1

Preparation

For many students, honing their research skills is an important component of their academic development. However, inexperienced researchers can be naïve in their approach, and may even attempt highly complicated studies that have little chance of being completed in the time, and with the resources, available. This chapter describes the thought and preparation needed to plan your project, particularly how to formulate your ideas into something structured and workable before going out into the field. In your research, you will search for explanations or patterns, make comparisons, predictions and generalisations, and formulate theories. Research is not simply an exercise in information gathering. Rather, research is about asking questions that go beyond description and require analysis. Your research will be highly individual, and there are no set outcomes. You will form your own opinion, even if this disagrees with previous work. This is because progress in science results from the continual testing, review, and criticism of other researchers' work. Do not expect your research project to answer all your original questions. It is much more typical to find that research generates more questions than it answers. Research submitted for publication or for examination should show evidence of originality. Even if your research is not wholly original, it can show evidence of original thinking. Although the prospect of carrying out original research may seem rather daunting, providing you do not exactly copy someone else's experimental design, methods, sites, etc., your research is almost certainly going to be original. There are several ways in which work can be original:

- Executing an entirely new piece of work (e.g. studying a plant or animal for which there is little or no information currently available).
- Adding knowledge in a way that has not been previously done before. Many empirical studies do not develop new topics to study but instead advance their work with the use of original experimental designs, new statistical methods, etc. For example, new insights might be generated from exploring the ecology of an otherwise well-studied animal at different sites to see whether a behaviour or food preference differs between locations.
- By testing somebody else's idea, or by carrying out an established idea in a new area, new experimental subject, etc., or by using existing data to develop new interpretations.
- Continuing an existing piece of work that is ongoing at your university or with a partner institution. For example, there are many long-term experiments that invite

Practical Field Ecology, Second Edition. C. Philip Wheater, Penny A. Cook and James R. Bell.

Companion Website: www.wiley.com/go/wheater/practicalfieldecology

students to participate in summer work. These opportunities can be symbiotic and provide both you and the scientist running the project with more data that could elucidate a mechanism or generate new hypotheses.

- Originality may only be apparent in the breath of the study. Increasingly popular is 'cross-disciplinary' science where, for example, soil scientists, botanists, and entomologists converge on a subject matter or site and work together to test an over-arching hypothesis.

All research, whether taking place in the laboratory or field, needs careful planning. It is perhaps self-evident that such planning should involve the correct use of equipment and choice of appropriate sampling methods and collection sites. In addition, a wide range of associated logistic, legal, and health and safety implications need to be considered. Although many of these issues are equally important in field or laboratory-based investigations, field research may be more limited by time and other factors (access to sites, time of year, weather conditions) than is laboratory research. Thus, field study may need more careful consideration prior to implementation. Chapter 1 details some of the issues involved in planning and designing fieldwork, and culminates in a checklist that may help to prevent problems once research is implemented. Chapter 2 deals with the techniques required for monitoring sampling sites and measuring physical and chemical factors. Chapter 3 covers the methods used to sample static or relatively immobile organisms, whilst Chapter 4 extends this concept to studying mobile animals. The latter includes a consideration of monitoring behaviour and of dealing with both direct and indirect observations, as well as covering the trapping and marking of individual animals. In Chapter 5 we summarise a large number of different approaches suitable for the statistical analysis of ecological data. Finally, in Chapter 6 we cover how to present your results and produce appropriate reports, posters, and presentations.

Choosing a topic for study

The first stage of a research project is choosing a subject area in which to research (see Box 1.1 for a list of some texts that include ecological project ideas). As you will be devoting substantial time to your project, it is important to choose a topic that interests you. You may also wish to make your research relevant to your current or future employment. A variety of organisations – local, national, and international – may be able to help you identify an area of interest that is also of current relevance. Such organisations include local authorities and wildlife trusts, nature reserves, museums, and (inter)national bodies such as the RSPB, Plantlife, and WWF (Naturenet provides a list of some such organisations that may be able to help).[1] Pick a topic of the right size: neither too big nor too small – typically the fieldwork component of most undergraduate projects are less than 6 weeks in length. Looking at successful previous projects may assist you in judging how much can be done in the available time (ask more experienced researchers for examples of good projects to look at). Finally, your proposed project has to be feasible; for example, in terms of equipment, access to sites, and realistic timescales. Once you have selected your subject and provisional title, be prepared to be flexible and, if necessary, to change direction. This may happen for a variety of reasons; for example, if a pilot study reveals a more interesting avenue for research, or if your original ideas turn out to be unfeasible. You should note that the planning process should involve a consideration of the

