UNDERSTANDING SCIENTIFIC INVESTIGATION

Victor Lau

The Chinese University of Hong Kong

The methods scientists use to investigate a problem are collectively called **scientific methods**. But scientific methods are not specific methods; they are a variety of methods that can make **evidence-based conclusions**. **Experiments** are one of the commonly used scientific methods. In the following, the basic design of an experiment will be explained with an example about fertilizer and plant growth.

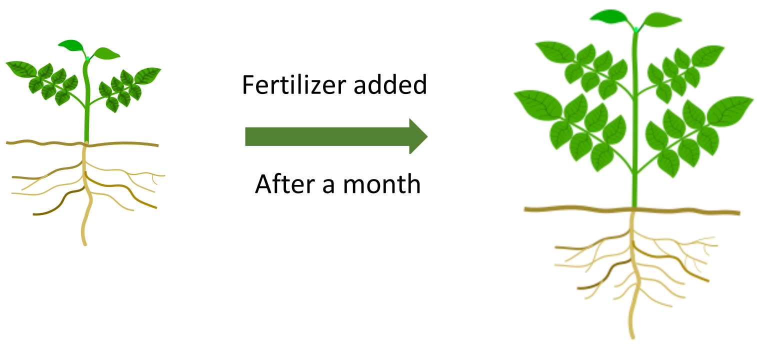
Scientific investigations start from asking a **clear question**.

Question

Does the addition of fertilizer into soil promote the growth of plants?

An experiment is done to collect **evidence** to answer the question.

Experiment



A reasonable conclusion is made from the results of the experiment.

Conclusion

The addition of fertilizer in soil can promote the growth of plants.

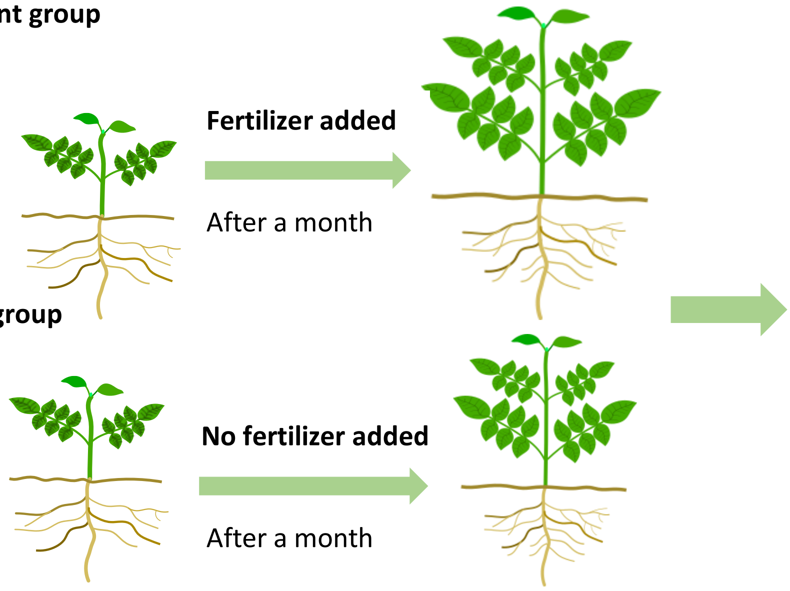
Critical thinking

The plant will grow even without the addition of fertilizer. How can you be sure the growth of the plant is caused by the addition of the fertilizer?

(A) A controlled experiment

Treatment group

Control group



Results
The plant added with fertilizer grows much larger than the plant without fertilizer added.

In this improved experiment, a **control** is added. A control is the **replication of the experiment without the treatment** – addition of fertilizer. A control is important by providing a **reference for comparison** with the treatment so that we can know the effect of the fertilizer.

Conclusion

Since the plant added with fertilizer grows faster than the plant without fertilizer added, we are confident that the fertilizer can promote plant growth.

> **Critical thinking**
> The plant in the control group may grow slower due to poorer soil, less water, less sunlight, etc. How can you be sure that it is caused by the lack of fertilizer?

(B) Independent, dependent and control variables

Between the treatment and control group, there could be many **differences** that may affect plant growth, such as the plants, the nutrient and water contents in soil, the amounts of sunlight, etc. But we are only interested in how fertilizer affects the growth of plants. Therefore, we must keep **all other conditions between the two groups the same**. These conditions are called **control variables**. Addition of fertilizer or not is the **independent variable** that we change it intentionally to see its effect on the **dependent variable**, the plant growth. If we have 'controlled' all variables except the independent variable, we can be confident that it is the independent variable that **causes** the change in the dependent variable.

