Chapter 2
The Scientific Method

Overview

Ever believed something to be true, because everyone else says it is? That's a phenomenon known as the bandwagon effect, and is one of many **cognitive biases** that we humans are constantly under the influence of. Cognitive biases are very common and may be a result of personal, social or cultural beliefs. They distort how we interpret an observation or fact, and they can affect our rationality and logic in decision-making.

For example, people tend to listen more to information that confirms the beliefs they already have. The truth is we are more likely to accept evidence that confirms our beliefs and to reject evidence that contradicts them. This kind of cognitive bias, known as confirmation bias, is a prominent feature of many of today's societal debates, including climate change.

The **scientific method** is a way to cut through the noise and minimize the influence of cognitive bias when we try to make sense of the world around us. It is an empirical method of acquiring knowledge, involving careful observation and applying rigorous scepticism about what is observed and how it is interpreted.

It is not infallible, but it is the closest system we humans have to extracting and understanding the truth about the world around us in an objective, neutral way. In a world driven by social-media opinion and 'fake news', well executed science has never been more important.

This chapter explores the principles of the scientific method, and explains how you can design, conduct and write-up a research project that will pass peer-review and have the maximum impact to your readers.

Contents

2.1

Introduction

Introduction to the Scientific Method

In any field of research, scientists are faced with the need to compromise certainty.

With most research questions, there is the ‘ideal scenario’ to get the perfect data-set, but in reality this is almost always logistically and financially impossible.

A certain amount of certainty will almost always need to be compromised to account for the ‘cost’ and feasibility associated with a project. The amount of compromise that is acceptable- or necessary- will depend on what the data will be used for, what resources are available, and what will be permitted or acceptable (i.e. by ethics committees, national legislation and international norms).

For example, we might ideally want to put GPS collars on all 5000 elephants in a national park to study seasonal movements, but:

(a) it would be prohibitively expensive and take too long;

(b) park authorities are unlikely to approve such invasive sampling, and;

(c) it is probably not necessary to know what every single individual in a population is doing, in order to make reasonable assumptions about that population.

Introduction to the Scientific Method

This may seem like common sense, but this is the essence of the scientific method.

The scientific method helps us **ask a question**, as objectively as possible, on a small, and manageable part of the world, and to think carefully about the **implications** of the answer at a **wider scale**.

As we are forced to work within limitations, the scientific method helps us to take a representative sample, and extrapolate, with as little error as possible.

Introduction to the Scientific Method

Let’s say you have been awarded a grant of $20,000 for 1 year to measure carbon stocks in a 10,000 ha community forest reserve. There are a range of methodological options available to you.

Which method is the best one and how do you choose?

Method 1	Cut several hundred trees in the concession, and measure the carbon directly with an automated carbon analyser.
Method 2	Measure the diameter and height of each tree in the concession, and estimate carbon stocks with equations derived for other species.
Method 3	Measure the diameter and height of a subset of trees in the concession, estimate the carbon with the equations developed for other species, and extrapolate for the entire concession.
Method 4	Use Radar images to estimate carbon stocks based on the amount of canopy visible from satellites.
Method 5	Extrapolate carbon stocks from data published in a study conducted in a nearby forestry concession.

Introduction to the Scientific Method

The options range from (1) intensive and costly to (5) easy and cheap. You can start by narrowing down your choices based on practical feasibility. Method 1 will be labour intensive, expensive and challenging to obtain permission for, due to its destructive nature. Method 2 is logistically unfeasible: at a tree density of 250 stems/ha there will be about 2.5 million trees in the concession – it would take years to measure every one of them! Method 5 is also a last resort, as with a project budget and one year you are able to collect your own data specific to the concession and improve on accuracy.

Therefore, within the limitations of your project you could reasonably consider Methods 3 or 4.

Method 1	Cut several hundred trees in the concession, and measure the carbon directly with an automated carbon analyser.
Method 2	Measure the diameter and height of each tree in the concession, and estimate carbon stocks with equations derived for other species.
Method 3	Measure the diameter and height of a subset of trees in the concession, estimate the carbon with the equations developed for other species, and extrapolate for the entire concession.
Method 4	Use Radar images to estimate carbon stocks based on the amount of canopy visible from satellites.
Method 5	Extrapolate carbon stocks from data published in a study conducted in a nearby forestry concession.

Introduction to the Scientific Method

As we hone in on the best method for our study with the constraints we have, the scientific method guides us in this process, to ensure that:

- our sample is representative of the true population we want to measure
- our extrapolations incur as little error as possible.

More specifically, the scientific method guides us to:

- Formulate research questions and hypotheses.
- Choose the most practical, useful, and objective means to obtain a representative sample of the population, to answer the research questions that have been formulated. This is *sampling design*.
- Choose the best tools (methods) to collect the data to answer the research questions.
- Analyse the data with the most appropriate statistical tests.
- Extrapolate, as carefully as possible, on a larger scale, to interpret the results. This is usually presented in the discussion of a report or publication.