1 http://naturenet.net/orgs

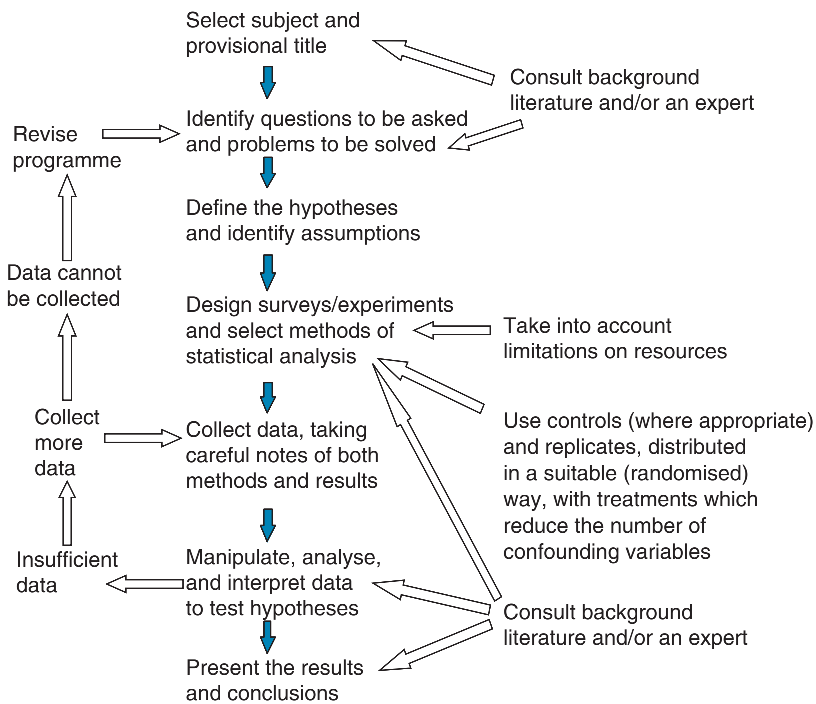


Figure 1.1 **Flowchart of the planning considerations for research projects.**

whole project to enable you to identify and deal with any potential problems before they become major issues (see Figure 1.1). In all aspects, reading around the subject will allow you to use appropriate techniques, build on existing knowledge, and avoid reinventing the wheel. Inevitably, there can be logistical problems that influence your choice of site, or species, or otherwise prevent you from proceeding exactly as you would have wished. Although you can avoid many such problems by careful planning, there are some aspects that you will not think about until you implement the research. A pilot study will help to identify such issues and may allow you to refine the study in advance of full implementation. If you are unable to do a pilot study, at least try to explain your approach to someone who has spent time doing ecological fieldwork. They may highlight some obvious deficiencies. This is particularly useful for field trips abroad, which preclude a pilot and can be fraught with difficulty.

Box 1.1 Some sources of ecology projects

There are a number of journal papers that list topics of interest in ecology and associated fields (e.g. Sutherland et al. 2006; Sutherland et al. 2009; Pretty et al. 2010; Sutherland et al. 2013). Even if the projects themselves are not of direct interest, they may inspire other project work since there is plenty of debate stemming from these papers on a variety of blogs and forums.

Natural England has a list of research ideas on their national nature reserves https://www.gov.uk/government/publications/research-opportunities-on-national-nature-reserves-in-england

There are many other resources that give examples of feasible ecological research projects. The series listed below are examples of some of those that cover either a wide range of habitat types, or a range of organisms.

- The Practical Ecology Series provide project ideas associated with grasslands (Brodie 1985), freshwaters (Gee 1986), the seashore (Jenkins 1983), and urban areas (Smith 1984).
- Routledge Habitat Guides each include a section (section 5) giving project ideas for the habitats associated with grasslands (Price 2003), uplands (Fielding and Howarth 1999), urban habitats (Wheater 1999), and woodlands (Read and Frater 1999).
- The Biology of Habitats Series covers a wide range of global habitat types and includes practical aspects of working within the habitat and some of the types of studies that are possible within them (e.g. Rydin and Jeglum 2013 for peatlands and Little et al. 2009 for rocky shores).
- The Naturalists' Handbook Series (now published by Pelagic Publishing) contains many ideas related to studying a group of species (e.g. Gilbert 2015 on hoverflies or Roy et al. 2013 on ladybirds), or different habitats (e.g. Hayward 2015 on sandy shores or Wheater and Read 1996 on animals under logs and stones), or implementing different techniques (e.g. Unwin and Corbet 1991 look at insects and microclimate, whilst Richardson 1992 examines pollution monitoring using lichens).