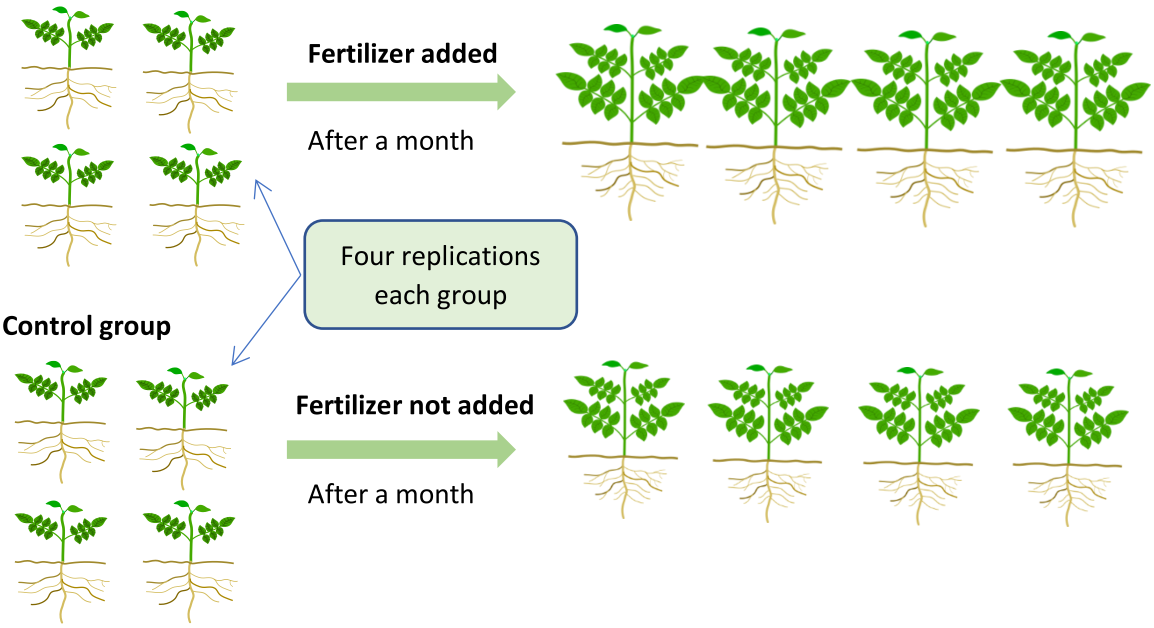
Three types of variables in a controlled experiment

	Treatment group	Control group
Independent variable (change intentionally)	Fertilizer added	No fertilizer added
Dependent variable (the results to look at)	Faster growth	Slower growth
Control variables (keep the same between groups)	Types and size of the plants, soil nutrients, soil water, sunlight, etc	

Critical thinking
We can choose the same type of plants in the same size, but they are **two individual plants** with different **genes**. The plant in the control group may grow slower due to its genes but not because of no fertilizer. Moreover, there must be small differences between the two groups that cannot be fully 'controlled'.

