MODULE 4: PROFESSIONAL SKILLS FOR BIODIVERSITY CONSERVATION LECTURE NOTES

Learning Objectives

Students will be able to:

- Scientifically collect and record observations about the environment.
- Collect data based on natural history and species identification guides.
- Describe experimental design and the scientific approach.
- Understand the importance of research in tackling conservation problems.
- Understand how the tools of science to inform conservation policy and protected area management.

Key Topics

- What is scientific research and why is it important for conservation?
- The scientific approach
- Statistics Basics
- Survey Design
- Reading Scientific Literature

Key Topics

- What is scientific research?

Scientific research is the systematic investigation of scientific theories and hypotheses. A hypothesis is a single assertion, a proposed explanation of something based on available knowledge, for something yet to be explained. One that is subject to further experimentation. Research in conservation is essential for understanding how the natural environment is affected by human activities. **Monitoring** is the observation of long-term trends.

- Why is scientific research important for conservation?

Science is constantly evolving. As we continue to gather more data and observations, we can continually advance our protection of the natural world. Scientists, naturalists, zookeepers, and wildlife enthusiasts rely on research studies to help fuel conservation management. Research is conducted by many individuals, testing many different theories from many different angles.

Conservation management cannot only be based on one study. When you're trying to solve a problem, you don't just look at one angle: you gather your resources and pick a solution that fits best. Managers often mix multiple research findings to decide best practices to use. Large samples (i.e., many animals or plants) and multi-year studies paint a more accurate picture, too. The more we learn, the better management we can produce.

Comparatively less research is undertaken in the world's most biodiverse countries, the science conducted in these countries is often not led by researchers based in-country, and these scientists are also underrepresented in important international fora. Mitigating these biases requires wide-ranging solutions: reforming open access publishing policies, enhancing science communication strategies, changing author attribution practices, improving representation in international processes, and strengthening infrastructure and human capacity for research in countries where it is most needed.

Scientists try to understand how changing abiotic and biotic drivers will influence growth rate and abundance in plant/animal populations. Environmental changes alter both the distribution and the abundance of organisms. Ecological research on distribution documents either historical range shifts or attempts to predict future occurrence patterns.

Examining the distribution and intensity of threats and species' distribution and density helps when designing protected area networks, determining reserve boundaries or creating corridors that link isolated populations. It is also important when deciding where to invest time and effort in protection or research activities, and provides empirical data to evaluate existing management strategies. Finally, these data are essential for IUCN Red List of Threatened Species assessments, which should be based on actual population size and status.

To manage a protected area, scientific research helps us understand…

- How habitats function;
- How human activities affect these habitats.

An ecosystem is an extremely complex web of interactions. Our actions perturb this web and the interactions. We need to find out HOW and TO WHAT LEVEL human actions create disequilibrium in the ecosystem. Research helps us determine the effects of these human actions.

- The scientific approach

The scientific method entails systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses.

PRACTICUM: "Making Observations"

  - Hypothesis formulation

Type of Question	**Type of Methodology**	**Type of data**
What?	Qualitative	species lists, habitat lists, types of settlements, land uses, NTFPs
How many?	Quantitative	numbers, lengths, areas, weights, temperatures, "how many per unit - DENSITY"
Where?	Spatial	where is x in relation to y
When?	Temporal	phenology, sociological and geological history of a site, breeding periods, lunar cycles
What changes?	Can be all of the above	

To determine what kinds of biological changes have occurred in a study site, you can look for historical baseline data and monitor change; or using the status quo as the baseline, you can plan future monitoring. All of the above parameters can be observed over a longer time period. It is critical to distinguish seasonal changes (or cycles that last several years) from long term, unidirectional changes.

To determine what kinds of biological changes have occurred in a study site, you can look for historical baseline data and monitor change; or using the status quo as the baseline, you can plan future monitoring. Distinguish between seasonal changes (or cycles that last several years) from long term, unidirectional changes. For instance, how has poaching changed over the past 10 years? Ask people to think of examples.

- Causality: Explaining cause and effect

Science seeks to discover the causes of observed phenomena. That is, why does something happen? For example, we might ask ….

- Why do some species of plant fruit at certain seasons?
- Why do some species of animal plant prefer certain habitats?
- Why do some species of tree no longer produce seedlings when the habitat is changed by humans?
- Why has there been a reduction in elephant numbers in Liberia since 1980?
- Why are some fruits red and some smelly?

Answering these kinds of questions involves the data collected in the descriptive phase of research (i.e., what, where, when questios) and combining these different datasets.