Ecological research questions

Having decided on a provisional topic, the next step in the successful planning of any research project is to identify those questions you wish to ask and then to formulate the aims and objectives. There are various reasons for researching particular plants, animals, or environments and this section provides a quick overview of the scope of ecological projects.

Many studies involve monitoring the number of species, number of individuals (relative abundance), estimates of population size or density (absolute abundance), or community structure (diversity, evenness, and richness). Additionally, studies on animals may require observations of the behaviour of individuals or groups and their interactions with each other and their environment.

Monitoring individual species and groups of species

Sometimes, projects may be targeted at a single species (i.e. an autecological study). For example, where a species is judged to be important for positive reasons (including its conservation or commercial value) you might require information about its distribution, population size and dynamics, age structure, behaviour, etc. Where there is a more negative view of a species (e.g. because it spreads disease, competes with native fauna and flora, or is an invasive species that dominates a habitat to the exclusion of other species), you may need information about its distribution, dispersal ability, vulnerability to disturbance and predation, etc. A biogeographical study that might be of interest is the examination of species' distributions where species are expanding or contracting their ranges – perhaps as a result of climate change or other factors (either natural or human-influenced, e.g. habitat disturbance and fragmentation). Conversely, you may be interested in groups of organisms (i.e. community ecology or synecology), examining the diversity of communities, the interrelationships between plants and animals in protected areas, or in establishing ecosystem function in relation to environmental legislation (e.g. the EU Water Framework Directive).[2] Studies spanning a wide range of different taxa can be particularly valuable in understanding complex environmental systems, although they may be difficult to implement and the subsequent analysis and interpretation of the results can be complicated.

Monitoring species richness

In studies examining species richness, you might be interested in the presence or absence of one or more species (or other taxonomic group) in order to investigate the links between such species and aspects of the environment; for example, in terms of the ecology of the species concerned, or in studies of pollution where the species may be useful as a biological indicator of certain toxins. Here, simply listing the plants and animals present may suffice. Although this may appear to be a quite simple approach, care needs to be taken to ensure that sampling techniques are used that are appropriate to both the organisms under consideration and the habitats in which they are found. For example, studies on bird species richness in urban parks may be complicated if some parks are dominated by relatively open habitats of amenity grassland and formal flower gardens, whilst others feature dense shrubberies and even woodland. Observations of the species present may be easier in the open habitats than under dense canopy. Care will thus be needed to ensure that all species are counted (as accurately as possible) using the most appropriate method for that site or habitat type. For these reasons, issues around surveying habitats and sampling organisms are considered in the next three chapters.

Monitoring population sizes and density

In population and density studies, it is the number of plants or animals of particular species that is of importance. Such studies may look at the number per unit area (i.e. the density) or calculate estimates of population sizes. Densities are taken from the estimated population size divided by the size of the area sampled. However, for mobile organisms it may be difficult to identify the spatial limits of the population (e.g. in studies of butterflies, some species may be highly mobile with individuals not being restricted to defined small sites). Under such circumstances, densities may be less useful than estimations of population sizes of the animals using particular sites. If populations of several species are being studied, then it is important to ensure that the sampling methods used are appropriate to all the species being monitored. For example, in rainforests, some species of butterflies are found mainly within the canopy and are only occasionally caught at ground level and, conversely, some are predominant at ground level. Clearly, any survey comparing such study sites should incorporate sampling at both levels.

The spatial distribution of organisms can also be of interest, whether that is of static organisms (see Chapter 3) or more mobile individuals, populations, or communities. In the latter case, movement obviously complicates the monitoring techniques. It is often the case that individuals (and indeed populations) are restricted to defined areas either

2 http://jncc.defra.gov.uk/page-1375

because of limits to their ability to disperse and colonise, or because of resource limitations (including competition between members of their own or other species). Biogeographical studies can provide information about distributions in geographic space and over time, whilst a landscape ecology study would look at spatial patterns over a range of scales in relation to landscape function and processes. At the smallest scale, individual home ranges and territories can be studied. Plotting the movements of animals diurnally and seasonally can indicate how individuals and groups of individuals make use of the environment in which they live.