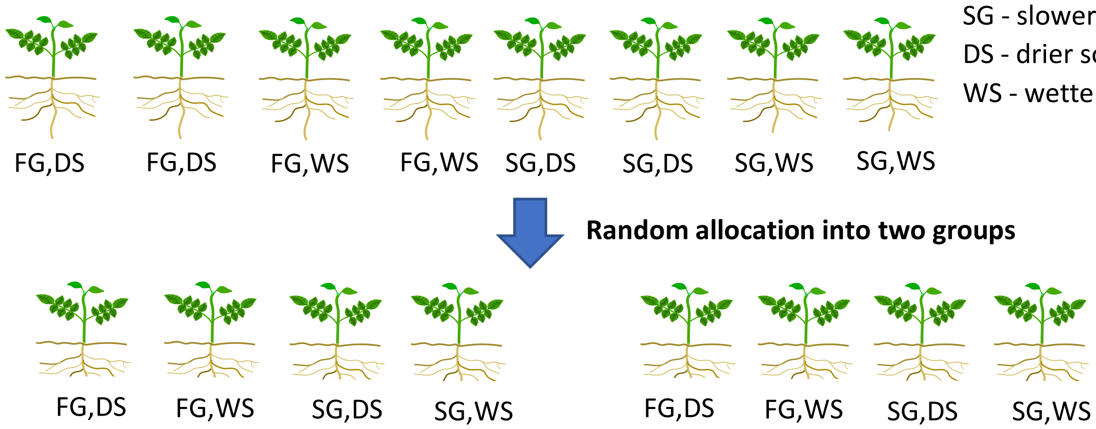
(C) A controlled, replicated experiment

Treatment group



The experiment is **replicated** four times in each group. If the eight plants are **randomly allocated** into the treatment or control group, it is likely that each group will have some plants that grow faster naturally and some that grow slower naturally due to their genes. With enough replications and random allocation, the '**averages'** of the plants and the conditions between the two groups would have no big differences.

FG - faster growth plant
SG - slower growth plant
DS - drier soil
WS - wetter soil

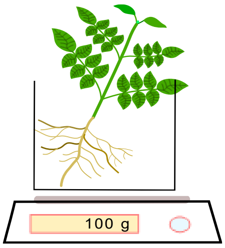


Treatment group **Control group**

(D) Measurement

When there are many plants in each group, it is difficult to tell the difference between the groups by their appearance. Moreover, scientists want to know **to what extent** the fertilizer promotes growth, rather than a **'yes' or 'no'** answer. To do that, we need to **measure** the mass of the plants in each group and obtain the **data**.

Measurement needs a **measurement instrument**, such as an electronic balance. Different instruments have different **limitations and uncertainties**. We need to **choose the right instrument** and **use it correctly**. We better use **more than one instruments** so that we know if anyone is defective. Also, we should make **repeated measurements and get the averaged values** to reduce the **random errors** during measurement.



Measure the mass of a plant with an electronic balance

- Do you choose the right instrument e.g., balance vs electronic balance?
- Do you use it correctly e.g., how to set zero?
- Do you use one more balance?
- Do you measure the mass of each plant for several times and get the averaged value?

Apart from the accuracy, another challenge of measurement is **what to measure.** What is the best measure that can show the growth of a plant? A **dry mass** is always preferred over **wet mass** since water content is not considered growth. Should we weight the whole plant, the shoot, or the root? How about counting the number of leaves? Sometimes there is no definite answer, but scientists need to clearly state their methods of measurement in an experiment.

(E) Data presentation and analysis

The data obtained from an experiment need to be **recorded accurately.** The data are then **analysed** and **presented** in tables and graphs.

A table recording the raw data

Dry mass of the plant, measured for three times (g)				
	Treatment group (fertilizer added)		**Control group** (no fertilizer added)	
Plant	At the start	After one month	At the start	After one month
1	10, 10.1, 10	55.1, 55, 54.9		
2	11, 11, 11.1	52.9, 53.1, 53		
3	13, 13, 13	57, 57, 56.9		
4	7, 7.1, 7	58.9, 58.9, 59		
5			9, 9, 9	33.1, 33, 33
6			11.1, 11, 10.9	30, 29.9, 30
7			11, 11, 11	35, 35.1, 35
8			8, 8.2, 8.1	34, 34, 34.1

The raw data above need to be processed further for analysis. First, the three values of repeated measurements are **averaged**. The **increase of dry mass** of each plant in the month is calculated to show their growth. Then, the **average increase of dry mass** in each group is calculated. After the data processing, a clearer picture of the data is shown to answer the question of the investigation.

A table presenting the processed data

	Dry mass of the plant - average of three measurements (g)					
	Treatment group (fertilizer added)			Control group (no fertilizer added)		
Plant	At the start	After one month	Change	At the start	After one month	Change
1	10.0	55.0	+45.0			
2	11.0	53.0	+42.0			
3	13.0	57.0	+44.0			
4	7.0	58.9	+51.9			
5				9.0	33.0	+24.0
6				11.0	30.0	+19.0
7				11.0	35.0	+24.0
8				8.1	34.0	+26.0
Average of the plants in a group			+45.8			+23.2

