Online publication

Beyond the methods manual

Some things to know while you work on your first research project

Donald B Zobel

Department of Botany and Plant Pathology, Oregon State University, Oregon, USA
For correspondence, email: zobeld@science.oregonstate.edu

Experienced scientists have learned how they can best do science, but not from books and at the cost of much time and much wasted effort. Some scientists have tried to pass on their experience of doing science; for example, Karban and Huntzinger's book "How to Do Ecology" (2006) is a recent, useful attempt. I recognized the need for this type of assistance for beginning scientists in Nepal in 1985, while teaching at Tribhuvan University, Kirtipur. A methods manual we wrote then (Zobel et al. 1987) has been well used, but covers only specific methods. I now have 40+ years of experience doing plant ecology research and teaching, including lots of time working with graduate students on their first attempt at research, and several years experience in the Himalaya. So here is my attempt to help you answer some of the vexing questions you will have while setting out to complete a research project. Although my experience and examples involve plant ecology, the general ideas apply to many areas of science. I hope these few pages will make your research easier and more productive.

Getting started as a scientist is a daunting task: there are so many new words, such complex and interlinked ideas, such a wealth of information that is interesting and useful! And then there is another giant step in your development as a scientist, starting to do research yourself—so many possibilities to choose from, so much information to gather, so many things to plan. How can you find out what you should be doing? Fortunately, there are many books and journal articles about research methods. But these usually give only details of individual procedures that you can use to get a single kind of data. Using one method correctly is far from "doing science". What should be done first? Then what? How should you put all the pieces together? Where can you find help when you are unsure what to do next? Methods manuals do not provide answers to such questions.

Choosing a research topic and objectives

One of the most important steps in research is to define the topic to be studied. You can define your topic in different ways: you may use your research to answer a question, to clarify a pattern, to solve a problem, or to test an hypothesis. Select the topic carefully, because research is expensive and takes much time. Your topic should arouse your curiosity and be important to you, in order to keep you motivated enough to complete the project and to do a good job.

Many topics are suitable for research in ecology. You must select both a biological system and an aspect of its ecology that you will study. Ecological systems range in complexity from a local population of a single species to a landscape including many ecosystems to the whole earth. In some cases, the emphasis of research is on a biological system that is of special interest (for example, a specific species, population, community, or ecosystem). In other studies, emphasis will be on a property of a biological system (for example, its distribution, structure, development or function), so you need to select a system that allows effective investigation of the property you chose to study.

Once you have selected a general topic, you need to define the area of study more narrowly, identify the effort it will require, and make sure that it is of a size appropriate for a single research project. For example, if you wish to study autecology (that is, how a single organism or species responds to environment), you can choose among four general types of characteristics that affect the adaptation of an individual (Grubb 1977):

a) Tolerance of mature plants to physical and chemical factors;
b) Life form (size, shape, three-dimensional structure);
c) Phenology (timing of life-cycle processes); and
d) Method of reproduction.

A given autecological project may study only a few of the possible subjects within one of these four subdivisions, or may include parts of more than one subdivision.

Himalayan Journal of Sciences Online Publication
Publication No. 2011-01: 1–6

When developing a study plan, consider doing the research in a *comparative* manner, using two or more systems (populations or species or communities or ecosystems). Using more than one system increases the generality of your results. Also, it is often easier to interpret data for several systems in a comparative manner than to interpret data for one system that you study in isolation.

Once you have chosen the research topic, select specific objectives you wish to fulfill. If you can, state the objectives as testable hypotheses—but remember that a topic needs to be well understood before useful hypotheses can be stated. If no good hypothesis is available, explicitly state the question you wish to answer, or the problem you wish to solve, or the pattern you wish to investigate. In some cases, there may be more than one reasonable hypothesis to explain a given observation. If so, you should state "multiple working hypotheses" (Chamberlin 1965) and design the study to determine which hypothesis among those you are considering is the most reasonable.