- Statistics Basics
    - Sampling and Sample Size
    - Dependent vs. Independent variables
    - Accuracy and Precision
    - Random vs. Systematic Errors
    - Bias

Sampling helps you decide the amount of data you need to collect in order to track a resource. Because sampling always includes some uncertainty, estimation is also required of the precise number of samples needed to confidently conclude that management did or did not work. **Undersampling** (taking too few samples) prevents one from detecting a change even if a change has occurred. Oversampling will let one identify a change in response to management but results in an unnecessary waste of effort. Because monitoring in the field is often very expensive and time consuming, it is important to optimize sampling.

The sampling scheme you select depends on the following:

(1) The entity to be measured

This might be a direct measurement of species/community or an indirect measure of a habitat indicator. Monitoring may involve measuring the change or condition of some aspect of the species itself. If you are monitoring the species, the objective should include its scientific name. If the objective will address a subset of the species (e.g., only flowering individuals, only females), this should also be specified. Monitoring may also measure indicators that function as surrogate measures of species success. There are four classes of indicators: 1) indicator species that correlate with the success of the target species and are easier to measure; 2) characteristics of the ecosystem the species inhabits; 3) threats; and 4) indices of abundance. Monitoring indicators may be less expensive, provide more immediate monitoring feedback to management, and focus on the aspect of the species or community over which you actually have management control, such as habitat quality or intensity of threat. Monitoring indicators may also be problematic, however, because the relationship between an indicator and a particular species is usually hypothetical, or at best only partially understood. Monitoring an indicator may thus result in false conclusions about the condition of a biological resource.

(2) Attribute to be measured

This is the specific attribute of the entity to be measured. Often this will be a parameter such as size, density, cover, or frequency. It might also be condition or a qualitative measure (e.g., many, few, none).

To sample, you repeatedly measure the environment in a quantitative fashion to determine if your management plan has succeeded in changing (or maintaining) the state of the resource in question. In other words, the sampling objective for conservation is to estimate the parameter in the population (the “what, where, when”) under management and compare this estimated value to the threshold value desired.

Different minimum sample sizes are required to meet different survey objectives. Also, sample size suggestions can vary depending on the techniques used to survey the targeted community. These must be carefully considered prior to fieldwork. In general, the greater the sample size the better the precision at estimating species or population trends. A preliminary or pilot survey would help to test survey techniques and provide an assessment of the site (e.g., determine distribution patterns) so as to determine (or modify) sampling effort. Even a single specimen representing a species is sufficient for information related to presence-absence. When the sample size varies among populations or species, an evaluation of this variability should be investigated in order to gain a realistic assessment of abundance, richness, or other estimator.

Modern surveys try to cover the entire area of interest using an evenly spaced sampling plan, so that the sampling is representative of the whole site (whether it is a protected area, logging concession, community forest, or other type of land use).

You may want to know something quantitative. The question may be …

- “How many pygmy hippos are in Sapo National Park?”
- “How often do zebra duiker come to this area?”
- “How many Azobé (*Lophira alata*) trees are there per hectare in a typical wetland in SE Liberia?”
- How old is the average *Famira* tree?

It is impossible to count all the individuals in a population. So, we must **sample. Sampling** means to count the number of things or events using a pre-determined sampling methodology. This can provide an accurate **estimate** of the “number of things” you want to measure in the study area. In order to produce an accurate estimate, you have to develop a robust **sampling design**, which may be unique for each new study.

If you are sampling **density** of plants or animals, the bigger area to which your final estimates apply is called the ***sampling area*** (e.g., a national park, the area within 5km of a road, all habitat types in a wetland area). S**ampling units** are smaller sections within which you count the plants, animals, humans, or events. These can be plots or transects (area, length) or fixed amounts of time per hour, or units such as households.

Because it is impractical or impossible to count every individual in most populations or communities (groups of populations), biologists measure biodiversity by first sampling the organisms and then extrapolating to estimate the total number of organisms. For example, to compare the number of bird species in different types of forest, biologists record the number and species of individual birds encountered at randomly selected locations within each forest type. Population biologists compare the average density of the individual species in each forest type. Community biologists compare the average number of species in a given area, such as a square meter or square kilometer, or the diversity index in a given area. The higher the diversity index, the more species and the more even the distribution of individual organisms among these species. Biologists interested in genetic or ecosystem diversity rely on similar sampling procedures and diversity indices.

Sample size …is important because the bigger your sample size, the better your precision. This is because your variance goes down.

