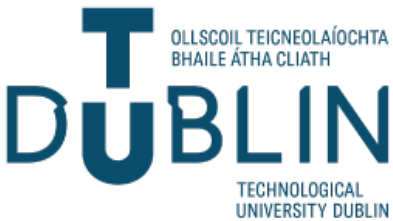
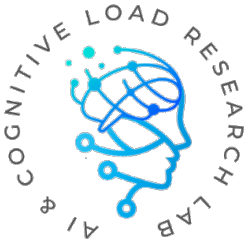
Demystifying the research hypothesis

Dr. Luca Longo

Artificial Intelligence and Cognitive Load Research Lab
School of Computer Science, Technological University Dublin

ADAPT, The global centre of excellence for digital content and media innovation
AIRC, The Applied Intelligence Research Center
D-Real, SFI center for research training in digitally-enhanced reality
CeADAR, EI Ireland's centre for applied Artificial Intelligence
ML-Labs, SFI center for research training in machine learning





✉ **luca.longo@tudublin.ie**
http://www.lucalongo.eu/

A research hypothesis is the most important element of a **scientific research project**.

It is a **predictive testable statement**

Hypothesis: a precise statement

IF _____ [*I do this*] _____,
THEN _____ [*this*] _____ will happen.

Hypothesis and directionality

A hypothesis can be directional or non-directional.

- **directional**: it indicates the nature and direction of the relationship/difference between variables
- **non-directional**: it only states that relationship/difference will occur

Hints for forming hypotheses

- single sentence
- simply stated
- at least one variable
- variable/s clearly stated
- relationship/difference precisely stated
- testable

A research hypothesis is always tentative and it is the focus of the research.

What you think will happen

What you ‘think’ will happen, of course, must be based on your understanding of the science (through a literature review) and scientific principles involved in the experiment you are proposing.

Note: you don’t simply ‘guess’.

It is not a random guess rather an ‘**educated guess**’ based on what you already know/learnt about a phenomena.

In order to write a robust hypothesis, you must understand what the **variables** are for your project.

Reminder:

hypothesis is a testable statement &
the variables in it must be testable.

Hypothesis:

IF machine learning is used,

THEN the performance can be better predicted.

It seems like an obvious statement, isn't it?

The above hypothesis is too simplistic for a post-graduate project!

You should find some problem for which the answer is not obvious or already known.

The hypothesis must be:

- based on your ‘educated guess’ **not on known data** or facts;
- written **before starting with experimental procedures** and not after.

Tips for writing the research hypothesis

The research hypothesis:

- comes **after the research question**
- is **NOT a question** but a statement
- must be **clear**, in simple language & **precise**
- is followed by one **experiment** & not many
- must be **TESTABLE**

Testable hypothesis

Testability requires measurable variables:

to **accept/reject** a research hypothesis,
an **experiment** needs to be planned/executed &
measurements or observations made
to check how planned variables are related

A hypothesis is testable if there is a possibility of deciding whether it is true or false based on experimentation **by anyone**.

Leading to Replicability & reproducibility

A good research hypothesis leads to:

- **replicability:** the experiment must be replicable multiple times and by other reseachers
- **reproducibility:** the scientific findings obtained in the experiment must be reproducible by other researchers

Null hypothesis

The null hypothesis The word ‘null’ means that it is a commonly accepted fact that researchers work to ‘nullify’. This does not mean the statement is null itself (null $\sim$ nullifiable) (H_0) is the **commonly accepted fact**.

Researchers and scientists work to **reject, disprove or nullify** the null hypothesis.

Null hypothesis

The null hypothesis [1] (H_0) is the **commonly accepted fact**.

Researchers and scientists work to **reject, disprove or nullify** the null hypothesis.

[1] The word 'null' means that it is a commonly accepted fact that researchers work to 'nullify'. This does not mean the statement is null itself (null $\sim$ nullifiable)

Alternate hypothesis

An **alternate hypothesis** (H_1) is a complementary (often inverse) statement to the (null) hypothesis.

Researchers come up with an alternate hypothesis, (one thought to explain a phenomenon), and then work to reject the null hypothesis.

Why testing the Null hypothesis?

It's an essential component of the **scientific method**

Philosophical disclaimer: approach to investigation, combining thinking and doing

Why testing the Null hypothesis?

