientific discoveries. And these paths still have some common characteristics tying them together.





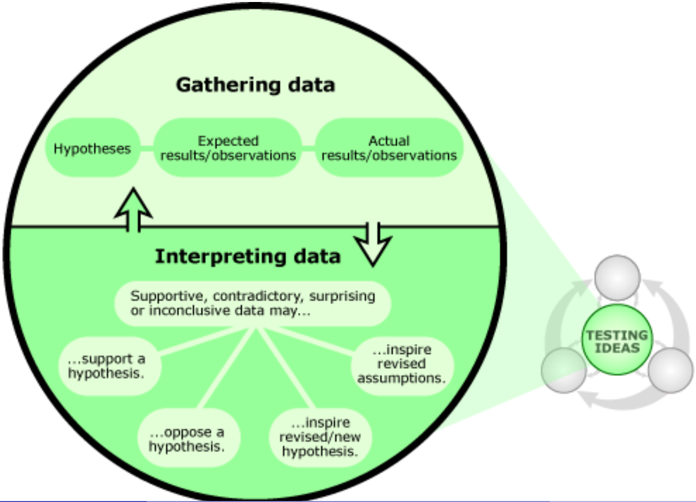
Science as an Interactive Process – Breakdown

Where do ideas come from?



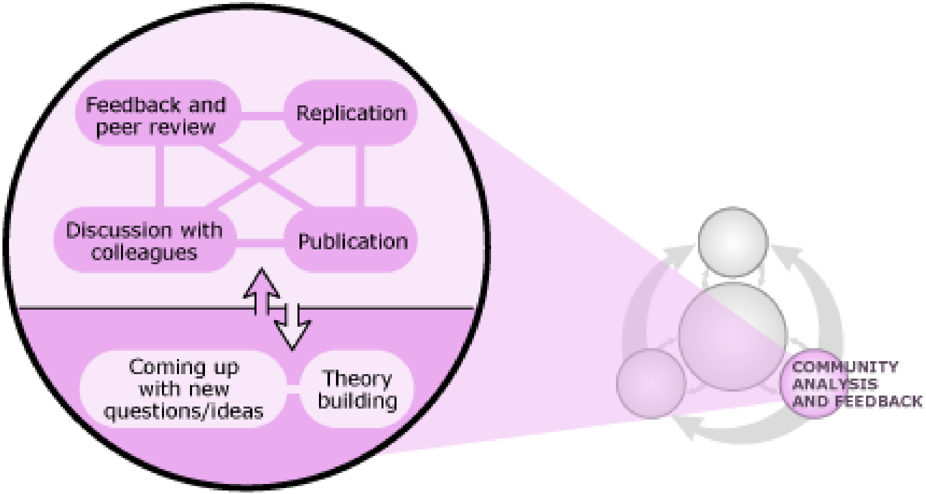
Science as an Interactive Process – Breakdown

Testing and Experimentation



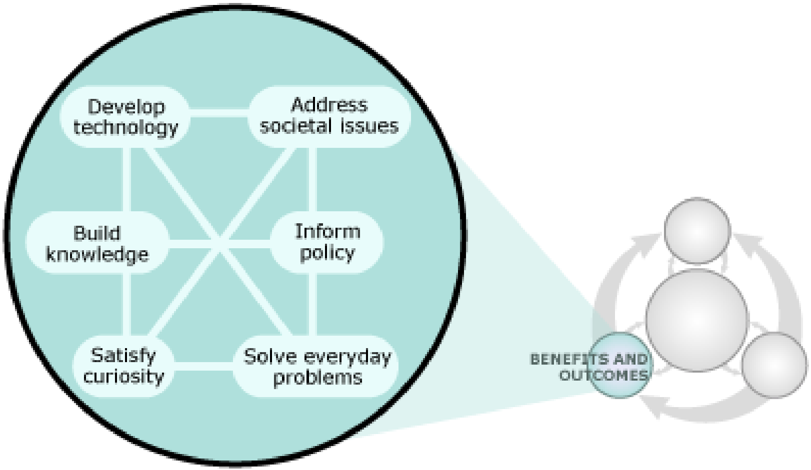
Science as an Interactive Process – Breakdown

The Scientific Community



Science as an Interactive Process – Breakdown

Science and Society



Wrap up – What is Science?