Monitoring community structure

Another type of study of interest would be to establish the structure of the community of a specified area or habitat type (e.g. the community of fish in a lake, or the community of insects inhabiting a certain species of tree). Such studies may involve sampling a large range of quite different organisms. Organisms differ in size, distribution (both spatially and temporally), their use of microhabitats and, in the case of many animals, mobility. As such, care needs to be taken to ensure that the methods are as comprehensive as possible and are not biased towards or against any particular species or groups of species. For example, sieving soil to examine the communities of animals living within different layers (leaf litter, humus layer, the 'A' horizon of the soil, etc.) may underestimate larger animals that are found at low densities (e.g. large ground beetles), and may overestimate species that are found in large aggregations if sampling happens to coincide with these groupings (e.g. some woodlice). Several different techniques may need to be used together during a single study in order to obtain a broad understanding of the community structure of such habitats.

Monitoring behaviour

Studies on animals may involve monitoring the behaviour of individuals, even if this is not the primary purpose of the study. Knowing whether rabbits are feeding, being vigilant for predators, etc. may be useful if numbers are being counted in particular sites. Of course, other research projects will focus primarily on animal behaviour. Such behavioural studies may involve the observation of a number of individual animals in a variety of settings, or the interactions that animals have with others of the same, and/or different, species. Activity levels and habitat/resource utilisation may change with time of day and season according to changes in the weather. Assessing the time that individuals devote to different behaviours (time budgets) can help to identify such changes. It is essential that the location and methods used by the observer do not influence the behaviours being monitored. Working too close to large mammals with young may mean that the major behaviour monitored is vigilance directed against the observer – and may be dangerous for the researcher. Behavioural changes may occur even with subtle changes to the environment, such as the use of head torches. The emitted wavelength of white light attracts flying insects like moths, but changing the torch to red light solves the problem. What this does show is that all actions in the field, however apparently trivial, need careful consideration before fieldwork begins.

A note of caution

Whilst focusing in on the main aim of the research will help you to formulate the procedure to be followed, you will also need to understand the limitations of the approach that you take. For example, census methods (such as simple species counts) can be quick to implement and provide substantial amounts of data in a short time. In contrast, techniques to assess population sizes or community structure tend to be much more time consuming and may produce complex data sets. However, you should be aware that whilst it is usually possible to extract census information from population or community study data sets (albeit with a loss of detailed information), it is not possible to use census methods to assess community structure or population levels. In general, it is important to have at least some knowledge about the ecology and behaviour of the species or community under investigation when designing the research project, irrespective of the type of study being undertaken. Understanding the limitations as well as the potential benefits of any technique employed is essential to being able to critically evaluate the data gathered. As such, refining and developing new and existing methods can avoid or, at the very least reduce, biases and other problems with particular techniques. Case Study 1.1 describes one particular example of how a novel technique was developed.

Case Study 1.1 The development of a novel net for sampling bats emerging from tree roosts

Henry Andrews is an ecological consultant and the founding member of the Bat Tree Habitat Key (BTHK) project (http://battreehabitatkey.co.uk). The BTHK project was set up to achieve several objectives: (1) to establish the full range of different features of trees used by bats for roosting; (2) to understand which species used which features, when they used them, and how they used them; and (3) to build an app to help fieldworkers assess which species they might encounter in potential tree roost features. This case study describes the development of a new net by Henry Andrews of the Bat Tree Habitat Key project and NHBS Ltd to allow the user to sample bats emerging from tree roosts safely. This work involved a number of contributors: Henry Andrews, Katharine Clayton, Oliver Haines, Thomas Hamilton Koch, and Steaphan Hazell.

The BTHK Tree-Roost Net is supplied by NHBS Ltd (https://www.nhbs.com), a company that manufactures standard and bespoke marine, freshwater, and terrestrial survey tools.

Photo HA

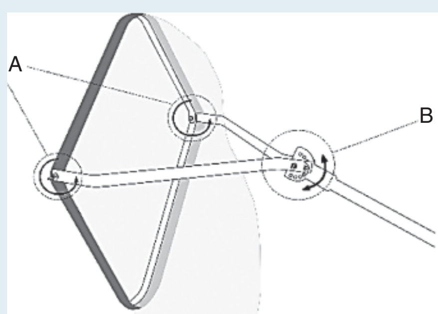


Image BTHKP

BTHK Tree-roost Net *in situ*. A schematic showing the pivot points to adapt the BTHK Tree-roost Net ensuring optimum positioning against the tree.