    - Random vs. Systematic

Random sampling means all the points in your study area have an equal probability of being chosen so that systematic bias will not affect the sample.

Systematic sampling. Things are not usually systematic in space in nature, so a random start point can be the corner for a systematic design (plots or transects). If you deliberately choose plots transects for convenience, you BIAS your sample

    - Bias

Bias can result in the overestimation of human signs and under-estimation of animal signs. For example, hunting and trapping is usually more intensive near roads so if a survey only takes place near roads, the sample will be biased. Surveys are also subject to bias if data collectors walking along existing roads and paths and if they avoid wetlands and other habitats difficult to traverse. Bias also happens when an intensive survey is conducted in small area and the results are extrapolated to a much larger area without knowledge of different habitats or hunting pressures in areas not surveyed.

What kinds of bias should we consider?

- Spatial: How might sampling near a village affect results from plots?
- Temporal: How might season affect our results? Time of day?
- Social: How might speaking only to men in a village affect a survey?

- Dependent vs. Independent Variables

Does variable x affect variable y?

INDEPENDENT: what might affect the thing you are measuring
DEPENDENT: the thing you are measuring

Examples of **independent** variables:

- rainfall
- temperature
- soil type
- depth of the ocean
- patch size of a forest
- distance from a village

Examples of **dependent** variables:

- size of tree density of elephants
- number of species in a forest patch
- seasonality of fruiting type of
- hunting technique used.

That said, independent variables may be dependent variables based on the kind of study that is being conducted. For example, rainfall is not always an independent variable. Rainfall may depend on something else such as temperature. You can also have two dependent variables e.g. populations of predators and prey, which are interlinked. And both of these may depend on a third, independent variable.

Some variables covary
“Why has there been a reduction in elephant numbers in Liberia?” There are a number of possible, and nonexclusive, reasons:

- Demand for ivory continues
- Abundance of highly efficient guns since and during wars in the region
- No control over poaching
- Removal of necessary foods or shelter due to logging
- Others?

- Hypothesis formulation

You have a set of observations. This comes from your descriptive data collection activities. Now you want to know the reasons for the patterns you observe in the data. You have formulated possible explanations. Hypotheses are expressed as statements. They are explicative, offering possible cause-and effect statements about the observed phenomena. This is what a **hypothesis**: an explanation for observations. The pattern that you observe may be explained by one of several hypotheses. You should construct these so that they are *exclusive* – they do not overlap. Hypotheses should be **testable**.

STEP 1: Formulate hypotheses
Observation: The density of *Haumania danckelmaniana* stems differs between two sites studied. One of the sites is on flat ground and the other is on a slope.

Hypothesis (H_1): Differences in soil moisture are responsible for differences in [species name] abundance.

Null Hypothesis (H_0): Soil moisture does not affect the abundance of [species name].

STEP 2. Make predictions based on these hypotheses.
You then have to make predictions as to what would happen if this hypothesis were true.

Prediction: There is a significant relationship between soil moisture content and [species name] abundance.

STEP 3. Design a test of the predictions
In other words, design a way in which you collect new data – or use existing – relevant data which would test your hypotheses.

Your chain of reasoning is as follows.

1. Assume that the averages of the soil moisture values at the two sites are the same. (H_0)
2. Take several measurements from both sites and see if the values from the two sites are, on the whole, close to each other or far apart. How to do this? By….

- Survey Design

PRACTICUM: COLLECTING DATA

- Data collection, entry, and analysis

When collecting data, you must think carefully about how to record observations and make the effort to do so rigorously. Making observations and recording information in a conscientious and rigorous way is something that must be learned and practiced, since poorly recorded observations can be useless or even misleading. *Always* carry a notebook into the field, to record interesting observations. You can't record everything, so you should make or follow existing guidelines. Also, record unusual observations. So, you need to know what is unusual. You should get to know the fauna (all mammals, as many birds as you can, all big trees) as possible so you know if something unusual turns up.

Write down exactly what you observe. Good notes last forever, for example, we still refer to Charles Darwin's notes from the 1840s. Your notes should be clear, legible, and ordered logically or even you will not be able to understand them later. The first step to taking good notes is to be a good observer. You must be clear about what it was you saw or heard, or just as importantly, what you didn't see or hear. If you see a monkey eating a bird, do not assume that the monkey killed the bird. Do not assume you know what will happen next. Record any interpretations separately from the observation itself. Focus on the important details first. On your data sheets, fill in all the boxes if you can but do NOT guess, say "I don't know" if you don't. Record absence as well as presence of animals, plant species, humans, etc. since negative data is useful, too.