The Scientific Cycle

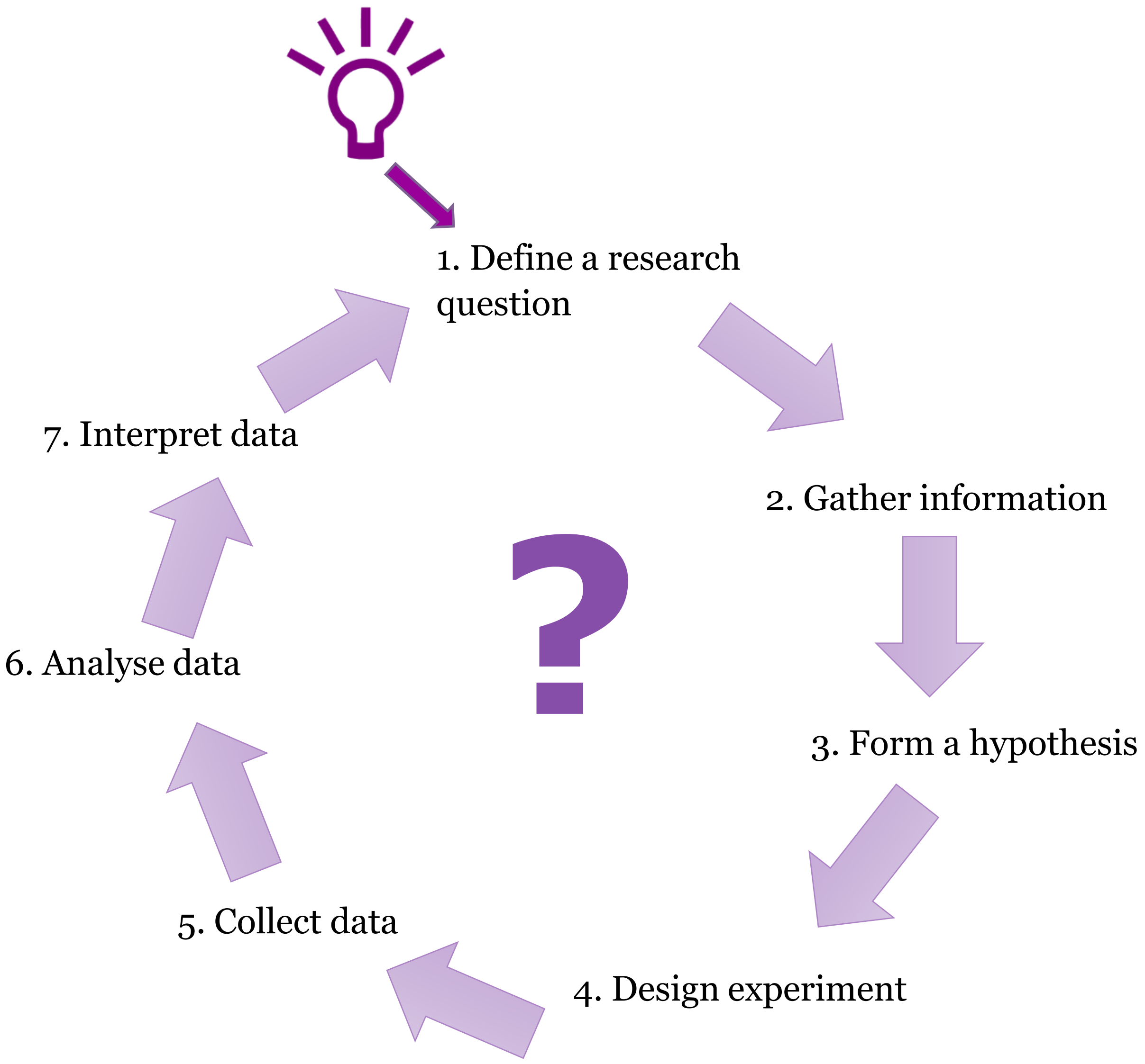
The scientific method is more a cycle than a series of steps.

The cycle always begins with an idea for the study. This idea gives rise to a research question, for which a conceptual design of the study is formulated.

This conceptual design allows the establishment of research hypotheses, for which a sampling plan is formulated.

The sampling plan directs data collection, which is then analyzed with the most appropriate statistical tests.

The results are interpreted and the reflections raised will provide the ideas for studies in the future…which feed into the next cycle of scientific research.



2.2

Defining research questions

Defining Research Questions

Research questions start with good ideas. The best ideas follow the cycle because they come out directly from another cycle. As a result, they build on the findings of previous studies – including preliminary results that require further investigation, the interpretation of biological patterns revealed by existing data and the exploration of more complex, probing questions to gain deeper insights into biological processes and phenomena. New studies should aim to be original, complementary to existing work and contribute to a more profound understanding of the subject.

The wording of your research questions is essential to get right, in order to conduct a successful project. A good research question should have the following characteristics:

- simple
- measurable
- attainable
- relevant
- limited to a particular time and place.