Model organism and research challenges faced

Of the 18 species of bat recorded in the UK, 14 are known to roost in trees (Bat Tree Habitat Key Project 2018). Little is known about inter or intraspecific differences in the use of tree features for roosting, or the extent to which different species use trees in different habitats or during different seasons. To fill this gap in our knowledge, the BTHK project collects and analyses records of tree roosting behaviour submitted by bat researchers, ecological consultants, arboriculturalists, and bat conservation workers. One problem identified by the BTHK project is the inadequacy of current methods to sample bats roosting in trees. Endoscopes are frequently used but provide limited data, because it is very difficult to identify bats to species level using this method (let alone record sex ratios!). Furthermore, bats frequently cluster in narrow crevices, making observation of more than one or two bats almost impossible. An alternative method involves sampling bats with a handheld net as they exit the roost, but no safe, reliable, purpose-built net existed.

How the challenge was resolved

The idea for a bat tree-roost net originated with Henry Andrews and NHBS helped develop the design and manufacturing process. The end-product (the BTHK Tree-Roost Net), employs a kite-shaped head that pivots in two places to ensure optimum positioning of the frame against the tree. The aperture of the net head can be adjusted to fit a wide range of tree sizes (marked A on the figure) and the net head pivots according to the height and location of the roost entrance (marked B on the figure). These features mean that the net can be adapted to sample roosts at a variety of heights and orientations, safely and securely. The net handle is made from aluminium, which is both strong and lightweight, and at 4 m long it can either be placed on the ground or braced against the users' body to provide a stable foundation.

The collection bag is made from soft white nylon mesh to prevent damage to the bats wings and feet and enable the user to see the bats more easily against the white background. The collection bag also has a flexible plastic attachment that extends both into and away from the aperture of the net. This allows the outside of the collection bag to form around the contours of the tree, and acts like a funnel trap to prevent bats escaping from the inside of the collection bag.

Once the net has been set up, a thermal imaging camera can be used to monitor the contents of the collection bag. To minimise disturbance, Henry Andrews recommends that two to three bats are sampled from each roost, as in most cases this will provide sufficient information about the composition of the roost.

Advice for students wanting to work with bats

Bats and their roosts receive full legal protection throughout Europe and should never be disturbed without a license to do so. If you are interested in getting involved in bat research, you could volunteer with your local Bat Conservation Trust group (www.bats.org.uk) who will almost certainly provide opportunities for you get involved in bat surveys. You could also contact academics or ecological consultants that study bats to find out whether you could assist them with any of their work.

Creating aims, objectives, and hypotheses

Once a topic for research has been chosen, you can work out the aims of the study. These are important, since tightly defining the aims helps to focus more clearly on the work in hand and can avoid problems in implementation. 'Woolly' aims, such as 'to investigate invertebrates under logs', may be a starting point for a more focused aim, such as 'to determine whether the number of invertebrates found under logs is related to the size of the log'. This then leads to further questions, including:

- Which invertebrates are to be examined, i.e. should they be identified to species, or merely counted *en masse*, or allocated to ecological functional groups (e.g. predators, herbivores, etc.)?
- What is a log (i.e. when is a fallen piece of wood a log rather than a twig?) and how many logs should be investigated?
- How should we standardise or otherwise account for the condition and type of the logs (degree of decomposition, species of tree, etc.)?
- Which measurements of size should be incorporated (e.g. length, width, surface area touching the ground, volume, depth of log in the soil, presence of other organisms such as fungi, etc.)?
- Where should we sample the logs?
- Which statistical method(s) should we use to analyse the data?

Once these questions have been answered, they become objectives that can be used to determine your methods. The aims and objectives lead us to the setting up of working hypotheses. For example, in our study of possible relationships between log size and the numbers of invertebrates found beneath them, we would set up a statistical hypothesis to be tested. It is common practice that the statistical hypothesis to be tested is a null hypothesis; in this context that 'there is no relationship between log size and the number of invertebrate animals found underneath them'. Most univariate statistical tests examine the likelihood of the null hypothesis being true (see Chapter 5). A null hypothesis should meet the following criteria:

- be a single, clear and testable statement – where more complex research questions are asked, you should break these hypotheses down into individual statements that are treated separately and tested in turn;
- have an outcome, typically either 'accept' or 'reject' the null hypothesis;
- be readily understandable to someone who is not a scientist.