- Making notes and records
    - Field books, sketches, data sheets

Field Books

Always write with sharp pencils because it will not be affected by water. Field books are good for non-numeric data and you can draw sketches/maps etc. in them. Field books should have the name and the title of the book on the cover; the start date; and, eventually, the end date. Number all the pages – don't rip out pages- if something is wrong, put a big "X" through it; otherwise future scientists may think information is missing.

Field notes

Field notes contain Information collected during day to day activities. You should write down any observations of note. For example, in addition to observations in the forest, information on road kills, bushmeat for sale, etc. can be recorded for wildlife studies. You must record each trip or forest walk whether or not you see animals. For example, if you are collecting data on animal abundance, you should, at the same time, collect data on human signs. This means that, even if your initial interest was animal abundance, you can later see if it related to the abundance of human signs.

Sketches

Emphasize key distinguishing features, such as color and shape, and accurate measurements should be written next to the sketch. Even if you can't draw, it helps a lot. Sketch tracks together with measurements of their length and width. Also note the distance between individual tracks and the track arrangement.

Data Sheets

Data sheets allow for standardized data collection between observers or over time. They also summarize a lot of info quickly, and inform eventual analysis. The top of every data sheet should have:

- type of data;
- names of observer(s);
- date, time, and location of observations
- weather
- habitat type(s)

Data sheets must always be stored properly.

Log Books

People should transcribe information from their field books to log books because personnel changes. You can have separate log books for specific species, e.g., “primate logbook” or issues, e.g., “village history” book, etc.

- Communication:
    - Data analysis, report writing and presentation

PRACTICUM: DATABASE EXERCISES

STEP 4. Make a preliminary summary of the data [CONTINUED FROM ABOVE]

Use a graph or table to report the mean and a description of the variability of the data, as well as the distribution. It often helps to visualise the data in some graphic form; a table often gives less clarity than a graph.

- Inputting and Analyzing data
    - Rarity
    - Diversity & representativeness
    - Trends

STEP 5. Analyse the data using statistical tests.

A research project is not finished until you have analyzed and interpreted the data**,** and made a report so that others can read it. **Data interpretation** is the process of giving biological meaning to the results.

- What do the data tell me?
- What do the data not tell me?

Look at the data to see if there the observed differences between data in two (or more) situations are statistically significant**.** Or if observed variables are **correlated. Data analysis** is the process of taking the numbers collected in the field and summarizing them into a form that is easy to interpret (graphs or tables). This summary is called the **results.** Results should be put into their biological context.

If the data were carelessly collected or if the research design was poor, the results are meaningless, i.e., "Garbage in, garbage out." You should always ask yourself the following questions

- **Accuracy?** *Were the data collected carefully?*
- **Training of assistants?** *Were data collectors adequately trained?*
- **Appropriate methods?** *Is this the right method to accurately capture observations?*
- **Assumptions of the methods?** *Do the methods employed fit the research question?*

Conditions assumed to be true while data are being collected in a plot, e.g., no tree is counted twice in a tree census; no tree outside the plot is counted; all trees inside the plot will be counted. I.e. good data collection. If you are counting moving objects, a plot may not work as you might count an animal twice. Different methods make different assumptions. Choose the one where you can meet more of the assumptions.

- **Was the timing right for the species? (time of day or year)** *Is the presence of the species in question affected by the seasons.*

The detectability and abundance of biodiversity values may vary temporally, including time of day, time of month (e.g., in relation to phase of moon), time of year (e.g., local or large scale movements or migration), seasonally, annually, and over periods of multiple years. Variation on longer scales is also possible, for example, caused by climatic phenomena such as El Niño events. It is good practice for baseline surveys to be structured in such a way as to help understand regular large changes in detectability and abundance of biodiversity values that may occur in the baseline study area over time (e.g., wet and dry seasons at tropical sites; some combination of spring/ summer/autumn/winter at temperate sites).

Statistics

Statistics give you a probability that your samples come from one or more different populations. They are not absolute (remember rejecting null but not necessarily completely accepting the alternative hypothesis. **Statistically significant results** in biology are those where the probability is less than 1 chance in 20 that your observed difference in the sample is due to random chance.

Conventions:
Less than 1 in 20
5% level
$P < 0.05$
$0.01<P>0.05$
*these are identical

We use statistical tests to find out if the trends we see in the data have a high probability of being representative of reality.