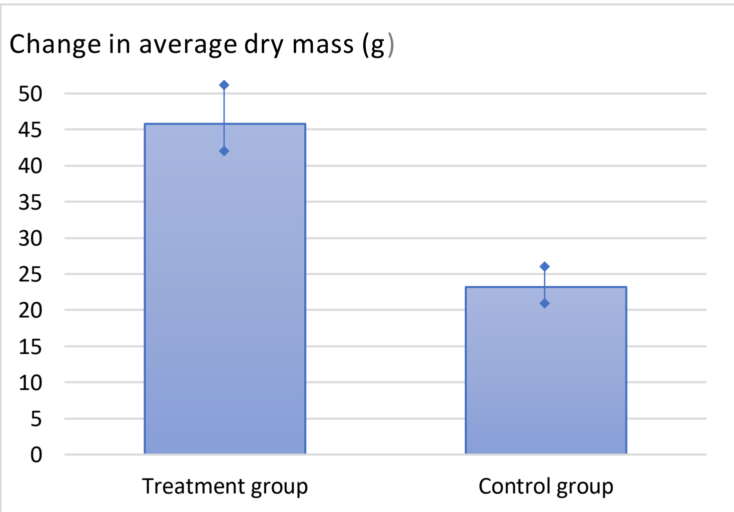
To make the data clearer for analysis, we may show important data only in the report. In such case, it is better to provide more information about the processed data, such as the **range** and **standard deviation**.

A table showing only the important data –changes in average dry mass of the two groups

	Treatment group (fertilizer added, n=4)	Control group (no fertilizer added, n=4)
Change in average dry mass (g)	+45.8 (42-52)	+23.2 (19-26)

n-number of individuals in a group; (…) range

A **bar chart** with the ranges may also be drawn to show the data visually for easier comparison.



(F) Data interpretation and drawing conclusions

A conclusion is an answer to the research question based on the **evidence**. Evidence in science refers to the experimental results or other observations. However, a conclusion is NOT an interpretation of the results by theories.

Conclusion drawn from the results
Addition of the fertilizer can promote plant growth.

Interpretation of the results
The fertilizer promotes plant growth because it provides nitrogen for protein production.

To draw a **valid conclusion**, the design of the experiment and the data must be examined critically.

Experimental design

Is it a **'fair' test**? Or does the experiment have any **uncontrolled variables** that may affect the results seriously? Below are some questions that may need to address:

- Are four replications in each group enough to control the differences of individual plants?
- Would the plants in the control group get less sunlight?
- Would the plants in the control group be affected by pests?
- Is the growth period of one month enough? Maybe the plants in the control group grew slowly at the beginning but would grow as well as the treatment group after six months.

Sometimes the best design may not be the one with everything under tight control in **the laboratory.** A **field study** may be better by showing the effects in the real environment. If the fertilizer experiment is done in the laboratory or greenhouse, the results may be different in the wild. So it all depends on the research question to answer.

Measurement errors

- Are the data accurately showing the dry mass of the plants?
- Is the electronic balance a **reliable** and **accurate** instrument?
- Did the scientists use the balance correctly?

The repeated measurements of the dry mass of plants did not make big differences in the results. It shows that the measurement is reliable. An electronic balance is an accurate instrument and its use is very simple and direct. There seems no reasons to worry about the measurement errors of the data.

Differences between groups

The plants in the treatment group had a much larger increase in dry mass than the plants in the control group: 45.8g compared to 23.2g. One can easily draw the conclusion that the fertilizer can promote plant growth. But sometimes data are not always such obvious. Look at the data below:

	Treatment group	Control group
Change in average dry mass (g)	+45.8 (42-52)	+42.2 (36-49)

45.8 and 42.2 are close and have overlapping ranges. They may have **no 'real' difference**. The small difference may be **by chance** because of the differences of the plants allocated to the two groups, some uncontrolled variables, some measurement errors, etc. In such case, a **statistical test,** such as ***t* test,** is needed to calculate if there is real **statistical difference** between the two groups. The test will look into all data and generate a ***p*-value** to show the **chance** that the difference is NOT real**.** ***p* < 0.05** means that there is **less than 5% chance** that the difference is not real and just by chance. Or alternatively, we have **95% confidence** that the difference is real. p <0.05 is commonly accepted to be a **statistically significant difference** between two groups. But some research may have higher requirement for a good p value, such as p <0.00001 in astrophysics.

Specific or general conclusion

Look at the following two conclusions:

1. (ALL) Fertilizers can promote the growth of plants (ALL)
2. The fertilizer used in this experiment can promote the growth of the plants used in the experiment.