It's an essential component of the **scientific method** [2]

[2] Philosophical disclaimer: approach to investigation, combining thinking and doing

Why not just prove the alternate hypothesis?

To guarantee **new hypotheses have no flaws**.

Clearly stating both null & alternate hypotheses is safer & it ensures the research is **not flawed**.Not including the null hypothesis is considered extremely bad practice by the scientific community

If the aim is to prove the alternate hypothesis (without considering the null hypothesis), we are set for failure & the experiment **won't be taken seriously**.

Why not just prove the alternate hypothesis?

To guarantee **new hypotheses have no flaws**.

Clearly stating both null & alternate hypotheses is safer & it ensures the research is **not flawed**.Not including the null hypothesis is considered extremely bad practice by the scientific community

If the aim is to prove the alternate hypothesis (without considering the null hypothesis), we are set for failure & the experiment **won't be taken seriously**.

Why not just prove the alternate hypothesis?

To guarantee **new hypotheses have no flaws**.

Clearly stating both null & alternate hypotheses is safer & it ensures the research is **not flawed**.[3]

If the aim is to prove the alternate hypothesis (without considering the null hypothesis), we are set for failure & the experiment **won't be taken seriously**.

[3] Not including the null hypothesis is considered extremely bad practice by the scientific community

Example: the flat heart society

Argument: Not so long ago (even nowadays), people believe/ed that the world is/was flat.

Null hypothesis

H0: The world is flat.

Alternate hypothesis H1:

IF photographic evidence is collected via satellite technology THEN the shape of the world is expected to be round.

Example: the flat heart society

Several scientists (e.g.. Copernicus) set out to disprove the null hypothesis.

Result: rejection of the null & acceptance of the alternate hypothesis.

Most people accepted it. Those who did/do not formed the Flat Earth Society `https://en.wikipedia.org/wiki/Modern_flat_Earth_societies`.

Example: the flat heart society

Several scientists (e.g.. Copernicus) set out to disprove the null hypothesis.

Result: rejection of the null & acceptance of the alternate hypothesis.

Most people accepted it. Those who did/do not formed the Flat Earth Society `https://en.wikipedia.org/wiki/Modern_flat_Earth_societies`.

Example: the flat heart society

Several scientists (e.g.. Copernicus) set out to disprove the null hypothesis.

Result: rejection of the null & acceptance of the alternate hypothesis.

Most people accepted it. Those who did/do not formed the Flat Earth Society `https://en.wikipedia.org/wiki/Modern_flat_Earth_societies`.

Example: the flat heart society

Several scientists (e.g.. Copernicus) set out to disprove the null hypothesis.

Result: rejection of the null & acceptance of the alternate hypothesis.

Most people accepted it. Those who did/do not formed the Flat Earth Society [4].

[4] https://en.wikipedia.org/wiki/Modern_flat_Earth_societies

What would have happened if Copernicus had not disproved the null hypothesis and merely proved the alternate?

No one would have listened to him!

In order to change the way people think,

we first must prove that their thinking is wrong.

The hypothesis is usually hidden in a word problem, and it is a statement of what you expect to happen in the experiment.

Educated guess: a researcher believes that if Support Vector Machine (SVM) is applied to classify emails into ham/spam then an accuracy of 85% or more can be achieved 85% because literature suggested this was the highest accuracy obtained that far.

The hypothesis is usually hidden in a word problem, and it is a statement of what you expect to happen in the experiment.

Educated guess: a researcher believes that if Support Vector Machine (SVM) is applied to classify emails into ham/spam then an accuracy of 85% or more can be achieved 85% because literature suggested this was the highest accuracy obtained that far.

Hypothesis formation

The hypothesis is usually hidden in a word problem, and it is a statement of what you expect to happen in the experiment.

Educated guess: a researcher believes that if Support Vector Machine (SVM) is applied to classify emails into ham/spam then an accuracy of 85% or more can be achieved [5].