- Types of measurements

Categorical (nominal) = color, sex, habitat type
Ordinal (ranked) = Bigger than, hotter than, faster than
Interval = real numbers

- Frequency distributions

Normal
Poisson
Bimodal – frequent in biology because of season)
Binomial (also common where there are two mutually exclusive possibilities : presence/ absence data)

Effect of sample size (draw the 4-histogram example, page 37): the distribution becomes clearer as N rises. Another way to improve your curve is to lump your data in classes.

- Continuous and discrete variables

Continuous = any fraction of a number
Discrete = whole numbers only

- Statistics Basics: mean, median, mode

Mean. Also known as the 'average'.

For example:
Dbh measurements of 85, 92, 73, 68, 101 cm.
N = 5
Sum = 419
Mean = 419/5 = 83.8

if the N of the original data is less than 100, record the mean to 1 decimal place. If N was 100-999, record the mean to 2 decimal places, etc. The biggest number strongly influences the mean.

Median is a resistant measure of the average because it is not skewed by exceptional data.

For example:
4, 8, 9, 12**,** 15, 25, 202
Here, **12 is the median.** It is resistant to the extreme observation.

If there is an even number of observations, the median is the mean of the values of the middle pair.

For example:
9.2, 11.5, 13.2, 19.7, 29.4, 50.1
Median= (13.2+19.7)/2= 16.45

The **mode** is the most common (frequent) value in a sample. It is the peak of a histogram. This is the only measure of the average that can be applied to observations on an ordinal (ranked) scale.

- Accuracy and precision

Accuracy is the closeness of a measured or computed value to the true value. You can increase accuracy by avoiding bias and ensuring your methods are appropriate. High accuracy is needed when estimating total population size, because the real number affects quotas as a function of viable populations.

Precision is the closeness of repeated measures to one another (like variance, standard error..)**.** You can increase precision by increasing sample size and using the same well-trained field staff each time. High precision is needed when looking at trends, otherwise you cannot detect change.

- Random Error and Systematic Error (Bias)

Each unit in the population should have an equal chance of being picked for the sample. **Random error** affects the **precision (variance)** but not the **accuracy** of a sample **b**ecause there will be the same number of erroneous readings above and below the true value. **Systematic error** affects accuracy, because it pushes the numbers to one side or the other of the real value.

- Statistical Errors

Sampling always involves some uncertainty because with sampling we are never entirely sure that we have properly estimated the true value of the parameter for the population. We have to consider the possibility that any difference that we see between two estimated parameters (before and after) could result from sampling errors. **A Type I statistical error** occurs when two populations sampled by chance give the incorrect appearance of being different when, in fact, they are not. It happens when our random sample is not representative of a population as a whole. We also want to be careful about committing a **Type II statistical error**, which involves wrongly concluding that there is no difference among data sets when in fact we simply failed to sample adequately to detect it. We guard against both of these kinds of errors by using increasingly smaller and more stringent levels of alpha or "significance levels."

Increasing the alpha level comes at a cost, however, because it generally involves more sampling to get higher confidence, and gathering samples takes time and money. In addition to balancing cost versus confidence we also need to worry about precision when formulating sampling objectives. How accurate do we really want or need to be? To some extent, the answer to this question depends on the importance of what we are monitoring. Resources of high value, such as quality timber and endangered species, may need to be estimated with greater precision than less critical resources like forage levels or water quality. General estimates, and hence less precision, may suffice for these latter two because the consequences of incorrectly estimating them are less dire (e.g., a slightly thinner cow versus an extinct species).

Once you have collected your monitoring data, you can compare it to the baseline data you collected earlier—presumably before management actions occurred. By contrasting conditions before and after management, you can evaluate whether your have met your management objectives, and hence whether you should continue or alter your management practices.

If you are estimating a quantity based on a single independent sample (i.e., you are not trying to relate the sample to another year or another site) then calculating the precision of your estimate using confidence intervals is the correct approach. Confidence intervals can be calculated for a mean, proportion, or a population estimate. Examples include total number of individuals within the sampled area, mean number of individuals per unit area, the proportion of quadrats (generally rectangular plots used in ecological and population studies) occupied by the species, the mean height or weight of individuals within your sampled population, the proportion of occupied nesting boxes, the mean number of motorcycle tracks per unit area, etc.

- Significance Tests

Significance tests are used to assess the probability of an observed difference being real or simply the result of the random variation that comes from taking different samples to estimate the parameter of interest.

One important distinction to make is whether the significance tests are for independent vs. paired samples. **Independent samples** are ones in which different sets of sampling units are selected randomly (or systematically with random starts) in each year of measurement.