- Science is a method, a way of thinking, and also a community. A process of search and discovery that is continuously changing our society.
- I highly recommend that you read the "Understanding Science" webpage, for a much more in-depth discussion. See the manaba links.
- **What about Computer Sciences**? Popular discussion of science usually focus on physics, biology, social science, etc. How does Computer Science fit in this framework?

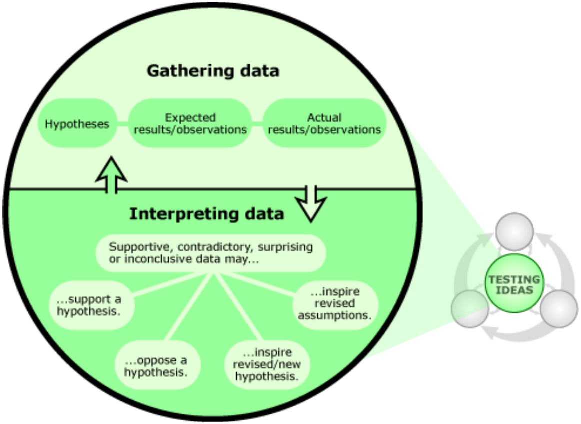
- In the next part of the lecture, we are going to focus on the role of experiments in science, which is also the focus for the rest of this class.

Part III: Experimentalism

The Role of Experiments in Science

Experiments occupy a central role in science. They are how we obtain information that help us answer our scientific questions.

So let's talk about what is an experiment, in detail.



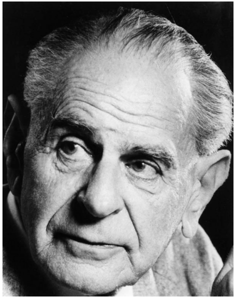
The Role of Experiments in Science

Philosophy and Experimentalism

- Philosophy of science: How can we know things?
  - Can we know things through logic? Can we trust our senses? Is knowledge immutable? etc.

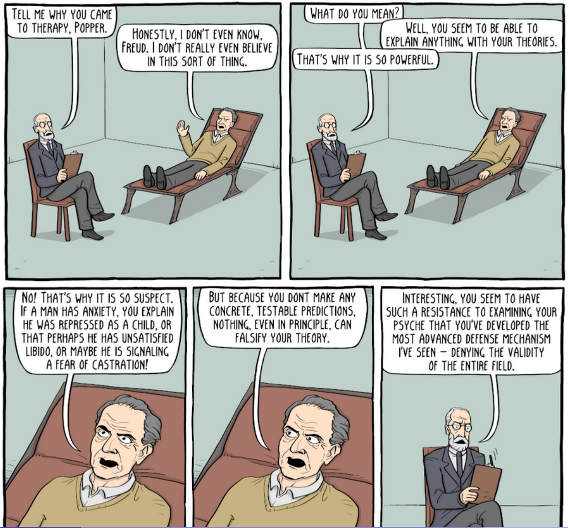
- According to Popper, the only way to obtain knowledge about the world is through rigorous experimentation (Experimentalism);
  1. Formulate a question as a scientific hypothesis;
  2. Execute an experiment following the hypothesis;
  3. Support or reject the hypothesis;

- The definition of scientific hypothesis is important. In Experimentalism, a scientific hypothesis is one that produces falsifiable predictions.



Karl Popper
(1902–1994)

Not everyone agrees with Popper, though...



Defining a "fair" experiment

Generally, we are interested in a "fair" experiment, that provides new evidence for the scientific question that we are trying to answer. A fair experiment has the following characteristics:

- It provides something to compare against;
- It controls the variables of interest;
- It avoids bias in the results;
- It is reproducible;
- It separates chance results from real differences;

"Provides Something to Compare Against"

Scientific Hypothesis

A scientific hypothesis is the question that we try to answer with an experiment. For an experiment to be useful, it is necessary that it has at least two possible results. Afterall, if you already know the result of your experiment, why do it?

Examples of scientific hypothesis:

- **Comparing different systems:**
  Chocolate cookies taste better than mint.
- **Comparing explanations to a phenomenom:**
  Bananas turn black in the refrigerator because of bacteria, not the temperature.
- **Test the characteristic of a system:**
  One apple tree is enough to feed our home for the entire year.

