Scientific Method and Scientific Experiments

Instituto de Computação
UFF

Vanessa Braganholo

E-Science

E-Science is the application of **computer technology** to the undertaking of modern **scientific investigation**, including the preparation, experimentation, data collection, results dissemination, and long-term storage and accessibility of all materials generated through the **scientific process**.

Source: Bohle, S. "What is E-science and How Should it Be Managed?", Nature.com, Spektrum der Wissenschaft (Scientific American), http://www.scilogs.com/scientific_and_medical_libraries/what-is-e-science-and-how-should-it-be-managed

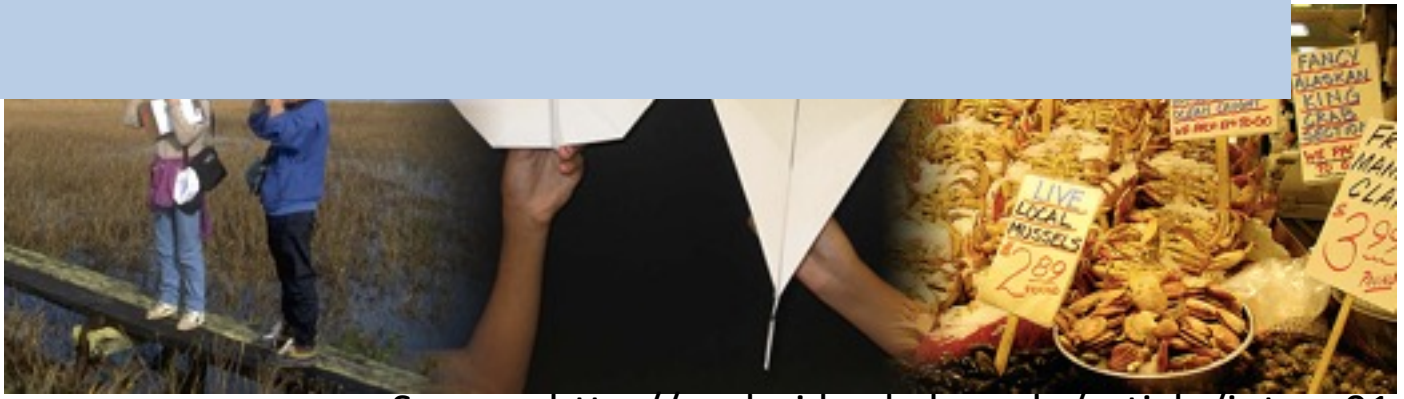
Agenda

- What is Science?
- Scientific Method
- Scientific Experiments

What comes to mind when you hear the word Science?

Source: http://undsci.berkeley.edu/article/intro_01

These images reflect only some aspects of Science, since it has several facets



Source: http://undsci.berkeley.edu/article/intro_01

What is Science?

- Science is both a **body of knowledge** and a **process**
  - It is not only isolated and static facts listed in a textbook
  - It is also a process of discovery that allows us to link isolated facts into coherent and comprehensive understandings of the natural world

Source: http://undsci.berkeley.edu/article/article/0_0_0/whatisscience_01

Other facets of Science

- Science is **exciting**
- Science is **useful**
- Science is **ongoing**
- Science is a **global human endeavor**

Source: http://undsci.berkeley.edu/article/article/0_0_0/whatisscience_01

Science is exciting

- Science is a way of discovering what's in the universe and how those things work today, how they worked in the past, and how they are likely to work in the future
- Scientists are motivated by the thrill of seeing or figuring out something that no one has before

Source: http://undsci.berkeley.edu/article/article/0_0_0/whatisscience_01

Science is **useful**

- The knowledge generated by science is powerful and reliable
- It can be used to develop new technologies, treat diseases, and deal with many other sorts of problems

Source: http://undsci.berkeley.edu/article/article/0_0_0/whatisscience_01

The Brick Wall of Knowledge

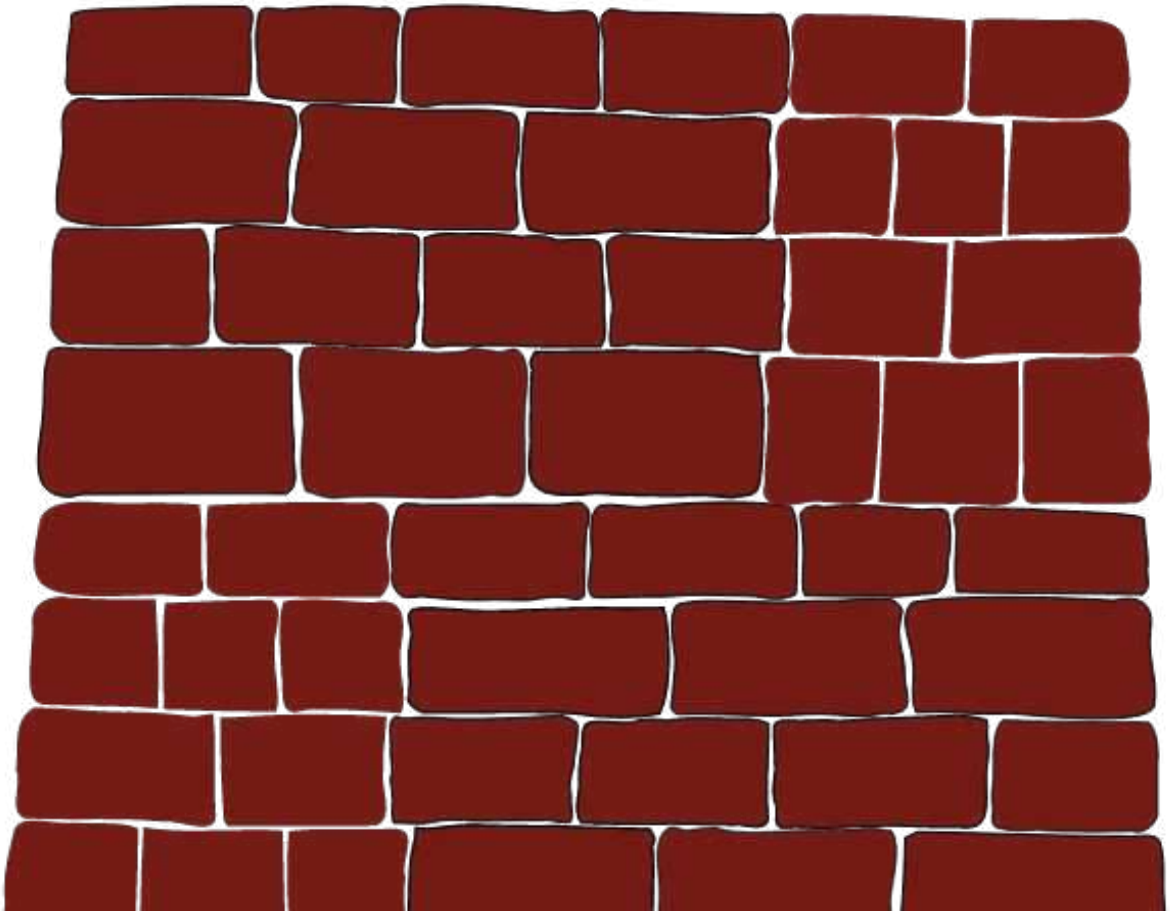


Image Source: https://clipartfest.com/categories/view/cc90dca431547f9c9d22265d24142adced0b8cf4/brick-wall-clip-art.html

Science is **ongoing**

- Science is continually refining and expanding our knowledge of the universe, and as it does, it leads to new questions for future investigation
- Science **will never be "finished"** (the wall to be built is infinite)

Source: http://undsci.berkeley.edu/article/article/0_0_0/whatisscience_01

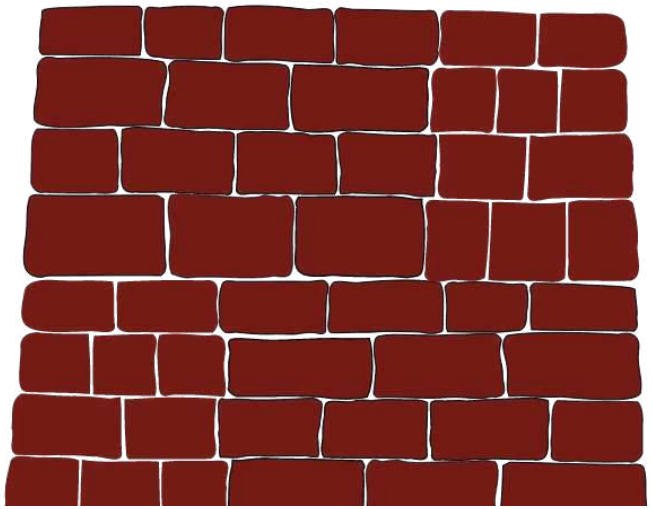
Science is a global human endeavor

- People all over the world participate in the process of science
- And you can too!