Choosing appropriate research objectives is difficult. Knowing the situation in the field, having a sound knowledge of biology, reading widely in the specific area of research, and consulting with more experienced scientists all can help you select good objectives. You need to assure that the objectives you choose:

a) are of scientific or practical importance, preferably both; that is, your report about the research should be useful to other scientists or to those who use the results of science;
b) will be a new contribution to science: that is, the results cannot be derived from work already completed;
c) are stated with enough clarity and precision that you can determine exactly which data you will need to fulfill your objectives, thus ensuring that you will not collect unnecessary data;
d) can be completed with the time and resources that you have; and
e) are of great interest to you, to keep you focused and enthusiastic about your research.

The general approaches to ecological research

Different approaches to research require different methods and often lead to different types of conclusions. After selecting a topic, but before developing a research plan, you need to identify the most appropriate general approach to research for your topic and objectives. For example, there are at least three approaches (Grime 1979) used to study the autecology of plants:

a) *Correlative research* relates two variables; the variation in one variable hypothetically accounts for the variation in the other. An example could be to determine whether there is a significant correlation between the number of herbaceous plant species present in forests and the amount of light penetrating through the tree canopy to the forest floor.

b) *Direct research* records the establishment, growth, reproduction, and death of individual plants in small plots, usually within natural vegetation; the ecologist observes natural selection as it occurs. This approach is important for the study of the population biology of plants (Harper 1977).

c) *Comparative* studies examine the growth, physiology or life cycle processes of groups of plants that differ in their ecology. Usually many different species are tested for one aspect of their physiology in laboratory or garden conditions, using standardized methods.

Other distinctions between research methods result from a series of choices you must make:

1. Do you wish to confine your research to *observation*, in which you do not change what occurs in the field, or do you wish to do an *experiment*, in which you manipulate nature (for example, you add fertilizer, or you remove competitors or herbivores)?
2. Will you work only in the *field* or only in *controlled conditions*, or in both situations? Use of results from studies of physiology or interactions among organisms in a garden, greenhouse, or laboratory may help you to understand the field situation; on the other hand, such results sometimes differ from what happens in the field. Drawing ecological conclusions based primarily on results from controlled conditions is risky.
3. Are you seeking a *mechanistic explanation* for a phenomenon; that is, will you seek to identify the full set of causal relationships that connects two variables? In some cases, research workers choose not to take the often enormous effort to find a mechanistic relationship (if it is even possible) (Wiener 1995), but rather seek only to define the relationship between two variables, one easy to measure and the other which is hard to measure but important (*calculation tools,* Peters 1991).

In any ecological research project, it may be important to have high levels of three aspects of quality for your data (Harper 1982):

a) *Precision* is the ability to detect small differences reliably;
b) *Realism* is the assurance that plants will react this way in nature; and
c) *Generality* is the likelihood of finding the same result for areas and organisms besides those you studied.

If you have the usual resources and time to do your research, you cannot simultaneously produce high levels of precision, realism and generality with the same research design. Therefore, you must choose which one or two among realism, precision, and generality are most important for fulfilling your objectives. For example, studies done in a greenhouse, to increase precision, lose realism. In contrast, intensive direct observations in the field increase realism and precision but, being necessarily limited to a very small area, lack generality. Finally, studies that scatter samples over a large area increase generality but lose precision.

Choosing specific methods

After selecting the objectives and general approach for a project, you must find the specific methods to produce data that best fulfill your objectives. The "best" method is not necessarily the most expensive, recent, precise, or accurate method. The best method is usually the one that allows the largest sample size possible while providing acceptable accuracy and precision. Although it may be possible to make very accurate and precise measurements, the time and expense required for each measurement reduces the number of samples that you can afford to take. Because natural systems are extremely variable, an ecologist usually needs as many samples within the study as he can afford to measure. For example, if you are studying more than one system (that is, population or ecosystem), you must decide whether to sample intensively within a few systems, or you can sample each system less intensively, allowing you to study more systems.