Defining Research Questions

When formulating research questions, avoid “why” questions. These are reflective in nature and very hard to answer directly.

e.g. "Why are there fewer buffalo this year than in previous years? "

This is a difficult question to answer with a single research project because ‘Why’ is reflective and open, and cannot be answered with a straightforward "yes / no", with a number or with a measurable answer. In our example the possibilities could be numerous: e.g. increased hunting pressure, drought, loss of food resources, epidemic disease or increased predation pressure. Testing each one of these possibilities is likely to be complex and time consuming.

"Why ..? Is better kept for the end of the study, rather than the preliminary question. It is very well placed in the discussion of a scientific article, because it allows the consideration of all the possibilities and the formulation of new questions. The ‘why’ of one study will feed into the scientific cycle of a new one by provoking new research ideas.

Similarly, other value-based question words such as “ought” or “should” should be avoided.

Defining Research Questions

In order to formulate your research question, it helps to first define your population and your variables.

In ecology, a **population** refers to all the individuals of one species in a given area (e.g. a population of beavers, leatherback turtles, oak trees, bees etc.).

A **variable** is any factor, trait, or condition that can exist in differing amounts or types (e.g. age, mass, length, quantity, temperature, distance, depth, etc.).

An experiment usually has three kinds of variables: **independent, dependent**, and **controlled**. The independent variable is the one that is changed by the scientist.

By using different combinations of populations and variables, three different types of research questions can be created. These are known as **descriptive**, **comparative,** and **correlative**.

Descriptive questions

Descriptive questions involve describing and/or quantifying parts of a natural system. They are good for pilot studies, or preliminary investigations in areas where not much is known.

They generally include **one population** and **one distinctive variable**. Examples of descriptive research questions are:

- What is the tree alpha diversity of Ashdown Forest, UK?
- What is the Alpine marmot’s distribution range?
- How frequently do manatees breed?
- How many carnivore species use forest corridors?
- What is the home range size of the Eurasian beaver?

Comparative questions

Comparative questions involve collecting data on different populations or under different conditions (e.g., times of year, locations, habitat types), to make a comparison.

They usually include **two or more populations** and **one distinctive variable**. Examples of comparative research questions are:

- Is there a difference in body length between male and female tortoises?
- Is there a difference in alpha diversity of fungi that live at different forest altitudes?
- Are there more primates in unlogged forests compared to logged?
- Do carnivores prefer secondary forest or primary forest corridors?
- Are corridors used more frequently in the dry season or the rainy season?

Correlative questions

Correlative field investigations involve measuring or observing two variables and searching for a relationship between them for a distinctive population.

They usually include **one population** and **two distinctive variables**. Examples of correlative research questions are:

- What is the relationship between tail length and age in humpback whales?
- How does a spider's reproduction rate change with a change in season?

2.3

Formulating hypotheses

Gathering information

After your research question is defined, and before you formulate your hypotheses, you need to gather information – by reading, asking advice, or making preliminary observations.

This information will allow you to make an informed prediction and formulate your research hypothesis.

It is very important that research hypotheses are based on theoretical knowledge, so make sure you read all the relevant literature and talk to all the right people who have intimate knowledge of your species, population and study site.

Formulating hypotheses

A hypothesis is a way of making a prediction. It is formulated as an explanation for a phenomenon that we expect to observe, based on prior knowledge and observation. In that way it is often described as an 'educated guess'. A scientific hypothesis has to be something that can be supported or refuted through careful experimentation. Upon analysis of results, hypotheses are either accepted, rejected or modified, however a hypothesis can never be proven to be correct 100% of the time. Through many repeated experiments, a hypothesis may be generally accepted as true, but there may be some observation somewhere in the world that does not fit the hypothesis.

Strong hypotheses are often written as, "If A occurs, then B will occur" and are presented as statements, not questions. Good hypotheses also are clear and keep variables in mind, defining them in easy-to-measure terms. Key features of a good hypothesis are that it is testable, researchable and simple enough to be tested with one or two experiments.

For example:

Research Question: *"Are chimpanzees hunted more in logged or unlogged forest?"*

Hypothesis: *Hunting results in a lower density of chimpanzees in logged forest*

The null hypothesis (H_0)is: "There is no effect of hunting on the density of chimpanzees"

2.4

Sampling Design

Sampling design

Sampling is a critical part of any ecological study and relies on the basic premise that a small portion of a population is studied in order to understand the whole population.

It is very important to remember that the results of any ecological study are only as good as the sample upon which the study is based. If the sample is NOT representative of the population being studied, the inferences made will be meaningless.

But how can you ensure that you sample is representative? The answer partly depends on using a sampling design or protocol based on an understanding of the population being studied.

The key point to remember is that the sample should be large enough to represent the range and pattern of variability in the population. Quite often these assumptions are ignored.

Population definition

It is important to accurately define your population, to ensure your statistical analyses will be conducted correctly.

In statistics, the population refers to the total set of observations that can be made. For example, if we are studying the heights of all women, the population is the set of heights of all women in the world. If we are studying the basal area of giant sequoia trees in Yosemite National Park, the population is the set of basal area measurements from all giant sequoia trees in Yosemite National Park.

However, because there is very rarely enough time or money to gather information from the whole population, a representative sample (or subset) of that population is selected.