Source: http://undsci.berkeley.edu/article/article/0_0_0/whatisscience_01

How can I participate?

How can I add a brick to the Wall of Knowledge?



Every baby knows the scientific method!

Scientific Method

1. Make an Observation
2. Ask a Question
3. Formulate an Hypothesis
4. Make a prediction based on the hypothesis
5. Test the prediction
6. Iterate: use the results to make new hypotheses or predictions

Source: https://www.khanacademy.org/science/biology/intro-to-biology/science-of-biology/a/the-science-of-biology

Example: Fail to Toast

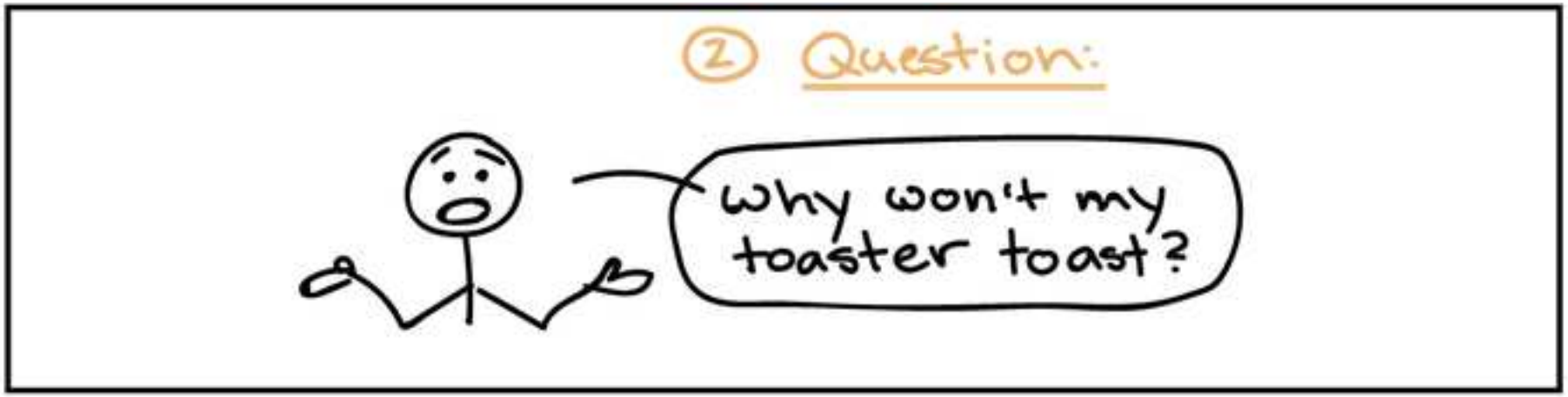
1. Make an observation

Source of the example: https://www.khanacademy.org/science/biology/intro-to-biology/science-of-biology/a/the-science-of-biology

Example: Fail to Toast

2. Ask a question

Why didn't my bread get toasted?



Example: Fail to Toast

3. Propose a hypothesis

A *hypothesis* is a potential answer to the question, one that can somehow be tested.

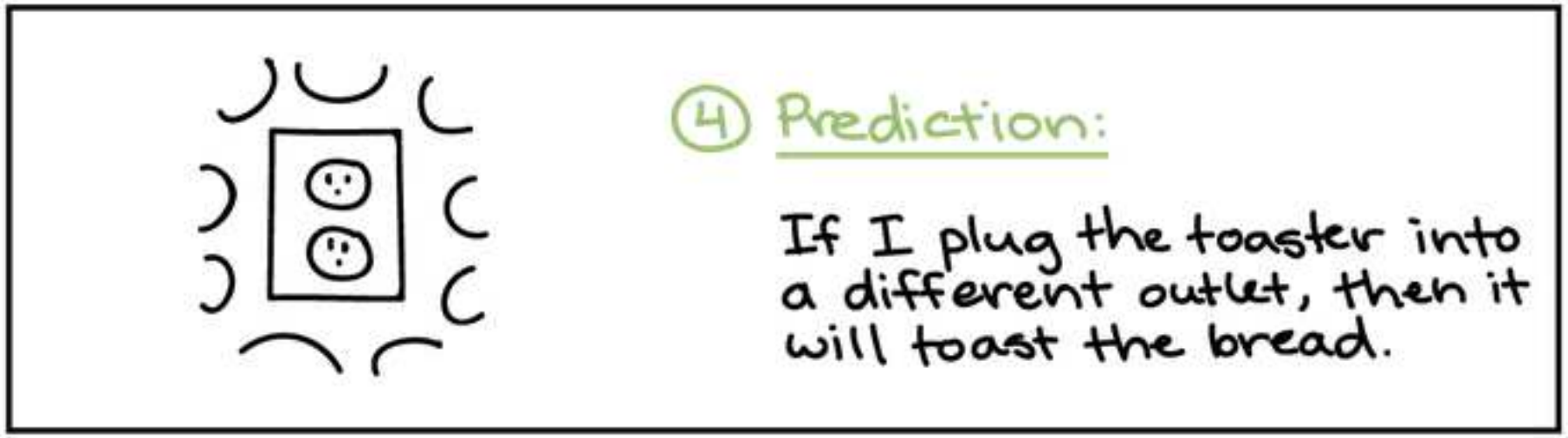
Hypothesis

- A hypothesis **is not necessarily the right explanation**
- Instead, **it's a possible explanation that we can test** to see if it is likely correct, or if we need to make a new hypothesis

Example: Fail to Toast

4. Make a prediction

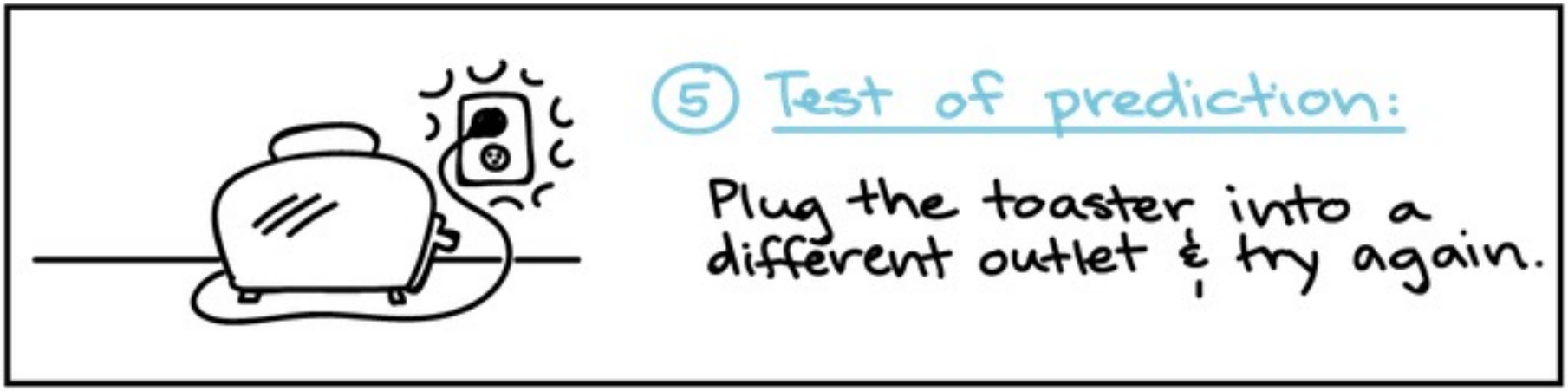
A prediction **is an outcome we'd expect to see if the hypothesis is correct**.