When measuring the environmental variables to which your study organisms respond, select only factors that are likely to influence the organisms and process you are studying. Measure only the aspect of an environmental factor that most directly affects the plant; for moisture, for example, measuring the water potential of the plant is more likely to be useful than measuring soil moisture or rainfall. Place the sensor of your measuring device in a place that is as similar as possible to the environment to which your study organisms are responding. For example, if you are studying the physiological or growth response of a leaf to temperature, measuring the interior temperature of the leaf tissue is best. If that is not possible, measure at the leaf surface or in the air adjacent to the leaf. Make measurements only at the time when the environment is affecting the process you are studying. Remember that environmental factors (even relatively invariable ones such as soil chemistry) change sufficiently with time of day or season to affect how organisms react. For some environmental factors such as leaf temperature and light, significant changes occur from minute to minute. To compare the environment measured in different places, measurements must be as nearly simultaneous as possible, or confined to the same time of day during similar weather conditions (for example, early morning during clear summer weather) or season.

Before selecting a method, check to see which methods have been recently used by other scientists for the type of data you wish to collect. You may wish to select an older method, due to the limitations of your budget or of the facilities at the institution where you work. In any case, be sure that results from the method you select are currently accepted by the research community.

Analyzing and presenting your results

While planning your research, consider how you will analyze and present your data. You need to decide early in your project how you will organize your data for statistical analysis. There are two traditional types of statistical analysis. You may plan to test differences between means of different types of sample units (for example, between oak leaves and pine leaves, or between fertilized pine trees and unfertilized pine trees, or between oak forests and pine forests). An alternative type of analysis is to describe a relationship between two variables (for example, light and stomatal conductance of leaves, or soil N concentration and plant growth rate, or soil temperature and relative density of a species in the vegetation). These two alternatives require different research designs and statistical techniques, the first often using *analysis of variance*, the second using *regression*. There are various ways to do each type of analysis; often you must select different methods of computation depending on the degree of normality of your data. Other specialized analyses are required if your data include many data values of zero, or if you sample the same item repeatedly over time. The type of replication you use depends on how you plan to analyze your data. For details, be sure you check with an expert in statistics or consult a book that you understand. Know what you are doing before selecting which parts of a statistical computer program you will use. In addition, there are many mathematical ways to analyze ecological data besides traditional statistics (for example, McCune and Grace (2002) for community studies and those with multiple variables), although many of them do not allow you to test results for significance.

As you develop your research plan, outline potential tables and figures to be filled in using your data. Use graphs and tables to present your data. Give data in the text only when there are too few of one type to make a table or graph effective. Usually, you need not present raw data; if necessary, they can be placed in an appendix, which is often possible in a dissertation. You should, however, try to present your data completely enough that future scientists can use them, or even re-analyze them—consider that your data, if well-collected, are your most important contribution to science. They should be available in a form that will allow them to be used as ecological concepts and methods of analysis change with time.

Usually scientists present mean values from several replications or from sub-samples of the same characteristic; when you do this, also include some measure of the variability among the subsamples (for example, standard error of the mean). When summarizing data, make sure that only data from similar sampling units are included in each mean; for example, it is not useful to compute an overall average for tree density for vegetation plots in a study that included three distinct types of forest. Instead, present the mean for each forest type separately. Means computed using data from different types of sampling units (for example, three kinds of oak forest) do not describe anything that really exists in nature.

Always indicate the units associated with your data (like grams, meters, number per hectare, or percent). Choose units of concentration that are related to a standard unit of volume, mass or area (for example, grams per Liter, trees per hectare) rather than amount per sample or per plot, or some other unit (for example, avoid “trees per 250 m^2” even if you used 250 m^2 plots). In ecology, units are often given per hectare (1 ha = 10 000 m^2) or per m^2. Be sure your description of your variables is complete—if you say, for example, that there are 1500 kg N/ha in the soil, you need to specify the depth of soil included in your calculations. Select units that produce data values including as few zeros as possible—use

2.6 m^2/ha for basal area of a tree species, rather than 0.00026 m^2/m^2 or 26,000 cm^2/ha.