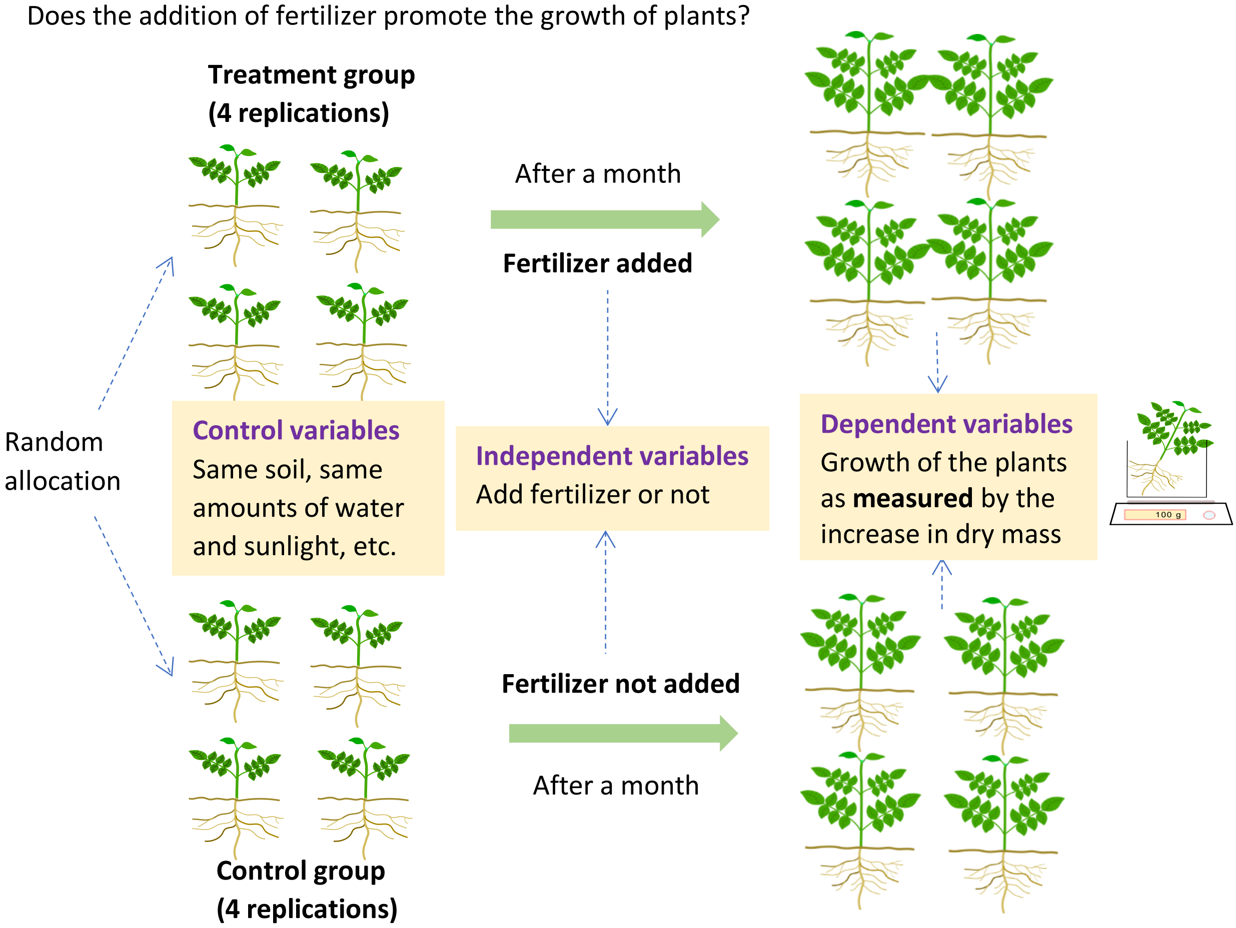
Conclusion 1 is applied to all fertilizers and plants in general, but conclusion 2 refers only to the specific fertilizer and plants used in the experiment. **Science is aimed at making general conclusions that can be applied to all instances**. To do that, many experiments using different fertilizers and plants should be done. For a single experiment like this, we can only make a specific conclusion.