Example: Fail to Toast

5. Test the prediction

To test the hypothesis, we need to **make an observation or perform an experiment** associated with the prediction.



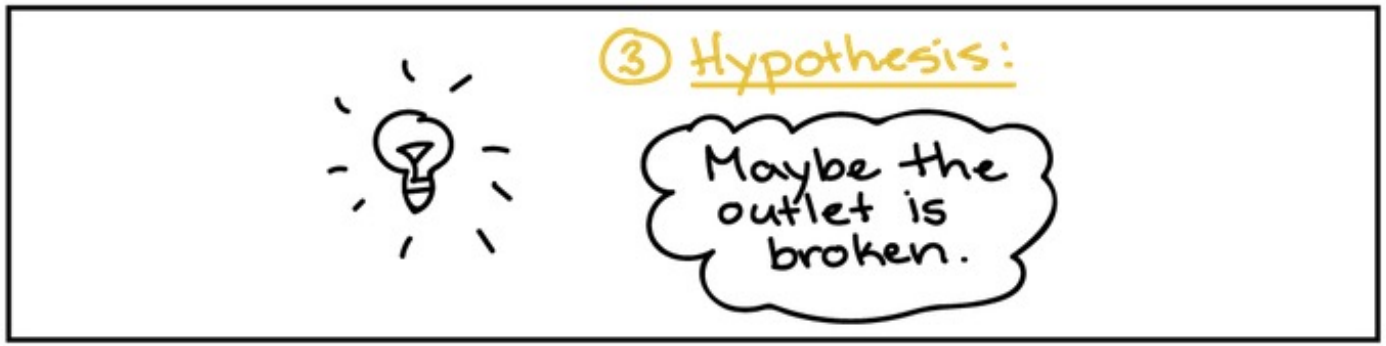
Test the Prediction Outcomes

- If the toaster does toast, then the hypothesis is supported—likely correct
- If the toaster doesn't toast, then the hypothesis is not supported—likely wrong

Supporting or Contradicting a Hypothesis

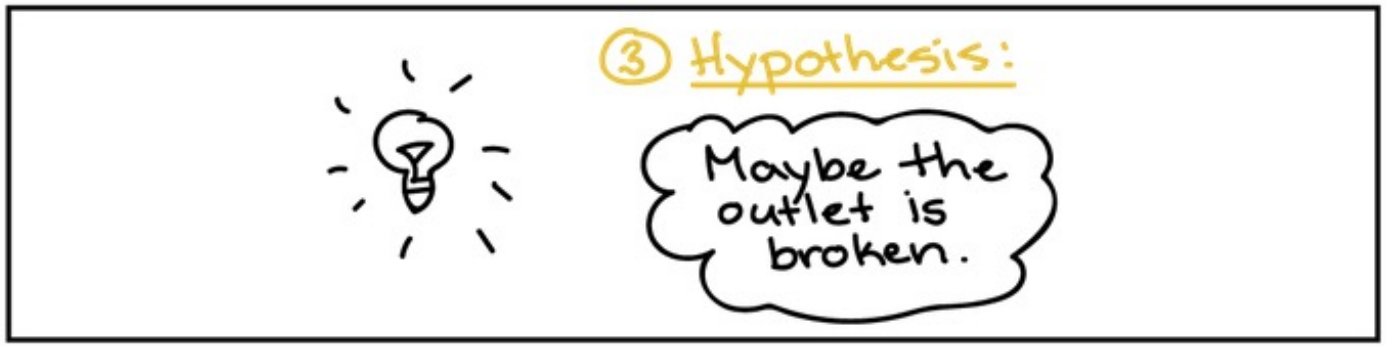
- **Results that support** a hypothesis can't conclusively prove that it's correct, but **they do mean it's likely to be correct**
- If **results contradict a hypothesis**, that hypothesis is probably not correct. Unless there was a flaw in the test—a possibility we should always consider—a contradictory result means that **we can discard the hypothesis and look for a new one**.

What can we say about the Hypothesis…



	First outlet working?	Second outlet working?
Toaster does work		
Toaster doesn't work		

What can we say about the Hypothesis…

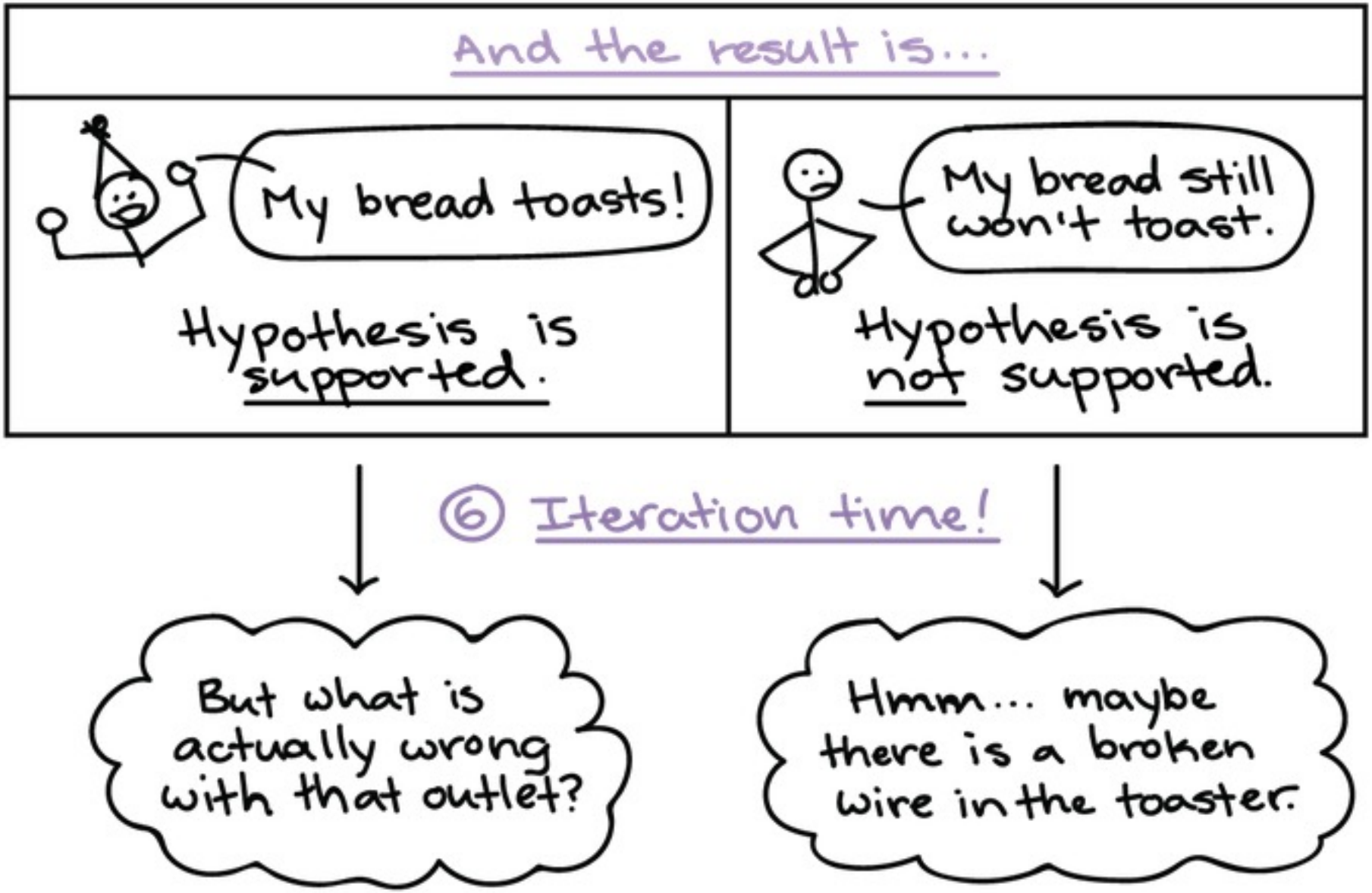


	First outlet working?	Second outlet working?
Toaster does work	Probably no	Probably yes
Toaster doesn't work	Cannot tell	Cannot tell

Example: Fail to Toast

6. Iterate

The last step of the scientific method is to reflect on our results and use them to guide our next steps.