Another important distinction to make is whether the data are **parametric** or **nonparametric**. Data that are parametric are typically those that form an approximately “bell-shaped curve” when their frequencies are plotted. Non-parametric do not display this characteristic pattern, e.g., count data that have many zeros and a few counts > 0. This distinction is important because different kinds of statistics are used to analyze parametric versus nonparametric data.

- Measures of variability and confidence limits
    - Variance

$$\text{Variance} = \frac{\text{sum (each value - average value)}}{\text{Sample size - 1}}$$

For example:
81, 79, 82, 83, 80, 78, 80, 87, 82, 82 (dbh of a sample of 10 trees)

1. First calculate the mean: this is 814/10 = 81.4

2. Get each deviation in turn by subtracting the mean from each observation in turn. Then square the deviation. This eliminates minus signs

$(81-81.4)^2=0.16$
$(79-81.4)^2=5.76$
$(82-81.4)^2=0.36$
$(83-81.4)^2=2.56$
$(80-81.4)^2=1.96$
$(78-81.4)^2=11.56$
$(80-81.4)^2=1.96$
$(87-81.4)^2=31.36$
$(82-81.4)^2=0.36$
$(82-81.4)^2=0.36$

3. Add up the 10 squared deviations. (56.4)

This is called the sum of squares of the deviations, or more simply the **sum of squares.**

4. Divide the sum of squares by one less than the number of observations (N-1)

56.4/9 = 6.27. This is the **variance**.

- Standard deviation

Standard Deviation = √Variance

Standard deviation is “the average deviation from the mean” – it measures the ‘tightness’ of the data around the sample mean. Standard deviation is calculated as the square root of variance

In the example above, **variance** was 6.27cm. Therefore the **standard deviation** is √6.27 or SD = 2.50

- Coefficient of variation

This presents the SD as a percentage of the mean.

CV = (SD/mean) x 100
So CV = 2.5/81.4) x 100= 3.07

So, the standard deviation is 3.07% of the mean.

Standard error is the average deviation of sample means from the population mean. Basically, it is dividing the variance by the sample size.

<u>Extrapolating from your results: what you can and can't say</u>

What you can say is "given the methods used", the results "from this area", "collected at this time", "under these conditions" are as follows. You cannot say anything more than that with certainty.

If you want to estimate elephant density across a whole park, you have to take samples throughout the park. If you only sample around a specific area – because it is easier – you may think you are measuring typical mixed forest, but what you may not know (and will fail to sample) are elephants in other areas.

You can speculate and discuss the implications of your results (what the results suggest but do not show conclusively), and **extrapolate** the results of your study to larger areas (draw conclusions, based on the data, about areas for which you do not have data), as long as you clearly explain the steps you took on the way to making your conclusions.

This kind of data interpretation is not the same as results, and should not be presented in a "Results" section of a report (see below). Results are based on field data and should be presented in the "Results" section without interpretation. A discussion of possible meanings of the results, and extrapolated estimates, are presented in the "Interpretation and Discussion" section, where the reader is free to agree or disagree with your ideas.

STEP 6. Accept or reject the hypotheses.
The **null** hypothesis (no difference between 2 situations) is the opposite of the alternative hypothesis. If there is a significant difference between two situations, you can reject the null hypothesis.

- Reading Scientific Literature

Primary literature is the first, formal publication of new scientific data or ideas based on the results of original research. Scientists publish their findings in the form of a scientific paper to share their data with others and advance their field. A scientific paper also outlines research methods so that readers understand how the study was conducted; this makes it possible to replicate and verify results. Furthermore, the paper should be published in a peer-reviewed journal that is accessible to others in the scientific community. In peer-reviewed journals, submitted articles undergo review by a small selection of scientists, so the research presented in the paper will already have been rigorously examined prior to publication. This process of peer-review is integral to validation of scientific findings, and helps scientists to understand and interpret their results from a variety of perspectives and to produce refined products.

Characteristics of primary literature include:
1) First publication of results from original research;
2) Replicable methods for peers to test the results; and
3) Publication in a journal easily available to the scientific community.

Examples of journals relevant to biodiversity conservation:
- *Conservation Biology*
- *Biological Conservation*
- *Ecological Applications*
- *Journal of Wildlife Management*
- *Trends in Ecology and Evolution*
- *Nature*
- *Science*

Secondary literature provides an overview, analysis, or interpretation – and sometimes a simplification – of the primary literature. Some secondary literature undergoes a peer review process, but some does not. Secondary literature is usually tailored for a particular purpose or audience that can consist either of readers with little knowledge of the sciences or of specialists in the field. In general, publications that are directly derived from primary literature can be considered secondary literature. Monographs (books), textbooks, and review papers are good examples of secondary literature. Review papers are those in which scientists summarize the historic and current literature on a topic.