Integration of learning

A replicated, controlled experiment

Question

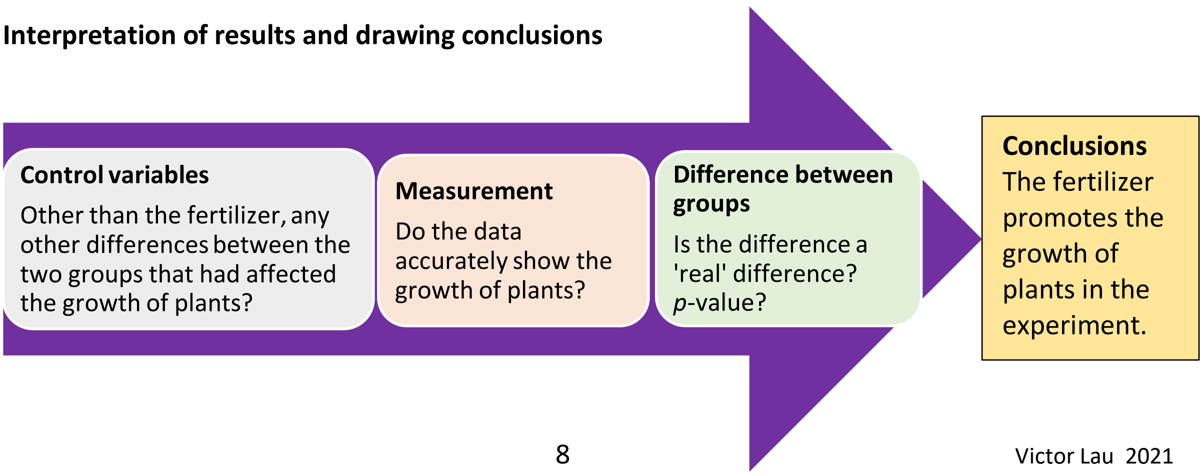
Does the addition of fertilizer promote the growth of plants?



Results

	Treatment group (fertilizer added)	Control group (no fertilizer added)
Change in average dry mass (g)	+45.8 (42-52)	+23.2 (19-26)

Interpretation of results and drawing conclusions



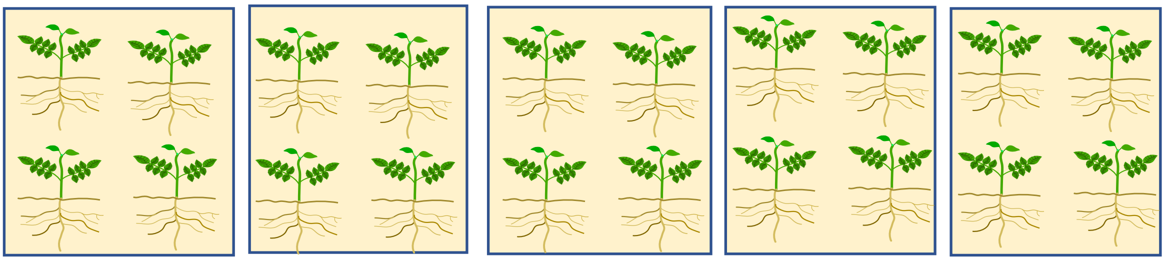
(G) A more complex design – different amounts of fertilizer

So far we know that fertilizer can promote plant growth. But we do not know **how much fertilizer** should be added into soil to produce the best plant growth. It may not be the more fertilizer the better. To find it out, we need to change the question and the experimental design.

Question – How is the plant growth affected by different amounts of fertilizer?