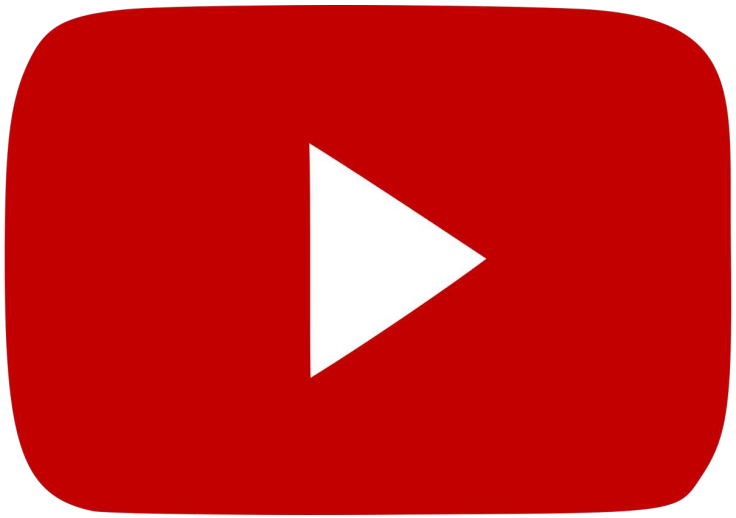

And the result is...
My bread toasts!
Hypothesis is supported.
My bread still won't toast.
Hypothesis is not supported.
⑥ Iteration time!
But what is actually wrong with that outlet?
Hmm... maybe there is a broken wire in the toaster.

If toaster still doesn't work

- New hypotheses:
  - Both outlets are working and the toaster is broken
  - Both outlets are defective and the toaster is working
- New prediction:
  - If I plug another device in the outlets, then it will turn on.

And so on…

Video



Link to the video: https://www.youtube.com/watch?v=N6IAzlugWw0

Making Assumptions

- All scientific tests involve making assumptions

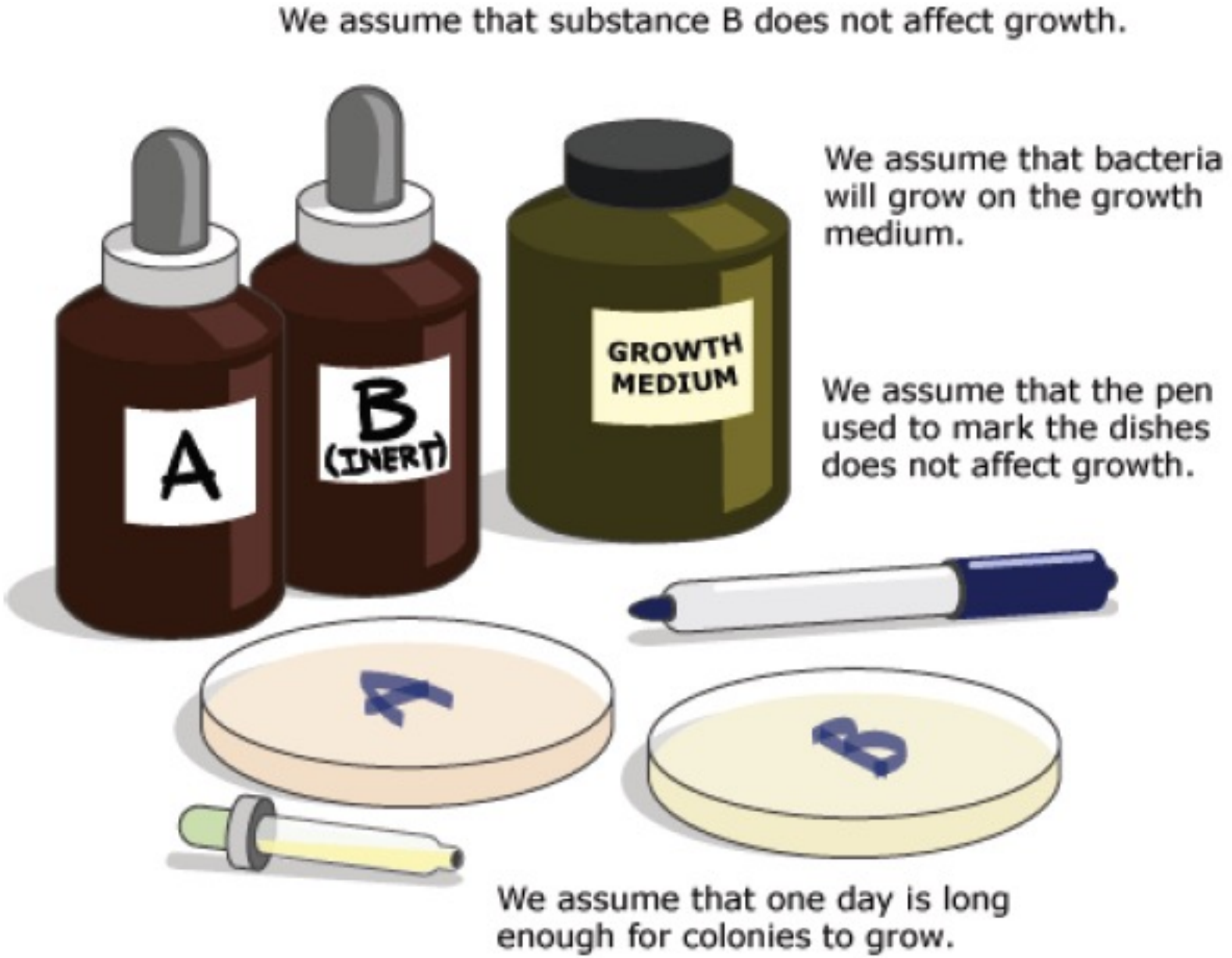
Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Example of a Hypothesis Test

- Suppose we want to test the hypothesis that **substance A stops bacterial growth**
- To test the hypothesis, we spread a mixture of substance A and some bacterial growth medium in a Petri dish, and spread a mixture of the inert substance B and some bacterial growth medium in another Petri dish
- We then wait 1 day to check the results

Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Even a fairly straightforward experiment will rely on some assumptions:



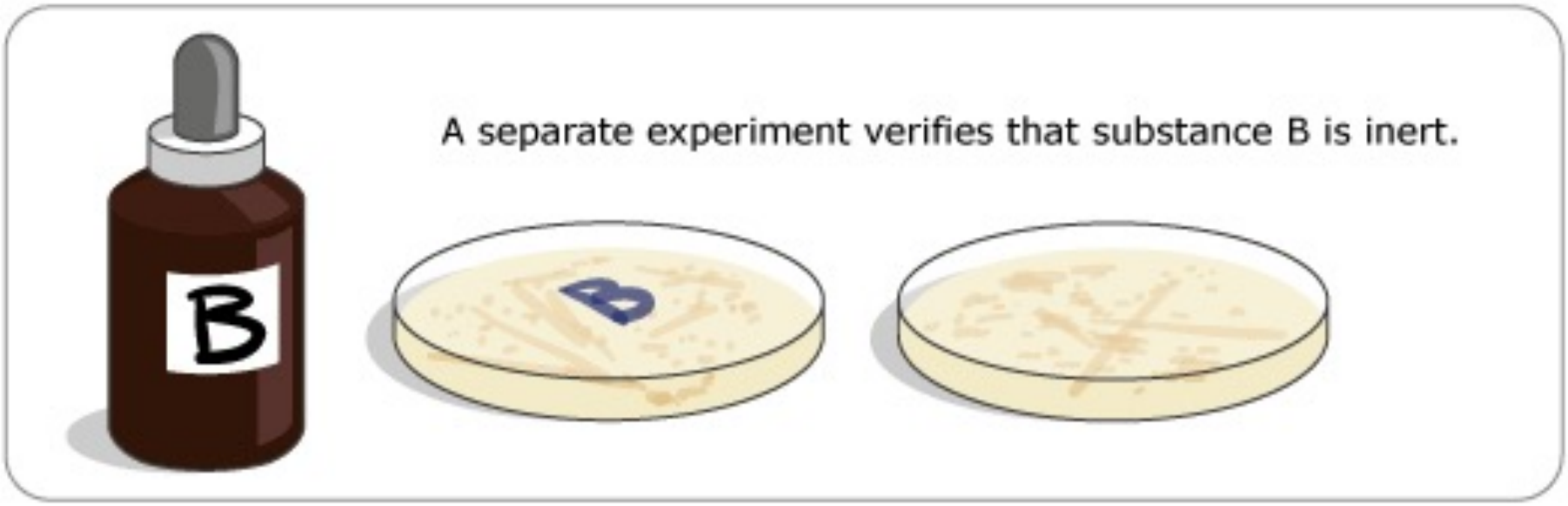
Source: http://undsci.berkeley.edu/article/0_0_0/Howscienceworks_01

Making Assumptions

- The previous example contained perfectly reasonable assumptions that **can be tested**
- All the assumptions **need to be justified** with
  - New experiments
  - Past tests performed by other scientists
- Some assumptions may remain untested because our knowledge in the field suggests that assumption is safe

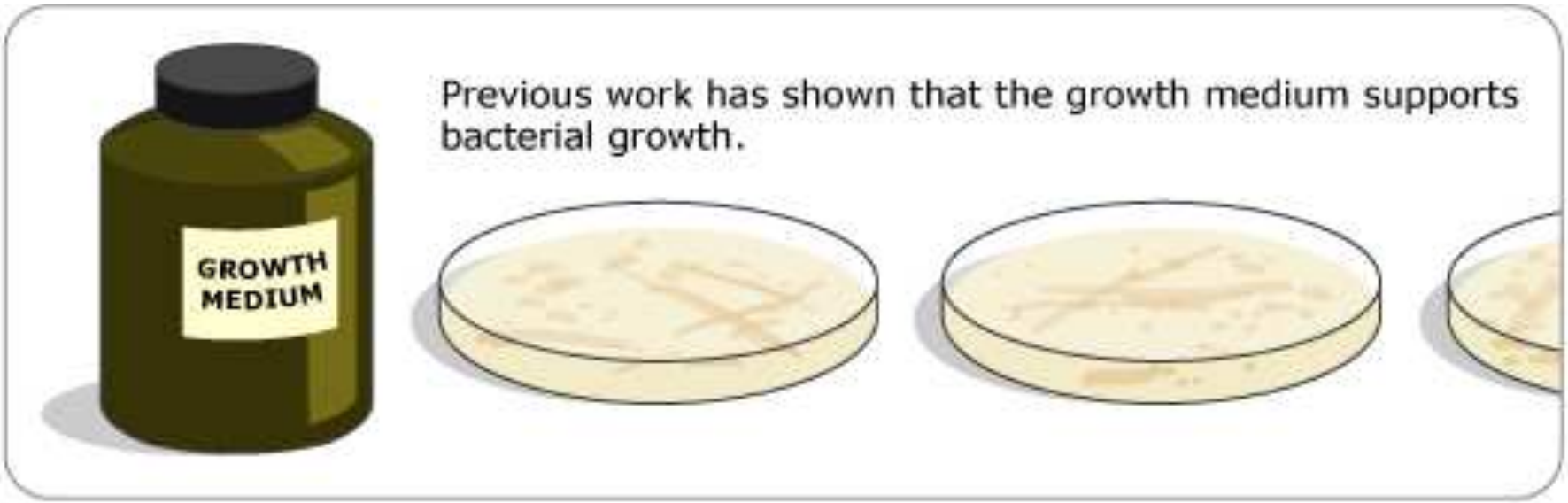
Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Check Assumptions



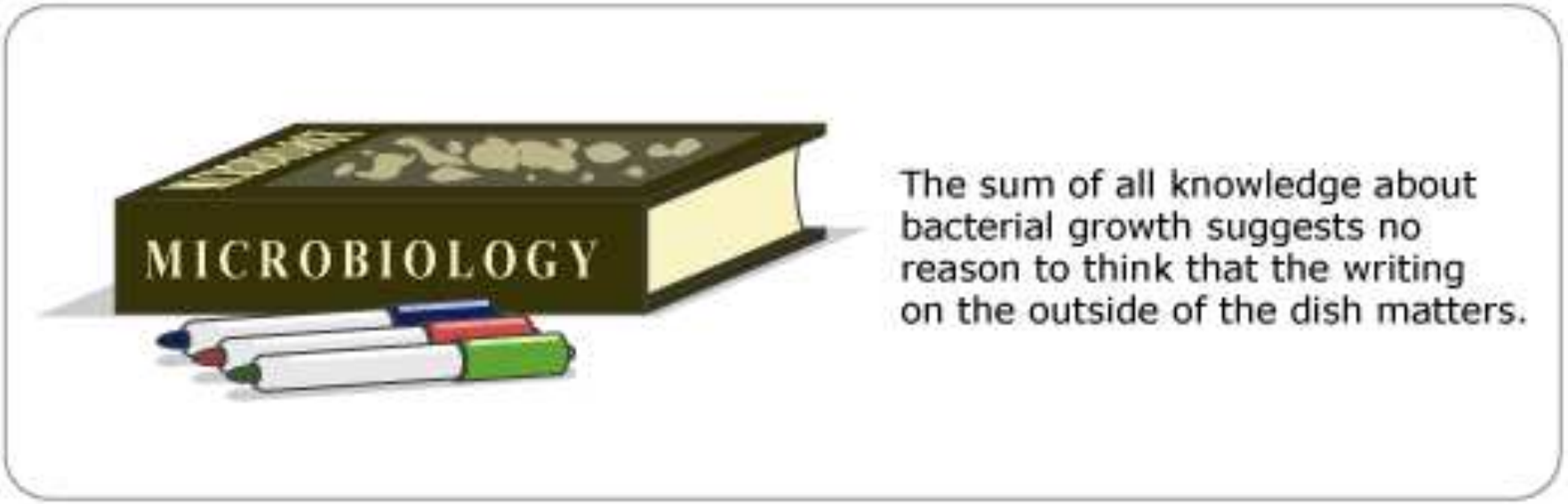
Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Check Assumptions



Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Check Assumptions



Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Observation beyond our eyes

- Making an observation does not mean we have to "see it with our own eyes" – in some cases, this may be impossible
- Several tools can be used to make observations: telescopes, microscopes, all kinds of sensors, radar, thermometers, etc.
- **Precision is a key** when making observations

Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Experiments

- An experiment is a test that involves **manipulating some factor** in a system in order to see how that affects the outcome
- They can be as simple as rolling a ball at different angles to check how that affects the speed of the ball, or as complex as removing a species from a controlled habitat to see how that affects other species

Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Types of Experiments

- ***in vivo*** (from the Latin *within the living*)
  - An experiment that is conducted on living organisms, usually in their own habitats
- ***in vitro*** (from the Latin *within the glass*)
  - An experiment that is conducted in a controlled environment, such as a laboratory
- ***in virtuo***
  - An experiment that is conducted in the form of a computer simulation, but allow agents to change the course of the simulation
- ***in silico***
  - An experiment where everything is simulated

Amigoni and Schiaffonati ”Multiagent-Based Simulation in Biology”. In: Model-Based Reasoning in Science, Technology and Medicine, 2007.