Note that in some studies, the population from which the sample is drawn may not be the same as the population about which information is desired. Examples might include a study about the effects of a particular drug conducted on rats, in order to gain a better insight into its effects on human health.

Sampling frame

A sampling frame is a complete list of all the key units (or items) in your population from which your sample will be drawn. Defining a sampling frame is important, as you can only apply your research findings to the population defined by the sampling frame. The sampling frame is where your sample is extracted from.

The difference between a population and a sampling frame is that the population is general and the frame is specific. For example, the population could be “Fish species in the River Lez.” The frame would name ALL of those species.

Ideally, the sampling frame should be identical to the population. For example, if your population is “All carnivores in Kruger National Park” but you take your samples only during the day, your sampling frame is “diurnal mammals in Kruger National Park”. Here, your sampling frame is different from the population.

Sample size

How do we choose a representative sample from your sampling frame? The larger your sample size is, the smaller the variance will be in your data, and the more accurate your results will be.

However, as we have seen, you always have to balance your sample size with the time and money available. Prior knowledge of the population helps, because by knowing the range and pattern of variability in the population, you can choose a sample that is large enough to represent the variability in that population. If the population is small, then it may be feasible to study all of its members.

When you only survey a small sample of the population, uncertainty creeps in to your statistics. If you can only survey a certain percentage of the true population, you can never be 100% sure that your statistics are a complete and accurate representation of the population. This uncertainty is called **sampling error** and is usually measured by a **confidence level**. For example, you might state that your results are at a 90% confidence level. That means if you were to repeat your survey over and over, 90% of the time your would get the same results.

There are different formulae you can use to determine sample size, depending on what you know (or don't know) about your population. Examples include Slovin's formula and Cochran's sample size formula. Many resources are available online.

Sampling units

Sampling units are also referred to as 'observational units' or 'experimental units' .

Choosing the right sampling unit is critical to the study- this is the unit that will be used to pick up data and represents the resolution of the study.

Sampling units must be clearly defined in terms of their spatial, temporal, and ecological/socio-economic attributes, and they must match the scale of the research question perfectly.

Sampling units might be natural: e.g. individuals, groups, houses in a village. Or they may be artificial: e.g. plots, transects, or quadrats.

Types of Sampling

A variety of sampling methods can be employed, and factors such as the need for accuracy, cost/logistical feasibility, and the nature of your sampling frame will all influence your choice of sampling method. It is advised that you seek expert advice when choosing your sampling method, and to discuss beforehand with statisticians: this may be a decision that involves considerable deliberation!

While not exhaustive, three common types of sampling in ecology include random sampling, systematic sampling and stratified sampling.

Random sampling

Random sampling is usually carried out over large, fairly uniform areas.

A numbered grid is placed over the study area and random numbers are chosen. Samples are taken within those grids.

Examples would include 0.25m^2 quadrat sampling in a meadow, 0.5 ha plot sampling in a national park, or 1 ha plot sampling across a whole landscape.

Systematic Sampling

Systematic sampling is often used when there is variation in the study area, or the sampling area is not very large.

Here samples are taken at fixed intervals, usually along a line. This method is useful for sampling across environmental gradients, e.g. for studying plant diversity along an altitudinal gradient, or moving from grassland to forest.

Line transects are often used for this sampling approach. Belt transects are widened lines and like long, extended plots- enabling sampling on abundance and density.

Stratified Sampling

Stratified sampling takes into account different strata within the population to be sampled, in order to avoid bias. Some important, but rare areas may not be picked up by the random sample and may be missed that you want to include. Alternatively, you may want to control for factors such as habitat type and seasonality. Stratification is normally carried out by an initial survey of the study area or analysis of any aerial photography, habitat maps or other data which are available.

For example, a researcher might estimate the densities of elephants in an area of forest using the following design:

60 line transects conducted between Jan and March (during the fruiting season), 20 each in recently logged forest, primary forest and secondary forest.

Here the population being sampled is divided into different strata, these being forest type and fruiting season.

2.5

Collecting Data

Data Collection

Proper data collection requires forethought, planning and patience. It is essential to the entire study that data is collected correctly. Incorrect data has no biological meaning and is completely useless. At the very least, it will mean a lot of wasted time and money. At worst, it will give false results, and could do more harm than good.

Imagine this scenario:

You are asked to conduct a study on lion densities, which you do, concluding that lion dung densities are significantly higher in one particular area (just outside a National Park) compared to a number of other sites. Based on your results, a conservation group decides to raise thousands of dollars and sets up a fund to try to protect lions in this previously unprotected area.

Money is spent on raising awareness in the local community, sending rangers to do patrols, and inviting tourists to visit the area and give more money for lion conservation. The group is so successful at publicising the conservation importance of this new 'lion hotspot' that it raises huge international interest.

The government sees the economic opportunities and increases the penalties for lion hunting, as well as putting pressure on the adjacent park to send more rangers to protect the lion rich area.

Tourists begin to flood into the country to visit these famous lions. At the beginning it is good for the country. However, after a while, tourists begin to complain that they never see the lions. They become disgruntled and rumours begin to spread that the government has exaggerated or lied about the lions, in order to increase revenue.