Characteristics of secondary literature include:
- Content: Summary and synthesis of existing information
- Audience: varied (though generally catered to a specific audience or purpose)
- Review: Sometimes peer-reviewed
- Publication format: books, journals, reports
- Literature citations: Provided

Relevant examples:
- *Annual Review of Ecology and Systematics*

- *Conservation Biology*
- *Trends in Ecology and Evolution*
- *Science*

    - Writing Science

PRACTICUM: SCIENTIFIC WRITING

Structure of a Scientific Paper

Most scientific papers start with a description of a problem and present results that lead to a conclusion about that problem. They generally follow this format: abstract, introduction, methods, results, discussion, conclusions. There are exceptions to this format depending upon the topic, where you want to publish, and your audience. However, although formats and conventions vary from journal to journal, you should always write clear, concise, and well-organized papers.

Major Sections of a Scientific Paper

The title

A good title is essential. While many will read the title, comparatively few will actually read the paper. The title is your opportunity to draw people in to read the paper. Good titles are a reasonable length (fewer than 15 words if possible) and accurately represent the content of your paper. The title should include the primary species or ecosystem covered, the geographic location, and the type of study. You want other researchers with similar interests to be able to find your paper, so think about what terms someone might use to search for articles like yours, and try to incorporate some of those terms in your title.

The by-line

The by-line provides the author and the institution where the author was when the paper was written, and/or when the investigation was conducted. When more than one person contributes to a paper, the question arises, who should be listed as authors and in what order? In general, to be considered an author, a researcher's contribution to the paper should be substantial, including participating in one or more of the following: developing the idea, planning the project, conducting the research, analyzing and interpreting the data, or contributing to the preparation of the article through drafting or revising the text. The relative contribution of each author to the manuscript should determine the sequence of authors.

Key words

Most journals require you to select five to eight key words that describe the main topics covered in your article. These key words should be terms that are commonly used in the literature and should be chosen with the objective that other researchers would think to search an index for those terms if they were looking for articles like yours.

READING: "Writing an Abstract"

Abstract

An effective abstract captures all the most important points of an article. It is short, generally 250 words or less. An abstract encompasses the paper's objectives and scope (why the study is important), results (what were the major findings), and main conclusions (what do the findings really mean). It might help to think of the abstract as a miniature version of the main paper (with the four major parts of the main paper – introduction, methods, results, and discussion –

reduced to one to three sentences). All information contained in the abstract should also be in the body of the paper. Abstracts should not have cited references, so be sure to write in a way that no references are required.

Introduction

The introduction should provide readers with a broad research context that sets the stage for justification of your study, and explains why they should be interested. It is critical that readers are convinced that the study was worth doing. State the problem/question clearly at the outset so the reader wants to read further to find what you discovered. Then, summarize the most relevant literature so the reader understands how your research relates to earlier research. This will put your research in context and help you more clearly articulate the problem you are solving. To sum up: start broad, set the context and justification, and then narrow in on the focus of your study. Introductions are generally not long – just long enough to draw the readers in and provide a context for the results and discussion.

Materials and methods

In this section, you should explain to the reader the process you used to obtain evidence to answer your central question. This will allow the reader to evaluate whether your methods were well designed – that is, whether your methods provide an effective means for obtaining evidence. Ideally, you should present enough information that if readers wanted to, they could replicate your study. An essential characteristic of a good scientific study is that the results should be reproducible if another researcher employed similar methods and materials. For field-oriented research you should provide a good description of the study area. The following topics are useful to cite: geographic location, altitude, general climate data, geology, soils, historical background, and general information on vegetation and or characteristic fauna. Describe your procedures accurately and in detail. Always include the dates (duration and sampling dates), location (if possible, provide latitude and longitude coordinates), as well as the time when the research was undertaken, be it a field or a laboratory study. Also of interest are details that might influence the data, such as the time of study and seasonal factors, such as temperature or weather. The materials section should also describe instruments/tools, sampling devices, or other equipment.