Avoid unjustified decimal places in your data. For example, assume that you measured tree circumference to the nearest cm and then calculated basal area. A change of 1 cm in circumference from 1 to 2 cm produces a change of 0.2 cm^2 in basal area; for larger trees, a 1 cm change produces a larger increment in basal area. There is no reason, therefore, to report basal area to the nearest 0.01 cm^2, even though the table converting circumference to area gives two decimal places. Likewise, you can calculate the appropriate degree of accuracy to use for reporting other types of data.

In order to indicate which statistical tests showed significance, give a *p*-value or place symbols in tables or graphs to indicate the *p*-value, or provide a separate table of *p*-values. Do not present details of statistical tests, like sums of squares. Often, you need to present only the means, degrees of freedom, and level of significance of the difference between means. For relationships between two variables, it is often sufficient to present the direction of the correlation (+ or –), coefficient of determination (R^2), degrees of freedom, and significance of the relationship.

Give units for all data, and repeat column headings on all pages of multi-page tables. On graphs, use a scale such that your data points fill up the space. Try to put different data sets that you wish to compare in the same table or graph, but do not make graphs and tables so complex that they become unclear.

Design headings for tables and legends for figures so that they can be understood without reference to the text. Identify all symbols, and label and give units for all axes of graphs. Place a scale on all photographs and maps, with a north-pointing arrow on maps.

When graphing the relationship between an independent and a dependent variable, place the independent variable on the horizontal axis, and make sure that it has ecological meaning. For example, rather than using a graph of basal area over plot number, the x-axis variable should be some environmental characteristic of each plot that may affect basal area, such as a soil property, temperature, or, if nothing more specific is available, elevation.

Details of data presentation and writing differ among journals and universities. Be sure to follow the "Instructions to Authors" for the outlet to which you plan to submit your report.

Limitations to drawing conclusions in ecology

When planning and interpreting ecological research, consider several properties of nature that restrict the generality and assurance with which you can draw conclusions. These restrictions apply no matter how effectively you carry out a study—they are the unavoidable consequences of choosing to study complex systems with closely interrelated parts, which is the task of ecology.

These properties include:

1. The environment is a complex of many factors; those factors may influence each other. Likewise, a biological system (for example, an organism or ecosystem) is a complex of interdependent functions. An organism reacts to the whole environment and, in turn, can modify the environment. Thus, the influence of an individual factor on an organism is often impossible to identify.
2. The response of a single organism depends on its present and past environments and its genotype. Individuals within most species of organisms vary significantly in genetic properties that affect their ecological behavior.
3. Complex ecological systems like vegetation may result from influences that present-day ecologists cannot measure or even identify. For example, the present species composition may result from past conditions, such as weather during seedling establishment. Often, the most important events were extreme in intensity and short in duration (like a landslide, fire, or disease epidemic); such events are very difficult to measure effectively, even at the time they occur.
4. Correlation between two variables does not mean that one causes the other. If you are tempted to discuss "cause and effect", remember that ecological phenomena are very complex. Any situation may result from a large number of possible causes, and often is determined by a combination of factors. When there is an apparent cause-and-effect relationship, can you identify which factor is the cause and which is the effect? For example, the nature of the soil determines whether a given tree species can grow in an area; yet, the vegetation at the site influences the chemistry of the soil beneath it. Therefore, if you find a correlation between tree species' occurrence and soil properties, you can safely conclude that they may be related, but cannot effectively specify which is the "cause" and which is the "effect".
5. Chance plays a major role in nature. In chemistry, scientists observe the average reaction of a large number of individual molecules, and from that result they propose "laws" that are unaffected by the chance behavior of a single molecule. While studying ecology, we usually observe only a few examples of the units that we study (for example, organisms or ecosystems), so the effect of chance on the behavior of even one unit may modify our results substantially.
6. A chaotic system is one in which a small change in one current condition produces a large change in some future condition of the system. It is very hard to predict the performance of a chaotic system. Some ecologists propose that ecological systems behave in a chaotic manner.
7. The methods that we select to use for a study determine which kinds of data we collect, and thus help to determine the results we achieve. Once we have selected our methods and designed our data sheets, the chances are low that we will discover an important phenomenon or relationship that is not part of the study protocol.
8. The "uncertainty principle" states that you can never be sure how much the act of making a measurement modifies the result that you measure. Correct methods