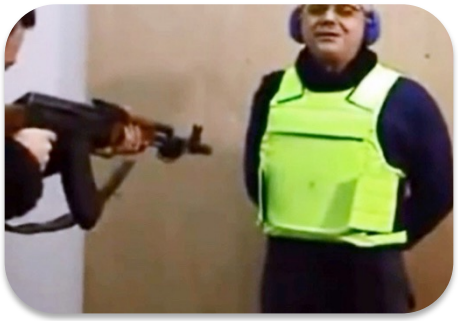
In the beginning

In-vivo

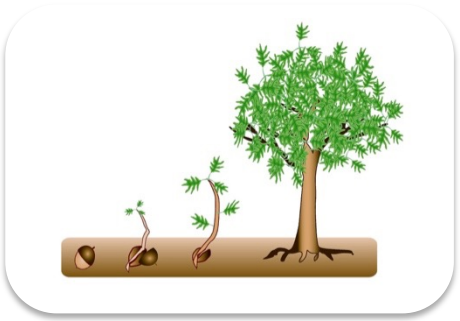
In-vitro

However in-vivo and in-vitro experiments are...

Costly



Risky

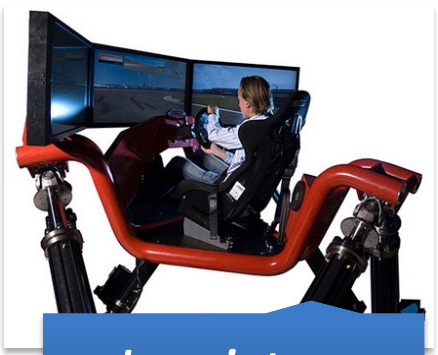


Slow

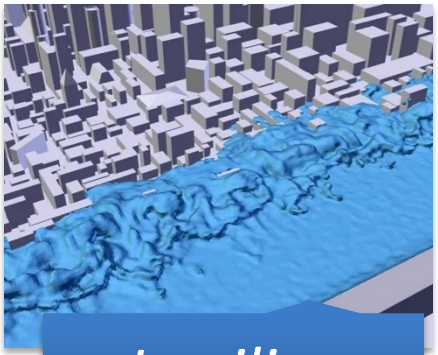
From in-vivo to *in-silico*

In-vivo

In-vitro



In-virtuo



In-silico

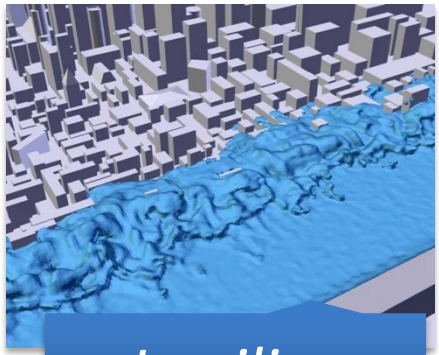
Travassos and Barros "Contributions of in virtuo and in silico experiments for the future of empirical studies in software engineering." WSESE 2003

Dealing with produced data

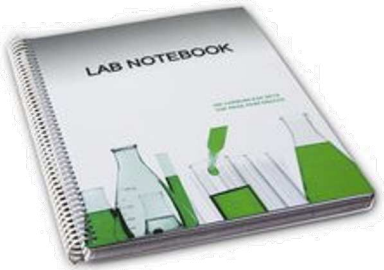
In-vivo

In-vitro

In-virtuo



In-silico



Data volume

Travassos and Barros "Contributions of in virtuo and in silico experiments for the future of empirical studies in software engineering." WSESE 2003

E-Science

- E-Science is closely related to ***in-virtuo*** and ***in-silico*** experiments
- It can also support other kinds of experiments, especially in the data analysis and acquisition

Observations and Tests leads to data

- Observations and tests yield data
- **Scientists analyze and interpret data** in order to figure out how those data inform their hypotheses and theories

Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

Digging into data

- Analyzing data is not a simple task
- Scientists usually use statistical calculations, tabulations, visual representations, computer models, etc.

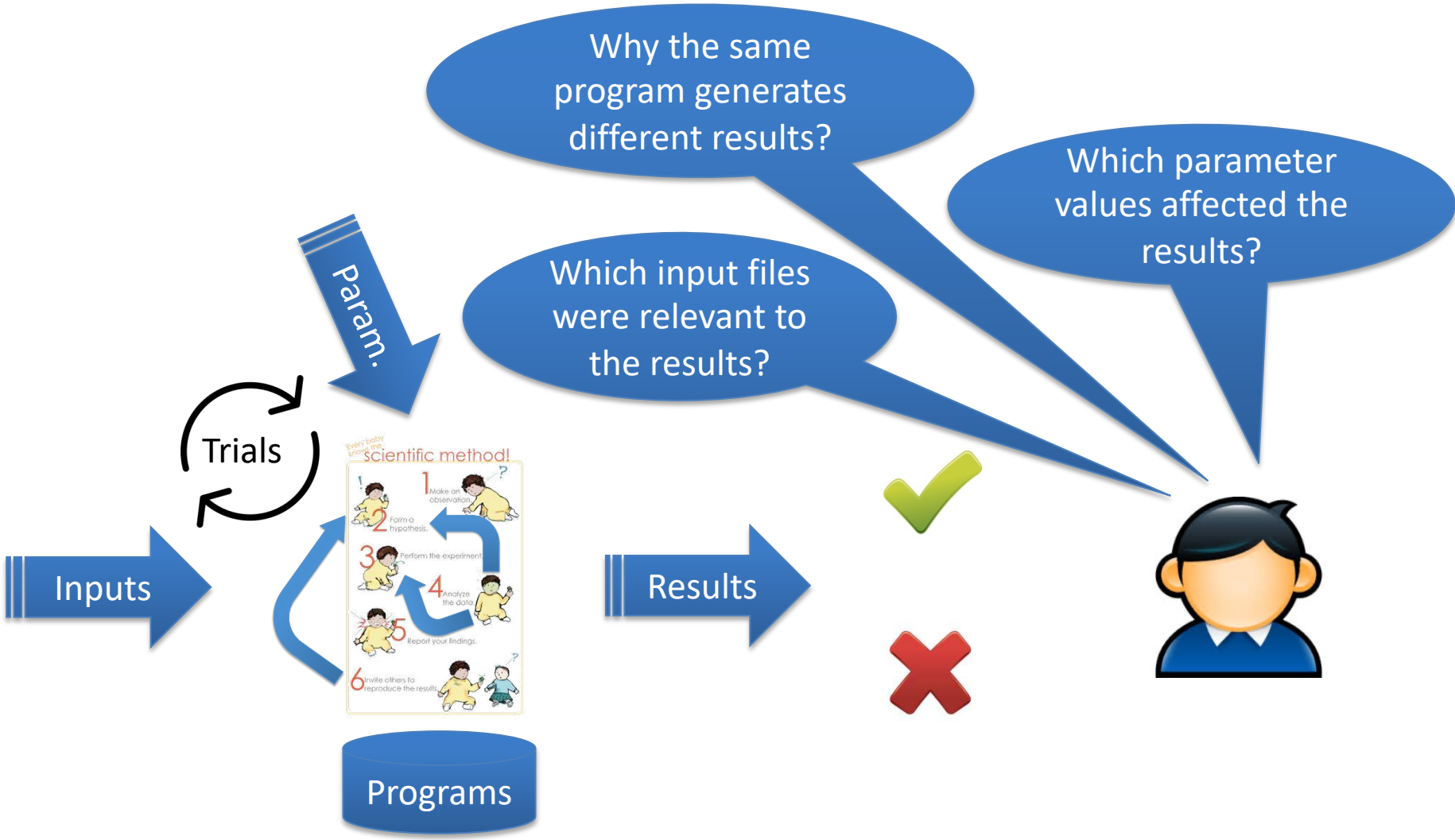
Source: http://undsci.berkeley.edu/article/0_0_0/howscienceworks_01