Data Collection

It becomes a political issue and soon tourists stop coming to the country: before long there are fewer tourists than ever before. The local community, who had come to depend on this new source of revenue, suddenly find themselves poorer than ever. They become very angry, and many start to hunt again. And because the lion-site is so well protected, they decide to hunt inside the National Park, where there are fewer patrols.

In fact, because patrols have decreased inside the park for some time, hunting had begun to slowly increase, without anyone noticing. When a monitoring study is finally carried out, it is discovered that elephant numbers have decreased by 40% inside the National Park, as a direct result of the flawed lion conservation project.

And when your original results are finally scrutinised, someone notices a small but significant error. The misinterpretation of a single decimal point on one of the field sheets meant that the lion dung density estimates appeared to be 10 x greater in one area (the lion-site) compared to all other areas.

This tiny error ended up having huge consequences.

Although this example seems extreme, you should never underestimate the consequences of erroneous data. Some human error is inevitable, but you must make every effort to ensure your data are correctly recorded. **Do not take shortcuts.**

Key Points for Collecting Data

1. Be Rigorous

Start your study with a short pilot study or practice run beforehand. This way you can find out if there are any problems with your sampling design or equipment and you can make modifications as necessary.

Make sure you take measurements with great care, consistently and always according to your protocol.

Do not allow anyone else to collect data for you unless you are 100% certain that that they will follow methods and record data in exactly the same way as you. If you work with field assistants, make sure you work together collecting the same data independently and compare your field sheets for discrepancies, before collecting data at different times.

Always record detailed information on field sheets. It is essential to follow your sampling design wherever possible but make detailed notes if you have to deviate. In addition to measurements, careful recording of meta-data are essential (but often lacking) – this will ensure your data can be referred to in the future.

Record the date in full (day, month, year) on every field sheet as well other crucial identifiers (e.g. site name, plot code, GPS locations) and the names of all data collectors and field assistants. Ensure your sample tubes are also labelled with the date, collector name, site name, sample number and other important identifiers.

Key Points for Collecting Data

2. Be Organised

Make lists and give yourself plenty of time to prepare your field kits and field sheets before you go into the field.

Check field equipment regularly.

Make sure you know *how* you are going to record your data.

Collect your data-sheets at the end of each day and organise them in labelled folders, storing them carefully in one place. Take photocopies or scans as soon as you can.

When you enter data into excel spreadsheets, make sure you save files using a clear file naming system, and keep them stored in clearly labelled folders. Have one ‘master’ copy of your raw data, and archive copies that you make- avoid having multiple active files of the same data.

Make regular backups on external hard drives.

Key Points for Collecting Data

3. Be Objective

Conduct your research with total impartiality. Avoid succumbing to the temptations of cognitive bias at all costs- particularly confirmation bias!

Remember that a scientist is looking for the ***truth***, whatever that may be. Negative data are just as important as positive data, sometimes more so.

It is tempting to try and find patterns that fit our expectations, but if you set out to ‘prove’ your hypothesis because it is more interesting, your data will be biased, inaccurate and may do more harm than good.

Key Points for Collecting Data

4. Show Integrity

Respect for the environment and the population you are studying is of utmost importance.

Create as little disturbance as possible to natural habitats when sampling, and always aim to use the least invasive or destructive protocol possible.

Be respectful of local communities and cultures, as well as other people working at the same site as you. Be mindful of other activities being conducted around you, and remember that while your study is of the utmost importance to you, it may not have the same priority to others.

Be aware of the impacts of your research locally- take guidance from trusted local partners and your hosts- your presence may have an impact on their visibility and reputation.

Key Points for Collecting Data

Remember: Be **ROOSI**!

Rigorous – **O**rganised – **O**bjective – **S**how **I**ntegrity

2.6

Analyses and interpretation

Cleaning and preparing your data

It is **essential** you examine your data thoroughly before analysing them. Data sets are more than likely to have errors from various sources. Common error sources include:

- faulty instrumentation (false or erroneous readings due to environmental anomalies or equipment malfunction)
- data logger 'artefacts' (e.g. during calibrations or equipment set-up)
- transcription errors (misreading field sheets, data entry errors from field sheets to excel)
- cut and paste mistakes (column/row/cell copy or sorting errors; watch out also for 1900 and 1904 date system differences in Excel : this can cause dates to change by 4 years and 1 day when copying and pasting between Excel files).

Data simply have to be cleaned and the best way to see if data are, in fact, clean is to look at them. Visualise the raw data by eye first of all, and look for any weird values, outliers or patterns.

Cleaning and preparing your data

Once you have cleaned your data, it is advisable that you prepare and examine a simple histogram for each variable under consideration for analysis.

Are there outliers or unexpected values? Is the distribution mound-shaped? Does the distribution have a different shape that makes sense? Is the distribution centered on what one intuitively expects to be the mean? Plots are a useful way to understand the structure of the data.

Many errors could be avoided through careful and thoughtful initial data visualization. Plot the data early and often.

Choosing statistical tests

There are many resources available to help you choose the most appropriate statistical tests for your study, and it is highly recommended that you complete a course on biostatistics. Discussion with supervisors, other experiences researchers and statisticians is also indispensable: if in doubt, reach out to others!