can reduce such an effect, but never eliminate the uncertainty.

9. Theoretically, there can be more than one effective biological solution to multi-factor requirements (like the need to gain carbon but simultaneously minimize water loss); apparently this is true in nature. That is, organisms with very different combinations of characteristics may perform sufficiently well to persist in a given environment. Theoretically, at least, the number of effective solutions to a problem increases as the number of important environmental factors increases, making it impossible to identify one set of "most-adaptive" characteristics.
10. Many characteristics of organisms appear to be selectively neutral. The amount and type of variation possible in some characteristics are constrained by the nature of the genetic line from which the organism developed. For these reasons, not every plant characteristic may provide an adaptive advantage, and you need not attempt to provide an explanation based on natural selection for each plant property that differs from that of other plants.

How might such limitations affect your research? Scientists seek to discover and to explain phenomena that occur in nature. The limitations listed here mean that determining causation is difficult and seldom completely possible; however, it remains a basic goal of research. You can perform useful and important research by working toward a causal explanation without expecting to find it, and you can complete a worthwhile project without identifying "cause and effect" relationships.

Using the literature

Knowing the work that others have done about your topic and in related areas is critical throughout your research effort, from planning, to doing the work, to interpreting the data, and, finally, to writing the report. Giving credit for what you learned from others is required in science. Every specific idea or fact that you derived from someone else's work should have its source indicated in your report. To fail to acknowledge the source of your information or ideas is a serious error. In the text where you present an item that you borrowed, give the author and date of publication of its source—for example, "*Quercus semecarpifolia* dominated the high elevation forests (Shrestha 1981)." If you directly copy text from another publication, enclose it in quotation marks and add the page number to the reference in the text, for example, " '*Quercus semecarpifolia* dominates above 2500 m' (Shrestha 1981, p 33)".

All citations given in the text must be listed completely in the Literature Cited section of the report, and all references listed in that section must be cited in the text. There are several acceptable ways in which information about a reference can be arranged—see instructions to authors of a good journal, like this one, for details. No matter what format you use, each reference to a journal article must include the family name and initials of all authors; year of publication; complete title of the article; name of the journal; the volume number of the journal; and the numbers of the pages on which the article begins and ends. There are many other types of publications, each with its own accepted format for a reference. Once you choose a format for references, follow it consistently.

Preparing a clear, accurate report

Writing is a very personalized skill. Few persons can easily write a clear, concise, error-free summary of a complex technical subject. Most persons can do this only after long practice and repeated critical review, and never without substantial effort—but most also can learn the skills necessary for successful writing, and everyone can improve with practice. Preparing a report about your results allows you to complete your project, but often requires much effort for a long time; plan for a long enough time for writing to do your report well.

The "style" of writing is the most difficult aspect to learn and to teach, but it is critical to producing an effective report. Clear organization and sentence structure are required for a successful style. Using words that are precise but widely understood is also important. An effective style produces a manuscript that is easy and interesting to read, easy to understand, and hard to misinterpret. Using an effective style is important. Good data and conclusions in an ineffectively written report usually have little effect on science and bring little recognition to the author.