No	Ranking dos algoritmos GRASP, desde o de melhor desempenho até o pior	Quantidade de instâncias com melhor valor			Média do gap
		mínimo	médio	máximo	
1	GRASP (det LS) + backward PR (eliteSize: 10) + restart (50 it)	20	19	18	0,00190
2	GRASP (det LS) + forward PR (eliteSize: 10)	20	18	19	0,00268
3	GRASP (det LS) + forward PR (eliteSize: 10) + restart (150 it)	20	18	19	0,00268
4	GRASP (det LS) + forward PR (eliteSize: 10) + restart (100 it)	20	16	17	0,00271
5	GRASP (det LS) + backward PR (eliteSize: 10) + restart (100 it)	20	16	16	0,00288
6	GRASP (rand LS) + backward PR (eliteSize: 10) + restart (100 it)	20	17	17	0,00302
7	GRASP (rand LS) + backward PR (eliteSize: 10) + restart (150 it)	20	18	18	0,00318
8	GRASP (det LS) + forward PR (eliteSize: 10) + restart (50 it)	20	18	19	0,00319
9	GRASP (det LS) + backward PR (eliteSize: 10) + restart (150 it)	20	17	17	0,00325
10	GRASP (rand LS) + backward PR (eliteSize: 10)	20	18	18	0,00328
11	GRASP (rand LS) + forward PR (eliteSize: 10)	19	17	18	0,00339
12	GRASP (det LS) + backward PR (eliteSize: 10)	20	16	16	0,00348
13	GRASP (rand LS) + backward PR (eliteSize: 10) + restart (50 it)	20	17	17	0,00352
14	GRASP (rand LS) + forward PR (eliteSize: 10) + restart (150 it)	19	17	18	0,00357
15	GRASP (rand LS) + forward PR (eliteSize: 10) + restart (100 it)	19	18	18	0,00399
16	GRASP (rand LS) + forward PR (eliteSize: 10) + restart (50 it)	20	16	16	0,00470
17	GRASP (alpha: 0.19, rand LS)	19	14	14	0,01669
18	GRASP (det LS) + external PR (eliteSize: 10) + restart (100 it)	19	15	15	0,01905
19	GRASP (det LS) + external PR (eliteSize: 10) + restart (50 it)	19	13	13	0,01917
20	GRASP (det LS) + external PR (eliteSize: 10) + restart (150 it)	19	14	14	0,01947
21	GRASP (det LS) + external PR (eliteSize: 10)	19	14	14	0,02006
22	GRASP (rand LS) + external PR (eliteSize: 10) + restart (50 it)	18	13	13	0,02405
23	GRASP (alpha: 0.19, det LS)	19	14	14	0,02409
24	GRASP (rand LS) + external PR (eliteSize: 10) + restart (150 it)	18	14	14	0,02438
25	GRASP (rand LS) + external PR (eliteSize: 10) + restart (100 it)	18	14	14	0,02442
26	GRASP (rand LS) + external PR (eliteSize: 10)	16	14	14	0,02490

Instituto de Computação

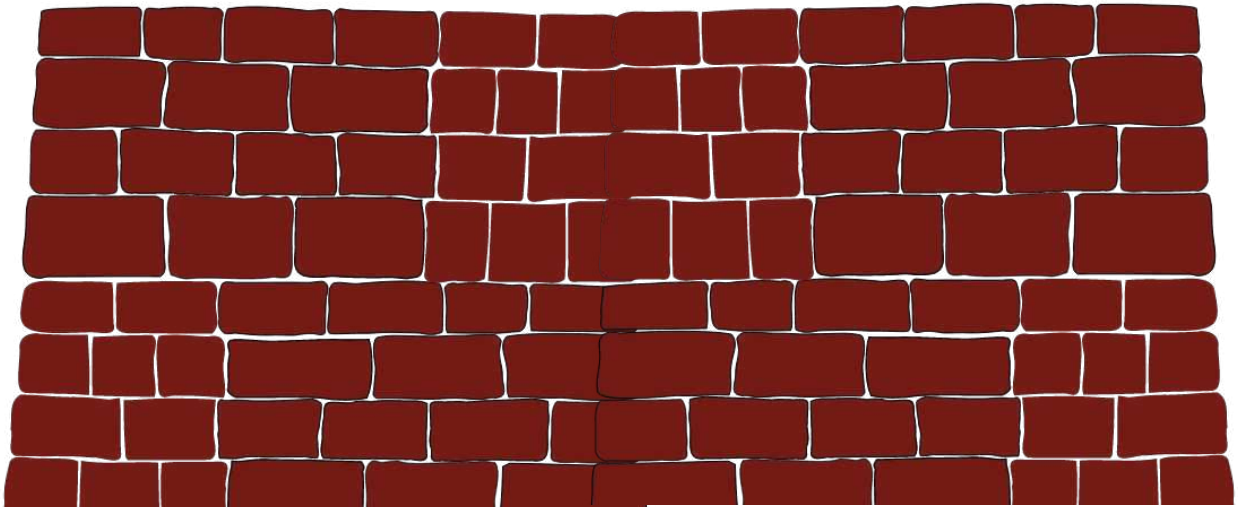
Dados de Ruslán Guerra Marzo

How to reason from the data?



Building a Strong Wall of Knowledge

- Accepted scientific ideas are reliable because they have been **subjected to rigorous testing**, but as new evidence is acquired and new perspectives emerge these ideas can be revised



Publishing

- Publishing results is a way of **making your findings available** so that other researchers can build upon them
- Publications describe a study and report any details that one might need to evaluate that study (background information, data, statistical results, graphs, maps, explanations of how the study was performed and how the researchers drew their conclusions)

Publishing

- The story does not need to be told **exactly** in the order it happened
- Changing the order of things may, in some cases, make your work easier to understand

UNTANGLING A TWISTED PATH

Walter Alvarez

In 1980, in the journal *Science*, Walter Alvarez and his colleagues published a scientific article describing their controversial new hypothesis that the dinosaur extinction was triggered by a massive asteroid impact. Despite its splashy and novel topic, the article laid out its hypothesis and evidence in the conventional way—linearly—which allowed colleagues in geology and paleontology to quickly understand and evaluate the research. Though helpful for scientific communication, this linear presentation can give the impression that an investigation has been plotted out from the beginning—but in fact, Alvarez's study was far from linear. He stumbled onto his hypothesis unexpectedly, originally setting out to study the tectonic movements of the Italian peninsula. After an intriguing series of twists, turns, false starts, inspirations, and rejected hypotheses, he and his colleagues found that they had completed a rather different, but compelling, investigation.