The statistical tests you choose must be directly linked to your research objectives and research questions. All statistical tests start with the formulation of a null hypothesis (H_0). This hypothesis states that there is no significant relationship or pattern within our data.

You will need to know if your data are **normally distributed** or not, if you can transform your data to conform to a normal distribution, and consequently, if you will apply **parametric** statistics or **non-parametric** statistics. The kind of data you have (interval or categorical, paired or unpaired) will also determine which statistical test you should use.

Due to the nature of ecological data, non-parametric statistics are fairly common in ecological studies.

Many software packages are available for statistical analyses, such as MINITAB and SPSS, however R is now the most widely used data analyses software used by ecologists. It is highly recommended you complete a course on R.

Intepreting your results

Interpretation of the results of statistical analysis relies on a basic question you need to answer, *do I or do I not have statistical significance*. This can be answered looking at one simple number: the *p* value.

Every statistical test that you report should relate directly to a hypothesis. But before you can determine if you have rejected or failed to reject your null hypothesis, you must designate the maximum probability of falsely rejecting the null hypothesis that you are willing to accept in your analysis. This is the significance level (also denoted as alpha) and is typically set at 0.05 in ecological research. A significance level of 0.05 means that you are willing to accept up to a 5% chance of rejecting the null hypothesis when the null hypothesis is actually true.

Once you conduct your analysis, you will get a *p* value. This number reflects the probability of obtaining results as extreme as what you obtained in your sample if the null hypothesis was true.

Statistical tests must always be reported with a p-value.

In reporting the results of statistical tests, you should report the descriptive statistics, such as means and standard deviations, as well as the test statistic, degrees of freedom, obtained value of the test, and the probability of the result occurring by chance (p value).

2.7

Write-up

Writing a scientific report

Whether you are writing up your research as a thesis, academic report or publication, there are a certain set of norms and standards you need to respect.

However, bear in mind that different institutions, organisations, journals and publishing houses all have different requirements- you will need to find out what particular format, structure and style is required by the bodies you are reporting to.

Doctoral theses are generally divided into several chapters, each one organised as a standalone scientific report: or more commonly nowadays, as individual publications.

Academic reports are usually written as a single document, but institutions may have their own reporting guidelines.

Scientific journals all have their own author guidelines and usually very specific requirements with strict word limits.

Structure

The first point to consider is the **structure** of your document. Most scientific research is written up in the following format:

- Abstract
- Introduction
- Methods
- Results
- Discussion
- Acknowledgements
- References

We will look at each of these sections in a bit more detail.

Abstract

The abstract is often considered the hardest part of a scientific report to write, and is best written last, once the rest of the document is finished. Usually limited to a few hundred words, it acts as a summary of your entire research document, and may be the only part other busy scientists may read. Therefore it needs to act as your research's publicity and contain the essential 'take-home' message from your study.

Abstracts need to follow the same logic as the main text, containing an introdution, aims and objectives, methods, results and conclusion. However, each section may only be one or 2 sentences long and so the whole abstract is usually presented as a single paragraph with no sub-headings.

Therefore, it is essential to be concise, and extract only the most essential pieces of information from each section of your main text.

Abstracts usually contain no references or citations, and have no figures or images in them.

Introduction

The introduction is the important lead-in to your research, and must thoroughly and adequately present the scientific background, theoretical concepts and state of current knowledge that is relevant to your study, allowing you to present and justify the research problem that your study will address. It provides the justification for your research and sets the scene, providing important contextual information to the reader to help them understand why you undertook this study and what it brings to the table to fill gaps in knowledge. It is the result of your literature review, and should be heavily referenced with the most recent and important publications relevant to your subject matter. Towards the end of the introduction, the aims and objectives of your research are usually introduced.

The introduction will vary in length depending on the kind of manuscript you are writing, but is usually a few pages long.

It does not usually have figure and diagrams in, and should **not** contain any data or results from your study.

Methods

Methods are the cooking instructions for your research ("we did this and did it this way"). They are often best written first, as you will already have all the necessary information to hand, even if you've not finished the analyses yet. The most important thing to remember when writing methods is that **any other researcher should be able to take them and replicate them**. Therefore, they need to be clear, explicit, and contain all the necessary information to allow your study to be replicated.

Methods can be difficult for people to write because as you are so involved in your own research you can easily forget that what is blindingly obvious to you may be completely foreign to someone else! Therefore, you need to explain all relevant details, including:

- Description of the study area, including geographic coordinates and main environmental characteristics, with justification for choice of study area
- the sampling units, sample size, study populations and sampling design chosen with justification for choice
- the kinds of equipment used to take measurements (model, make)
- the measurements that were taken, with SI units, and the frequency of data collection

Methods

Methods should also include a description of the analytical treatment of your data, including statistical tests conducted. A map of the study area with sampling locations is always a useful figure for the methods, or a table summarising the methodological design.

If you have used methods from another published study, you can simply reference the original study instead of re-writing them. However, if you adapted them in any way, this should be explained.