Reviewing the literature

You should always review the planning and implementation of each stage of your research project by using current information, either from others who have been involved in similar research, or using texts, papers in journals, or other information sources (e.g. the internet), or a combination of these. Be aware of possible biases in the information used, especially where this is obtained from websites belonging to individuals (rather than respected organisations) that have not been independently validated. Most papers in reputable journals and many textbooks have been examined by independent referees, although even these may contain factual inaccuracies and personal opinions that may not conform to current opinion. Although considered the gold-standard of information sources, even peer-reviewed journals are subject to bias against the publication of negative results. It is important to start your review of the literature as early as possible, since it is an ongoing process throughout your research and should inform each stage of your project. At the very least you should begin by reading the literature to establish that your proposed idea has not been already published and to define the gaps in knowledge that you will attempt to fill. It is likely that as you read one paper, you will find references to other work that may be important. Expect to read more than 100 papers in your field and to actively use a third to half of these in your thesis.

If you are new to a subject matter, you should first try and locate seminal piece(s) of work in the field or a recent review. Typically, this will be close to the top of a search list of highly cited papers and can be found by ordering a search by 'times cited'. Take a detailed look at the seminal paper(s), the reference list, and who is citing that paper. In journal databases (e.g. Web of Knowledge), citation networks can be viewed to examine the connectedness between a seminal paper and all those papers that cite it. This is useful because it can elucidate key papers in the field and reduce the search effort dramatically. Typically, your first search should include seminal works and a collection of the most recent papers in the field (i.e. from the last few years). It might be helpful to order these by journal impact factor (if available), since parochial journals may not contain as high quality science, although sometimes smaller research papers with less apparent impact can provide valuable information in the form of species lists, new methods, and negative findings that are often not reported in more mainstream journals. An additional word of warning: highly cited papers can also be poor papers in the field since other authors might simply be referencing them to make an example of that piece of work (e.g. 'Black and White's (2000) experimental design has been shown here and by others to be flawed'). Knowledge of the literature can assist in avoiding 'blind alleys' and unfruitful lines of enquiry or techniques. There are two main types of literature: primary and secondary.

Primary literature

This is first-hand information; for example, articles in specialist journals, reports, MSc and PhD theses by those who have done the work and generated data. Journals that publish only refereed papers (i.e. those that have been through a peer-review process) are the most important sources of primary, up-to-date information, and where possible your literature review should focus on this type of source. Other primary sources include technical reports, management plans, consultancy reports, and species lists (e.g. from annual recorder reports). All of these can be useful sources of information for ecological projects, but you should be aware that they may not have been edited or their quality controlled.

Secondary literature

This is prepared from other sources of information – including textbooks, review articles, opinion pieces, and letters to the editor – which often express views on primary research. If you are lucky, there may be specialised books covering your subject area.

These may provide a good starting point since books are secondary sources of information, while journals are a preferred source of reference for most research projects. The coverage in student textbooks is usually rather more superficial than that in specialist texts. If there are review articles on your subject, these may be useful to obtain an overview and as a source of new references. Reviews are found in edited book sections, journals that specialise in reviews (e.g. Trends in Ecology and Evolution)[3] and journals that have occasional review papers. Review papers in established scientific journals have usually been subject to peer review.

Other sources of information

Maps, personal notes, databases, museum collections, and archives may all contain information that can be useful in supporting your research. There are several software systems used within professional ecology that can be useful for research projects. These can help to set your work into a context grounded in practical conservation issues, as well as supplying data on either a wider spatial or temporal basis. One such package, the Conservation Management System (CMS) software,[4] is used by many countryside managers to produce management plans, whilst Recorder[5] is the software used by the UK National Biodiversity Network[6] to record, manipulate, and map biological records. **Note:** software is often subject to changes in version and may indeed stop being supported, either by the producers or by computer operating systems, over time.