Scientific Hypothesis

Specific Hypothesis

One characteristic of good scientific hypothesis is that they are specific. The more detail you give to the question you want to answer, the easier it is to design a good experiment:

Example:

- "my Optimization method is the best!"
- "The optimization method proposed in this paper is better than existing methods!"
- "Using Polynomial mutation is better than gaussian mutation for optimization methods!"
- "Polynomial mutation increases the convergence rate of an optimization method, when compared to gaussian mutation!"

Scientific Hypothesis

Falsifiable Hypothesis

A good scientific hypothesis is falsifiable. A falsifiable hypothesis is one that **could be proven true or false** by the result of a experiment.

- Non Falsifiable Hypothesis: "This Neural Network can solve any problem if we find the right parameters".
  If the method does not solve the problem, it just means that the parameters are wrong, so we do another experiment... and another... and another...
- Falsifiable Hypothesis: "Using algorithm A and parameters θ, we achieve a speed-up of 1.5 on benchmark B."

An experiment control variables of interest

Types of Experiments

There are many variables that affect the result of an experiment. For example, if you want to test if chocolate ice cream tastes better than coffee, the result may be different if you ask children or PhD students.

If your scientific question is focused on the opinions of PhD students (WHY?), you can control the variables of the experiment, and not include children or professors in your survey.

Depending on how much freedom you have to control the variable in your experiments, we can classify experiments in three types:

- Observational Experiments;
- Retrospective Experiments;
- Controlled Experiments;

Types of Experiments

Observational Experiments

In an Observational Experiment, you obtain data by observing a phenomena without interacting with it directly.

Example: you count the number of people who use the train with and without masks every day.

- Requires care to observe representative situations;
- Allows the researcher to choose general conditions for observation;
- The situation of interest may be too rare to observe naturally;

Types of Experiments

Retrospective Experiments

In a Retrospective Experiment, the researcher obtains data from historical records (newspaper, reports, other scientific papers).

Example: you search from the relationship between announcements of celebrity marriages, and total number of registered marriages;

- Generally cheaper, and may be the only way to gather data over a very long period of time;
- Susceptible to missing records or bias in recording;

Types of Experiments

Controlled Experiments

In a Controlled Experiment, the researcher is able to define several variables in the experiment, and perform it in the conditions desired.

Example: You develop a new algorithm, and test it on some selected data sets, on a collection of different computational architectures;

- Gives a lot of control for the researcher;
- If not designed carefully, allows for the introduction of biases into the experiment;
- Can be the most expensive kind of experiment (although not always in CS);

3. A fair experiment control for biases

"With great power comes great responsibility"

When you control the variables of your experiment (**or when you don't!**). You can introduce biases to the result.

For example, if you compare the speed of two computer programs, it may be hard to see a difference, if you run your experiment in a very powerful computer. Or the difference may be exaggerated if you use a very slow one.

To have a fair experiment, it is important to be aware of these source of biases, and change your experiment to avoid those biases. This is one of the main tasks of Experimental Design.

Some Experiment Design Questions

- Which methods we compare in the experiment?
- Which data sets are used?
- How many times do we interview each participant?
- In what order do we perform the experiments?
- Which data is reported, and how is the data summarized?
- What criteria determines that the hypothesis was accepted or rejected?
- What hyper-parameters do we use?
- How many times is the experiment repeated? How are these repetitions sumarized?

Experiment Design

Example: Controlling for Variation

Let's say you are comparing two computer programs by measuring their running time (wallclock time).

You know that the running time of a program is affected by other programs that are running in the background of the operational system. For example, if a software update happens in the background, it could make a run much slower.

To control for this variation, you make sure to run your experiment in a system with a minimum number of running processes, and you also repeat the experiment many times and take the average running time;

Experiment Design

Example: Controlling for Independence

Imagine that you are comparing two website designs with the following experiment: You measure the time for a user to find some information on website A, then you measure the time for website B.

If you make this comparison always in the same order for all users, you discover that the users are a bit faster for website B, because they get used to the testing environment and are more relaxed.

To remove this influence, you make sure that the test order is always random, or you make sure that each user tests only one website.

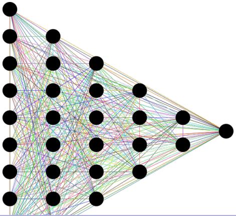
Experiment Design

Example: Controlling for Fairness

You propose a neural network architecture for a new vision problem, and you compare it against traditional architectures.