Methods are factual therefore they should avoid going into too much conceptual detail or theory (e.g. about the study species), which should be saved for the introduction.

Results

The results section is exactly this: it presents the results of your research and your statistical analyses- no more, no less. However, it is surprising how challenging this section can be to write.

First of all, you will likely have a lot of data to present, so the first challenge is to decide what results are relevant to your research objectives and questions, and what format they should be presented in.

Raw data is almost never presented in the results, if necessary it is added to an appendix or supplemenentary materials section. Summary statistics may be presented in table form or in the text, and the results of analyses are usually presented in graphical form, often as boxplots, histograms, scatterplots or line graphs. It is imperative that every figure is referred to in the main text, and that all figures are correctly labelled, with axes, titles, and legends. All numbers must be associated with their correct units of measure, and statistical tests and descriptive statistics must be reported correctly with standard errors, confidence intervals, p-values, degrees of freedom etc.

It is important to remember that the results section presents the **statistical** significance of the analyses as indicated by the analyses, but makes no biological inference beyond this. Interpretation of the results is done in the discussion. The results section usually contains several figures and tables, but few citations and references in the text.

Discussion

The discussion is the ‘meat’ of your study- it is where you interpret your data and make inferences about the biological significance of your results, placing it in a wider context.

The discussion can be the toughest part of the paper to write, as it should focus on the implication of your results in a wider context, and not get too bogged down with extraneous information. Often, discussions contain information that are better off in the introduction, so it is important to clearly separate background context from interpretive context.

The discussion should link back to the question(s), hypotheses and predictions in the Introduction, examine whether the findings support the hypotheses and compare your findings with those of previous studies. The Discussion should not repeat the results, but may summarise them. It should not include further results that are not reported in the results section.

It is often useful to address each major finding in a separate paragraph, comparing your results with previous studies, and giving potential explanations for any differences. Future directions should be precise and detail the exact information required to improve our understanding of the question further. End with the broader implications of your results.

Acknowledgements

It is highly unlikely you conducted your study completely independently, so it is important to acknowledge and give full and appropriate credit to all the institutions and individuals who helped and supported you.

Some of these people may already be co-authors- those who are not should be listed here, and include:

- institutions who hosted your study (not only your own institution but those who hosted field or laboratory work),
- All funding sources and grants (including in-kind support),
- national authorities who granted research permission,
- field and laboratory assistants,
- People who helped with sampling design and data analyses
- People who commented on the manuscript

References and citations

Your main text should contain citations, which are listed alphabetically in full in the References section at the end of the document. All citations in the text must be listed in the references, and all references must be cited in the text. Therefore it is essential you check thoroughly your citations and reference list to make sure they match perfectly. If you use a reference library software (e.g. Zotero, Endnote) automatic citation tools are available to use within Word – these are highly recommended as they are a great way to save you time and ensure your referencing and citations are standardised and have no mistakes in them.

The usual format for **citations** is to include them in parentheses in the main text as follows:

For one author: (Author surname, Year), for two authors (First Author surname and Second Author surname, Year) and for two or more authors: (First Author surname et al., Year). Multiple citations are usually ordered chronologically.

e.g. “the deleterious environmental impacts of logging can be substantially reduced if appropriate techniques are used (Bertault and Sist, 1997; Durrieu de Madron, 1998; van Kuijk et al., 2009).”

Be sure you correctly identify the **surnames** and do not cite the authors’ **first names.**

References and citations

The format for **references** varies with journal style, but a common style is as follows:

For journal articles:

First Author surname, Initials, Second Author surname, Initials,….last author surname, Initals (Year). Title of article, *Title of Journal*, volume (issue), start page-end page.

e.g. Cannon, C.H., Peart, D.R., Leighton, M., (1998). Tree species diversity in commercially logged Bornean rainforest. *Science,* 281 (5381), 1366–1368.

For book chapters:

First Author surname, Initials, Second Author surname, Initials,….last author surname, Initals (Year). Title of chapter, In: Surnames and Initials of Book editors (Eds). *Title of Book*. Publisher, place of publication, start page-end page.

e.g. Williamson, E.A., Feistner, A.T.C., 2003. Habituating Primates: Process, Techniques, Variables and ethics, in: Setchell, J.M., Curtis, D.J. (Eds.), *Field and Laboratory Methods in Primatology: A Practical Guide.* Cambridge University Press, Cambridge, pp. 25–39.

Other formats are used for books, theses, reports and other publication types and you should ensure you use a standard format and **be consistent** throughout.

Style and language

Once you have got the structure of your document clear, the next point to consider is the **style** and **language** of writing.

The style and language you use will dramatically affect how your message is conveyed and needs to be appropriate to your audience. If you are writing your results for the benefit of local community members, there is no point in using impenetrable over-technical language full of scientific jargon that only academics will understand! Neither is there any point reporting in English to local officials in a French-speaking country. Similarly, if you are writing for the benefit of experts in your field, you don't need to explain basic terms or concepts that they will obviously know about.

In general, most scientific reports and publications are destined for a technical, academic audience and so the style should reflect the level of specialist knowledge that your audience has.