Other details of writing are easier to learn than style and are also very important: for example, spelling, punctuation, capitalization and abbreviation. Strive for a report free of errors in all areas.

Proofreading means to read the manuscript carefully in order to find and correct errors. Proofreading serves several purposes, and the same reading will not suffice for all types of errors. The manuscript should be proofread first for errors in concepts and conclusions, again for errors in style and clarity of the writing, and finally for technical errors in spelling, data values, and other important details. You must proofread and correct errors in a manuscript each time before it is presented to a different type of reader—your co-authors; your colleagues who agree to review it before you submit the manuscript; the journal editors to whom you submit both the first version and (if you are successful) the revised manuscript. This is required as a courtesy to the reader and to assure that your manuscript will be judged on its scientific merit, rather than be rejected because it had many small, seemingly innocuous errors.

Proofreading is necessary to gain credibility. If you do not, for example, copy names of authors and organisms correctly, how can you expect the reader to believe that you have done a complex, difficult scientific study without error? When tempted to skip proofreading because you think there are not very many errors in the manuscript, remember "Murphy's Law"—the chances of an error remaining in a manuscript increase with the degree of embarrassment the error will cause the author. Few proofreading jobs are perfect, but there is no excuse for not trying to reach perfection. Proofreading should include all parts of the manuscript: text, literature citations, figures, and tables. Errors in numbers, especially in tables, and in scientific names, are especially easy to miss, and these should be checked with special care.

Are you ready to start?

The task of doing research is complex, sometimes mystifying. Some of my suggestions should help ease your work. Perhaps even more useful is for you to read and listen and talk about science. Read and read and read articles and books by scientists in your area. Learn how they did and interpreted their work, then use their ways if they fit your project.

Not all scientists work the same way. Many will disagree with some of what I wrote here and a few, with much of it. With practice you will find how you work best, but maybe this essay will give you a head start in a useful direction.

Ecological research has been a joy for me and I hope your research will be for you. I find continuing enjoyment when I see things in nature that I did not know were there, when I see the data accumulate and the patterns form, and when I can use the results of others to find ways to understand my results. And sometimes I find something completely new, whether I was seeking that specific thing or not, and then all the strain and effort cease to matter. For me, it remains a thrill to learn that my work has been accepted by other scientists and will be published and available to any scientist who can use it. I wish you the same satisfaction: when your work becomes available to others, you have completed the task of a scientific research worker.

Acknowledgements

This essay is derived from a manual first published in Nepal (Zobel 1985), with additions from teaching materials prepared at Oregon State University. I appreciate the many helpful comments of Pat Muir and John Metz.

References

Chamberlin TC. 1965 (1890). The method of multiple working hypotheses. *Science* **148**: 754–759

Grime JP. 1979. *Plant strategies and vegetation processes*. Chichester: Wiley. 222 p

Grubb PJ. 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews* **52**: 107–145

Harper JL. 1977. *Population biology of plants*. London: Academic. 892 p

Harper JL. 1982. After description. In: Newman EI (editor), *The plant community as a working mechanism*, Special Publication 1, British Ecological Society. Oxford: Blackwell Scientific. p 11–25

Karban R and M Huntzinger. 2006. *How to do ecology. A concise handbook*. Princeton: Princeton University Press. 145 p

McCune B and JB Grace. 2002. *Analysis of ecological communities*. Gleneden Beach, OR: MjM Software Design. 300 p

Peters RH. 1991. *A critique for ecology*. Cambridge: Cambridge University Press. 366 p

Weiner J. 1995. On the practice of ecology. *Journal of Ecology* **83**: 153–158

Zobel DB. 1985. *Preparing the Master's dissertation: a manual for students and teachers*. Kathmandu: Botany Instruction Committee, Tribhuvan University, Kirtipur campus. 21 p

Zobel DB, PK Jha, MJ Behan, and UKR Yadav. 1987. *A practical manual for Ecology*. Kathmandu: Ratna Book Distributors. 149 p