Independent variable – different amounts of the fertilizer added

0.2g 0.4g 0.6g 0.8g 1g



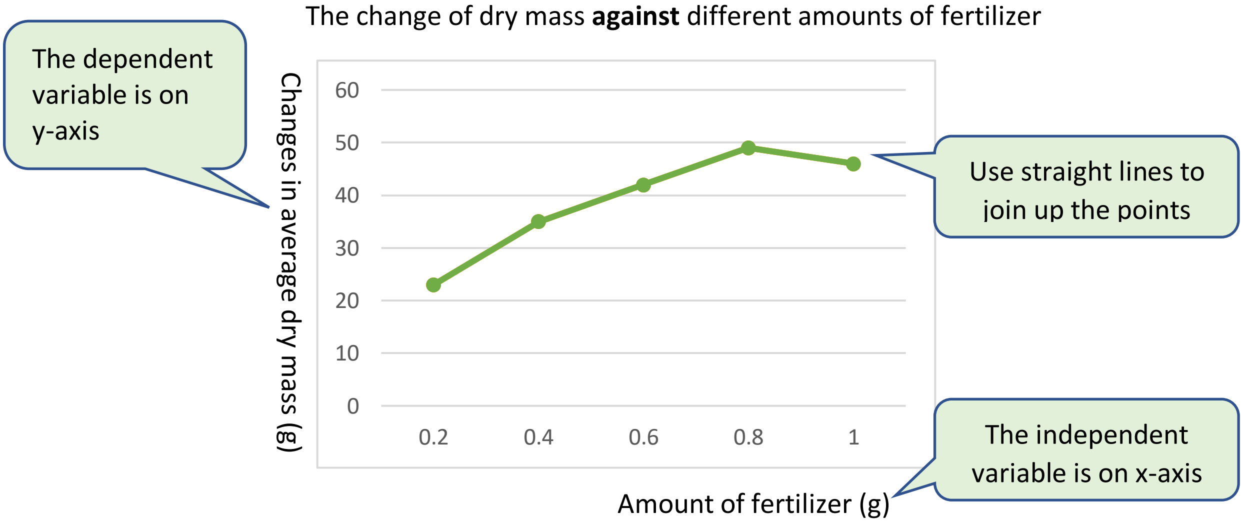
Five groups treated with different amounts of fertilizer. Each group has four replications.

Results

Amounts of fertilizer added (g)	0.2	0.4	0.6	0.8	1
Change in average dry mass (g)	+23	+35	+42	+49	+46

When both the dependent and independent variables are **numbers**, a **line graph** is needed to show their relationship.

The change of dry mass **against** different amounts of fertilizer



Interpretation of results and drawing conclusions

The dry mass increased more when more fertilizer was added. But the effect reached a **maximum at 0.8g**. The results show that addition of more fertilizer would promote more growth, but the **optimum amount** of fertilizer is 0.8 g.

We need to be careful when making **inference beyond the data** – what happens when fertilizer is more than 1 g or less than 0.2 g. It is justified when the data is having a clear trend and the theory supports it.

Control group?
There is no need to set a control group in this experiment because we are already comparing the groups treated with different amounts of fertilizer. But any **variables** between the groups still need to control, such as soil, sunlight and water added, and the plants that are randomly allocated.

(H) Investigations that seek to explain a phenomenon

- **Why** is too much fertilizer harmful to plants?
- **How** does the fertilizer work to promote plant growth?

These questions cannot be investigated without first making a **guess** of the answer. This kind of guess that gives a direction to a scientific investigation is called a **hypothesis**. A hypothesis needs to give a **prediction** that can be tested by an experiment.

Question: Why is too much fertilizer harmful to plants?

↓

Hypothesis: Too much fertilizer is harmful to plants because it makes the soil too acidic.

↓

Prediction: The soil added with more fertilizer will be more acidic.

↓

Experiment: Test the acidity of soil added with varying amounts of fertilizer.

↓

Results: The more fertilizer was added, the more acidic was the soil.

↓

Conclusions: The results **support** the hypothesis that too much fertilizer makes the soil too acidic and is thus harmful for plant growth.

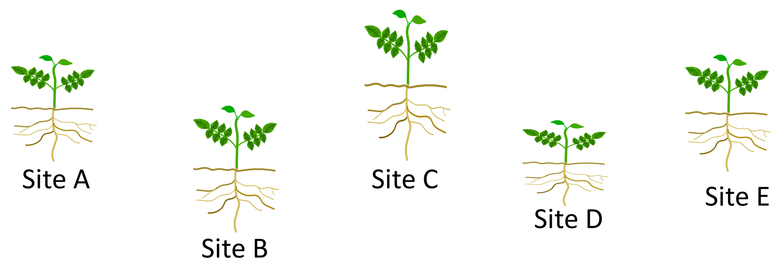
In the above reasoning, the **prediction is tested directly** by the experiment, while the **hypothesis is only tested indirectly** through the prediction. When the prediction is found incorrect, the hypothesis is **refuted**. When the prediction is correct, the hypothesis is **'supported'**, but not proved correct. It is because there could be many **other hypotheses** that produce the same prediction. Only after **many predictions** of a hypothesis have been proved correct, and other alternative hypotheses have been refuted, we can have higher **confidence** that a hypothesis is correct. Then, we may call this hypothesis a **scientific theory**.