In terms of language, for international publication English is the common language. About 80% of all peer-reviewed research is published in English. As a practicing ecologist, you will not be able to avoid this. If you don't want to publish in English you will at least need to read literature in English.

Also, be aware that the style of scientific writing in English may be different to what's expected in other languages. In general, scientific writing in English should be **concise** –this means using the fewest words possible to clearly convey your ideas.

How to be concise

Being concise means cutting out repetitive and redundant words, and making your sentences as straightforward as possible. It takes practice. Take a look at the examples below and try and see where the improvements have been made. Have a go at writing a more concise version yourself.

1. *Verbose:* A large number of stakeholders from several countries in the region had said that they were likely to be present at the annual meeting that had been arranged for the following month.

   Concise: Many regional stakeholders had planned to attend the upcoming annual meeting.

2. *Verbose:* The experiment was repeated due to the fact that it was necessary for us to take into account the effects of seasonal variation.

   Concise: We repeated the experiment to control for seasonality.

3. *Verbose:* Unpredictable and ephemeral changes in stormwater and currents that vary throughout the day present a tremendous challenge to ecologists hoping to study and monitor water quality in lakes, rivers and estuaries.

   Concise: monitoring water quality in aquatic environments is a challenge, as it varies from hour to hour due to stormwater and currents.

The message

The last – and most important point to consider is the **message**- or story- you are trying to tell. This can also be thought of as the ‘angle’.

In order to get this right, you need to be able to ‘see the wood for the trees’. It is very easy to get so caught up in the detail of your research that you can no longer see the clear, logical pathways leading your endeavors from the beginning to the end. But don’t forget that every story has to make sense.

The most important thing to remember is that you need to ensure that there is a clear and logical flow from your introduction, to your aims and objectives, through to your methods, results and interpretation, and finally your conclusions. This means that the results you present must answer the research questions you posed at the beginning, using the methods described, to address a problem described in the introduction, and interpreted in a way that opens opportunities for future research.

Sounds simple? It takes practice!

2.8

Follow-up

Communicating your research

One of the most important, but sadly most neglected parts of a research project is in the follow-up and communication. Research is meaningless if it is not communicated to the right people!

For this to be done effectively, you need to ensure you know **who** your target audience is, and **how** to best communicate your research to them.

While most studies are published in peer-reviewed journals, it is important to remember that scientific articles are really only likely to be read and understood by an English-speaking academic audience. Therefore, if your target audience is the English speaking scientific community, then publishing in academic journals is the best way to communicate your research.

However, in many cases, your target audience might be park managers, local officials, wildlife authorities or policy makers. Many may not speak a word of English. You may be required to share your data and provide a scientific report as part of your research agreement.

It is therefore essential you have a **post-research communication plan** to ensure all the right people and institutions have been contacted and informed about the findings of your study.

This plan needs to be appropriate to the audience in question. For example, posting links on social media to your new article is not an appropriate way to inform local authorities in developing countries, who may not have internet access and would probably be expecting more a direct, formal form of contact.

Communicating your research

Some ways you can communicate your research to a non-academic audience are as follows:

- Hold a follow-up meeting with stakeholders and give a talk or powerpoint presentation
- Write a report for stakeholders in a format that has been agreed
- Provide posters, leaflets or booklets

Nowadays it is common for studies to be publicised through social media, blogs, websites, press releases, media coverage etc. However, before you go down the route of publicising your work, you should make sure it correctly represents all the institutions that have been involved. This may be a sensitive issue, so it is essential that you contact all partner institutions in advance and ensure you have permissions to use logos, and that institutions agree to the content of what will be published and how they are representated.

The importance of doing this correctly cannot be understated: effective communication will ensure you maintain good working relations with your partners and engender positive collaboration. Failing to communicate only causes resentment and reduces partners' good will to collaborate in future studies.

2.9

Resources

Suggested Reading

- Steel et al., 2013. Applied statistics in ecology: common pitfalls and simple solutions. Ecosphere, 4 (9): 115 https://doi.org/10.1890/ES13-00160.1
- Sampling and the design of ecological studies: http://www.envsci-ed.brockport.edu/Documents/SS95/SS95_Ford_F.html
- Sampling and experimental design: https://www.zoology.ubc.ca/~krebs/downloads/krebs_chapter_07_2017.pdf
- How to write a science research question: https://www.ecologyproject.org/about/blog/how-to-write-a-science-research-question
- Slovin's formula to calculate sample size: https://prudencexd.weebly.com/#
- Sample size calculator: https://select-statistics.co.uk/calculators/sample-size-calculator-population-proportion/
- Sample size calculator: https://www.isixsigma.com/tools-templates/calculators/sample-size-calculator/
- Handbook of Biological Statistics: http://www.biostathandbook.com/testchoice.html
- Design and Analysis of Ecological Data: Study Design Concepts: https://people.umass.edu/sdestef/NRC%20601/StudyDesignConcepts.pdf
- Revising sentences for conciseness: http://www2.ivcc.edu/rambo/eng1001/conciseness.htm
- Effective writing: https://www.nature.com/scitable/topicpage/effective-writing-13815989