Study organisms

Give complete taxonomic information about the organism you are studying to the most precise level possible – genus, species, and subspecies. There has been much debate over the use of common names compared to scientific names in publications. Generally, one should avoid using common names for species because these names are often used inconsistently from one region to another. Therefore, it is better to use scientific names. A common name may also be provided in parentheses after the first use of the species name, if this is deemed useful, but from thereon the scientific name should be used consistently through the manuscript. Some publications (especially those aimed at a broader, non-scientific audience) may have a policy where common names are preferred; this may be more typical for well-known groups of species such as some species of birds or mammals.

Results

The results summarize your findings, drawing the reader's attention to especially important results, typically using tables and figures that present the data in more detail. The point is not to provide a long, boring list of information, nor is it to interpret what the findings mean. For example, when presenting statistics, reference them in support of your findings, but not as findings. You should focus on reporting major findings (versus statistically significant findings) that are directly relevant to the justification and objectives outlined in your introduction, and avoid including unnecessary details. It is good to include in your results what you *did not* find, as this can help the reader to better understand the problem. You should also note any anomalous results you obtained and discuss them further in the discussion section. Discussion of anomalous results in the results section should be limited to those stemming from methodological issues. No results should contain untrue information. Incorrect data is inexcusable.

Figures and tables
Tables, graphics, and statistics should not be used to present your results, but to provide supporting data for your results. They should be used when visual aids can more quickly and effectively convey key points or results, and they should make it easier for the reader to see important trends or comparisons that you extracted from your data to support your arguments and conclusions. The type of visual that is used, whether it is a table, graph, illustration, flow chart, etc. should be carefully chosen to appropriately and effectively convey the ideas or trends that you wish to highlight.

Methods for presenting different kinds of data

Types of data	Presentation methods
Exact values or totals	tables, lists
Complex relations between groups of data	diagrams
Differences and comparisons	bar graphics
Tendencies and interactions among variables	line graphics

Discussion
The discussion is where you interpret results, setting them in the context presented in your introduction. In fact, what you write in this section should validate the choice you make for your paper's title. What do your findings mean? Why did you get the results you found? Many papers that have high-quality data are rejected because the discussion is incomplete, poorly focused, or otherwise disorganized and unclear. Keep in mind that if your introduction is not well constructed, it will be difficult in this section to convince readers that your results provided you with relevant and applicable information about the problem to be solved. Below are some guidelines to follow:

- The discussion should lead off with the most significant result, and its importance; do not start by explaining exceptions or negative results. The idea is to present generalizations or trends that are supported by the results and reference key figures and tables where necessary.
- Do not repeat what you have already said in the results section.
- You should compare your results with those of other studies, noting how it extends or clarifies other research. Emphasize similarities or differences between your findings and other research and discuss the reasons for these findings.
- Do not provide in the discussion section data that were not presented in the results.
- If you think that your results tell you something that is not currently supported by the data, you can speculate on possibilities, but do not let speculation overtake your other

conclusions. This is very important – many papers are rejected because they include too much conjecture.

- State your conclusions clearly and make sure that they are fully supported by the data that you presented. It is tempting to overstate the scope or accuracy of your results. Provide alternative hypotheses where relevant.
- Scientists often summarize their findings and provide suggestions for future research and for application for managers at the end of the discussion section.

Literature Cited

The last section of your paper will be a list of the literature that you referenced throughout your paper. Note that “cited references” and “literature cited” are used interchangeably to denote the references (also called citations) that appear at the end of an article or book. Usually this list is presented in alphabetical order by author. Prior to submitting your manuscript, double check that all of the references you cite in the paper are in the literature cited section. Conversely, be sure that you only provide references that are cited in the paper. Sometimes you are asked to provide further readings in addition to the literature cited – if so, then you do not need to cite these in your paper; rather, this is meant to be a list of references related to the topic of your study where a reader could find additional information.

Many people fail to understand the importance of citing literature appropriately. You must give credit to the original authors for any quotations, ideas, images, or data that are not your own. Often papers contain too few references; at its worst, this can lead to allegations of plagiarism. Paraphrasing by taking someone’s sentences and changing a word here or there qualifies as plagiarism, as does not indicating a quotation by using quotation marks, or not crediting the sources of a quote. Do not forget that web sites are considered to be published and protected by copyright, just like print publications. Plagiarizing information from a web site is just as egregious as plagiarizing from materials published in print. Furthermore, figures, illustrations, and tables often draw from other’s work and these sources should also be referenced. Indeed, all facts, ideas, and data taken from other sources need to be appropriately cited in your work. The only exception is when you present information that is common knowledge – for example, the size or location of a country qualifies as common knowledge so these data do not need to be credited to a source. If you have any doubt about whether or not something should be cited, err on the side of caution and cite the source.