Because of the special characteristics of the problem, you fine-tune the hyper-parameters of your architecture to achieve the best performance.

To make sure that the comparison is fair, you use the same fine-tune techniques to the traditional architecture that you are comparing against, not using its old hyper-parameters from the literature.



Pre-registered Experiments

Pre-registration is the act of fully defining your research protocol **before you begin to collect or analyze data.**

By pre-registering your research, you avoid modifying your methods to fit your hypothesis (or modifying your hypothesis to fit your data)

Public pre-registration can prevent the loss of negative results. Private pre-registration can help you keep yourself in check.

Learn more: Center for Open Science
`https://cos.io/prereg/`

PREREGISTERED

Reproducible Experiments

Reproducibility is an important property of a good experiment:

- Others can confirm your results;
- Others can build on your results;
- Others can improve your results;
- Society can use your results;

Reproducible Experiments

How can we make experiments more reproducible?

Clear Experiment Design

Detailed steps taken to perform the experiment; Values of relevant parameters; How the results are processed and evaluated;

Open Data and Open Source

Data acquisition protocol is clearly defined; Raw data and pre-processing scripts are available; Data is well documented;

For CS, open source of proposed algorithms is essential;

Open Documentation

Code used for statistical analysis and data visualization;

Summary

Many scientific discoveries and scientific advancements (as well as your own master thesis!) depend on good experiments and data collection.

Unfortunately, good experiments are not as simple as simply going out on the streets and start collecting data.

Careful planning is necessary, not only to make sure that your experiment is not affected by biases, but also to make sure that your experiment is able to answer the question that you are asking.

Experiment Design is the proccess of carefully planning an experiment to make sure that it is fair and effective.

Part III: Outro

Summary of Lecture 01

- Experimentation is a key part of Science;
  - Experiments acquires data that can be used to validate or falsify scientific ideas, and to answer scientific questions;
- An experiment has to be performed carefully to guarantee its usefulness;
  - **Experimental design** defines the type of experiment, and how data is gathered;
  - Several factors can affect the **fairness and meaningfulness** of experiments;
  - **Reproducibility** is essential to guarantee the usefulness of an experiment;

Recommended Reading

Required Reading

- Understanding Science `https://undsci.berkeley.edu/article/intro_01`

Suggested Reading

- Videos: Crash Course Sociology and the Scientific Method, Sociology Research Methods; (link on manaba)
- Existential Comics `http://existentialcomics.com`;

About these Slides

These slides were made by Claus Aranha, 2022. You are welcome to copy, re-use and modify this material.

These slides are a modification of "Design and Analysis of Experiments (2018)" by Felipe Campelo, used with permission.

Individual images in some slides might have been made by other authors. Please see the following references for those cases.

Image Credits I

[Page 7] Marie Curie sketch from `https://www.irasutoya.com`
[Page 8] Curie sisters image from Wikipedia (public domain)
[Page 9] Marie and Pierre Curie image from Wikipedia (Public Domain)
[Page 11] Ambulance from `https://www.irasutoya.com`
[Page 16] CMB images from NASA (public domain)
[Page 18] James Lind image from Wikipedia, Public Domain
[Page 20] "Too simple" image: University of California Museum of Paleontology's Understanding Science
[Page 21] "Real Process of Science 1" image: University of California Museum of Paleontology's Understanding Science
[Page 22] "Real Process of Science 2" image: University of California Museum of Paleontology's Understanding Science

Image Credits II

[Page 23] "Real Process of Science Zoom 1" image: University of California Museum of Paleontology's Understanding Science
[Page 24] "Real Process of Science Zoom 2" image: University of California Museum of Paleontology's Understanding Science
[Page 25] "Real Process of Science Zoom 3" image: University of California Museum of Paleontology's Understanding Science
[Page 26] "Real Process of Science Zoom 4" image: University of California Museum of Paleontology's Understanding Science
[Page 30] Karl Popper picture from Wikipedia
[Page 31] "Freud and Popper" from `http://existentialcomics.com/`
[Page 42] Clock image from `https://pixabay.com`
[Page 43] Website images from `https://www.irasutoya.com`
[Page 44] Neural Network image by Gordon Johnson from `https://pixabay.com`