(I) Non-experimental, correlational study

Not all scientific investigations involve doing experiments. Careful and systematic **observations** alone can also lead to big discoveries, such as Darwin's theory of evolution. Observations are important methods in the areas where it is not easy to repeat the phenomena by experiments, such as ecology, geology and astronomy. One commonly used non-experimental method is **correlational study**. The fertilizer experiment can be replaced by a correlational study as below.

To answer if fertilizer can promote plant growth, we can test if plant growth is correlated with soil nutrients.

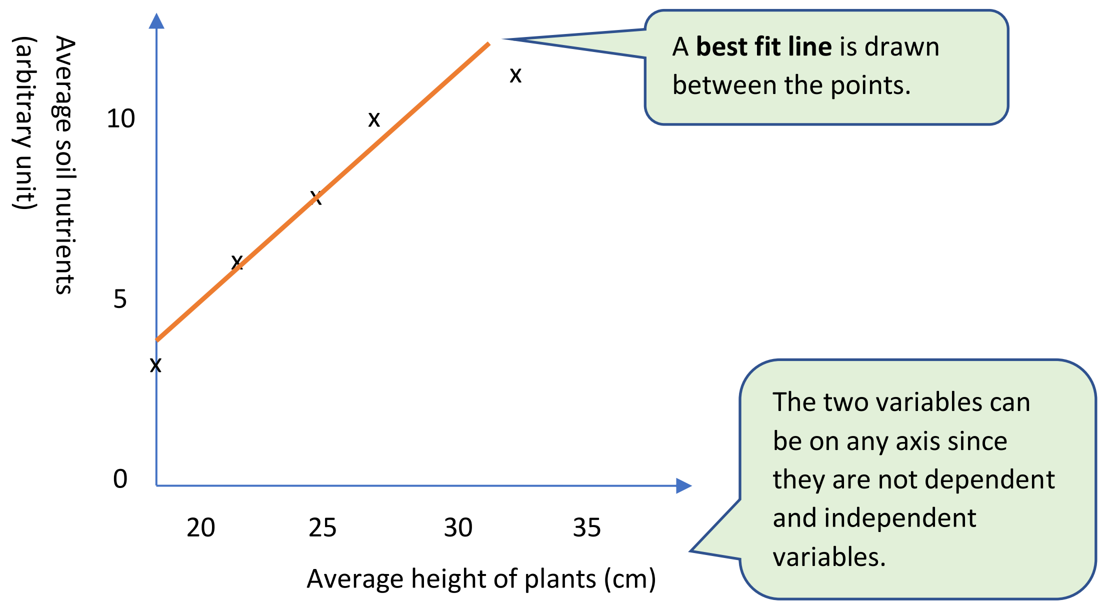
The soil nutrients and heights of the plants in **five sites of a forest** are measured.



Site	A	B	C	D	E
Average soil nutrients (arbitrary unit)	8	6	12	3	10
Average height of plants (cm)	25	23	33	18	27

Results

The two variables are plotted in a graph to find their relationship.



Data interpretation and drawing conclusions

A straight **best fit line** with positive slope can be drawn among the points. It shows that the two variables are having a **positive linear relationship**. That is, the more soil nutrients a site has, the taller are the plants there. The correlation can also be shown by the **p-value**. If $p < 0.05$, the correlation is unlikely a result of chance.

However, we cannot be sure which is the cause and which is the effect in the relationship. The cause and effect can be the reversed: the taller are the plants in a site, the more soil nutrients are there. This also makes sense – the larger the plants, the more leaves will fall to ground and enrich the soil. A correlational study **cannot tell the cause and effect** directly like in an experiment. It is because an experiment is changing a condition to see its effect, but a correlational study is only measuring two or more **naturally existing conditions** and try to find out their relationships mathematically. A correlation only forms part of the evidence for a cause and effect relationship.

Besides, many variables cannot be controlled in a correlational study. The five sites are different from each other in many aspects apart from soil nutrients, such as sunlight, rainfall, other plants and animals, etc. But we can control them better in the laboratory in an experiment.

	Experiment	Correlational study
Variables	• An independent variable is artificially changed to see its effects on the dependent variable • Other variables are controlled	• Two (or more) variables that exist naturally • Other variables are largely not controlled
Conclusion	A **cause and effect relationship** between independent variable (cause) and dependent variable (effect)	A **correlational relationship** (positive or negative) between two or more variables
Where it is done	Mostly in the laboratory	Mostly in the field

-END -