Publishing

- Publishing is also “quality control” of your research
- Publications go through a rigorous process of peer-review (journal or conference)

Conference

Autor

Chair

Program Committee
Reviewers

Conference

BEFORE THE DEADLINE

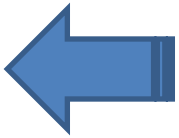
Conference

Conference

Conference

Author

Results



PC Chair

Conference

Results

ACCEPT
REBUTTAL*
REJECT

Journal

Author

Editor

Reviewers

Journal

ANYTIME
(CONTINUOS FLOW)

Journal

Journal

Journal

Author

Results

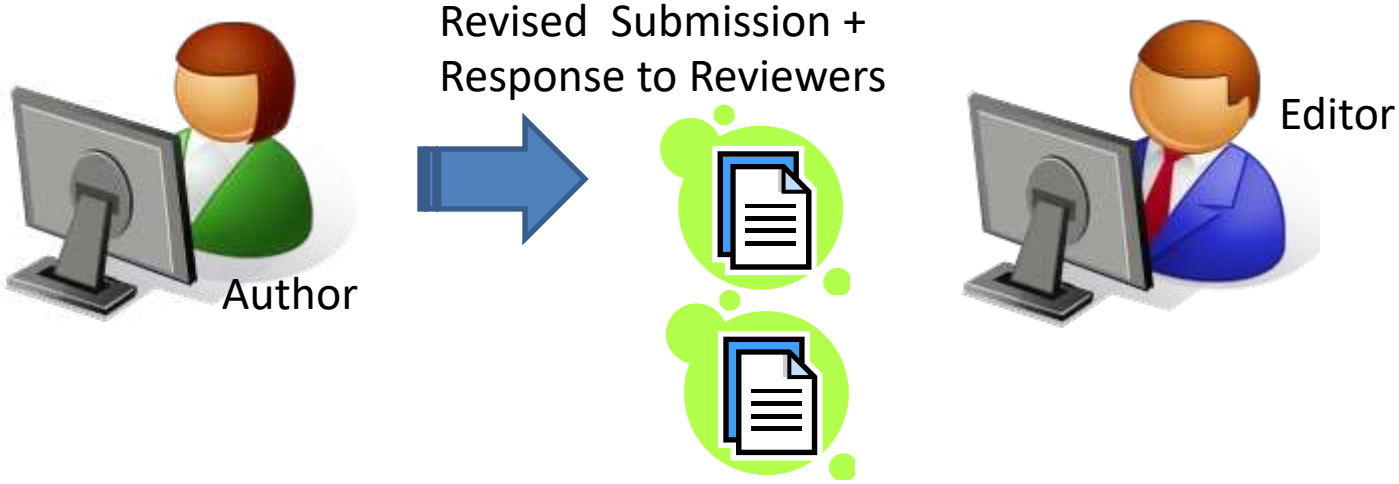
Editor

Journal

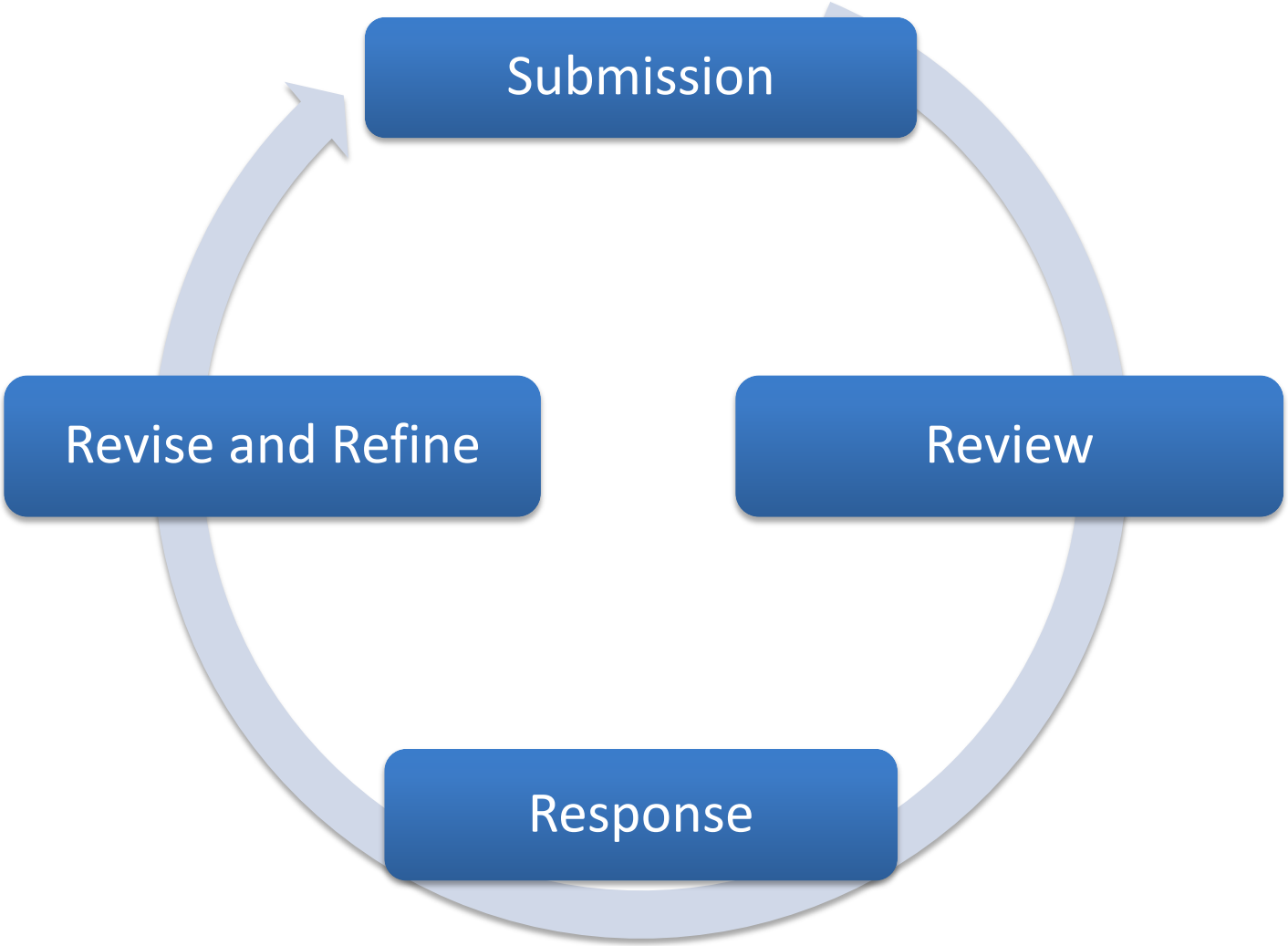
Results

ACCEPT
MINOR REVIEW
MAJOR REVIEW
SUBMIT AS NEW
REJECT

Journal



WITHIN THE DEADLINE ESTABLISHED BY THE EDITOR


Submission
Review
Response
Revise and Refine

Exercise

- You are a scientist and want to study the effects of low carb and low fat diets on patients with high cholesterol to find out which one works better to reduce cholesterol.
- How would you address this investigation in light of the scientific method?
  - What are the hypotheses you could formulate?
  - How would you design experiments to test your hypotheses?

Exercise

- You designed a new sort algorithm in your thesis. Your hypothesis is that your algorithm performs better than the state of the art algorithms.
- How would you design the experiments for your thesis?

Bibliography

- The scientific method, Khan Academy
- How Science works, Understanding Science
- Wikipedia, Scientific Method
- MATTOSO, M. L. Q. ; WERNER, C. M. L. ; TRAVASSOS, G. H. ; **BRAGANHOLO, V.** ; MURTA, L. G. P. ; OGASAWARA, E. ; OLIVEIRA, D. ; CRUZ, S. ; MARTINHO, W. Towards Supporting the Life Cycle of Large Scale Scientific Experiments. International Journal of Business Process Integration and Management (Print), v. 5, p. 79-92, 2010.

Scientific Method and Scientific Experiments

Instituto de Computação
UFF

Vanessa Braganholo

Hypothesis x Theories x Models

- **Hypothesis** are proposed explanations for a fairly **narrow set of phenomena**
  - Example: a certain species of butterfly A evolved from a result of hybridization of two other butterfly species B and C

Hypothesis x Theories x Models

- **Theories** are **broad explanations** for a **wide range of phenomena**
  - They often integrate and generalize several hypothesis
  - Example: the theory of natural selection

Hypothesis x Theories x Models

- Some Hypothesis and Theories may be so complex that are best **described by a computer program or mathematical equation**
- In such cases, they are called a **Model**