[5] 85% because literature suggested this was the highest accuracy obtained that far

Convert hypothesis into math

Null hypothesis (H_0):

accuracy of ham/spam classifiers $<=$ 85%

Alternate hypothesis (H_1):

IF a ham/spam emails classifier is trained with SVM

THEN its accuracy $>$ 85%

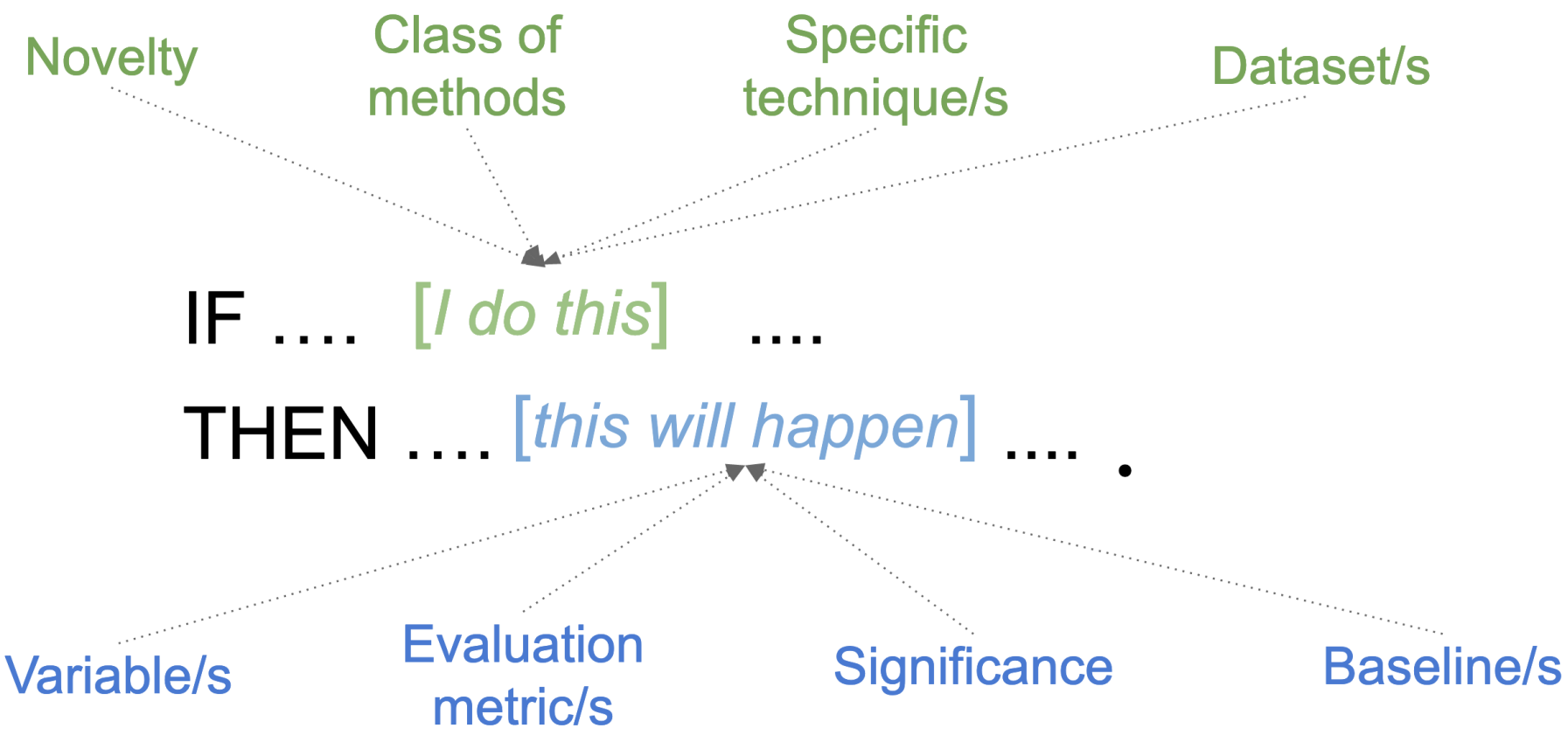
What if we do not know what will happen?
State what will happen if the experiment will not make any difference.

Alternate hypothesis (H_1):

IF a ham/spam classifier is trained with SVM
THEN its accuracy $\neq$ 85%

Live example

Hints for hypothesis formation



Give the reader all the elements to design an experiment alone!

Suggested readings

- Creswell, J. 1994 **Research Design: Qualitative and Quantitative Approaches**. (Sage)
- Leary, 2014 **Introduction to behavioural research methods** (Pearson)
- Marder, M.P. 2011 **Research methods for science**. (Cambridge university press)