Search terms

The best way to locate relevant papers is to systematically search journals using your library's electronic databases using key words. The use of targeted search terms can identify appropriate works and avoid too many articles or irrelevant papers from being selected. First, identify a list of key words or phrases that could be included in the title, abstract, or key words of an article. Begin by using simple combinations of terms (e.g. woodland beetles, hedgerow birds). If you are finding too many papers, then either restrict your search to more recent publications or use more complicated combinations (e.g. by adding 'predation' to the key terms 'woodland beetles' you can significantly reduce the number of articles returned by a search). Note that most databases allow wildcard entries (e.g. an asterisk) to truncate key works (so that space, spaces, and spatial can all be covered by spa*), although such terms can increase irrelevant returns (e.g. spa* will also cover any word beginning with spa; e.g. Spain, Spanish, etc.). You can use Boolean operators to refine your search and reduce the number of irrelevant works identified. For example, the search 'woodland beetles' OR 'hedgerow birds' will select those papers which have either 'woodland beetles' or 'hedgerow birds' in the title, abstract, or key words. Conversely, the search 'beetles' AND 'predation' will select only papers with both terms in them. Once appropriate papers have been acquired, you can use them to find others of interest in two ways. First, you can examine the literature that the paper cites to see if any older works are relevant. Second, you can use most search engines (e.g.

3 http://www.cell.com/trends/ecology-evolution/home
4 https://www.software4conservation.com
5 http://jncc.defra.gov.uk/page-4592
6 https://nbn.org.uk

Google Scholar) to identify those works that come after the paper in question and have cited it. If you have access to a librarian, they can help you to set up a search. This search can be saved and re-run during your project so that you can detect any new papers published whilst you have been carrying out your research.

Reading papers

You will find more source papers than you have time to read. It is easy to get bogged down in the wealth of published material. Keep your subject area in mind and do not read everything indiscriminately. Skim-read a new reference to decide how much attention it deserves. Start by reading the abstract, skimming the subheadings and then the first paragraph or so of the introduction and the last paragraph of the discussion. Only read in detail those papers that are particularly relevant.

Keep a copy (photocopy, scanned image, or electronic copy) of key references and make notes of (or highlight on photocopies) any useful information. Save the full reference, since all material cited in a research report must be listed in full in the reference list; there are few things as annoying as having to re-find the details of a reference that you read much earlier. For journals, the full reference includes the authors' names and initials, year, article title, journal title, volume number, and page numbers. For the other types of references, see the guidelines for the reference list in Chapter 6. For books, in addition to noting the authors, publishers etc., take down the library classification number for your own reference in case you need to return to it at a later date. Increasingly, many papers are available online. The level of access depends on the services subscribed to by your organisation. All libraries offer an interlibrary loan service to provide access to works that are difficult to get. When using such a service, note the date on which you request any interlibrary loans to help to keep track of your requests. If you have access to reference management software (such as EndNote,[7] Mendeley,[8] or Zotero[9]) then it is worth using this since it will make creating your reference list easier when you finalise your report.

Be critical as you read; do not accept everything as true just because it is published. Look at the evidence and decide whether the conclusions are justified, or whether the results could be interpreted differently. It is, unfortunately, not uncommon for assertions to be made with no supporting evidence. You will find that different authors in the same field may disagree. It is particularly important to distinguish opinions and speculation from evidence. You can make your own interpretations and conclusions from the work of others and cite them using expressions such as 'an alternative explanation for the results of Green and Brown (2010) is that …'. Read critically and keep your use of the information relevant for constructing your own account.

Practical considerations

Research, especially field research, can be an unpredictable business. However, with careful thought it should be possible to ensure that most eventualities are covered. Legal

7 http://endnote.com
8 https://www.mendeley.com
9 https://www.zotero.org

and ethical aspects – for example, access rights to land and the impacts on protected species or habitats – should be top of your list of practicalities to consider. Health and safety is another obvious concern and it is essential that you ensure that there is no danger either to you or to those around you through carrying out an appropriate risk assessment. In addition, practical approaches – such as effective time management, efficient data recording and security, and the appropriate use of equipment and techniques – will also help you to deliver a research project successfully.

Legal aspects

When planning fieldwork, it is important to take into account your responsibilities and any legal implications of the work. At an early stage in the planning of the project, always seek permission to work on a site from the landowner and any other interested parties (for example, relevant statutory bodies if the sites have some form of special protection, e.g. SSSI or National Nature Reserve). Keep disturbance to a minimum and remove as few plants or animals as possible. Identify specimens *in situ* (if you can) so that they need not be removed from the habitat. Whole plants should not be taken without the express permission of the landowner. It is good practice in the field not to pick plants of any kind, unless absolutely necessary. In many countries there are a number of protected species (for example orchids) that should not be uprooted, picked, or harmed in any way. You should check with the appropriate governing organisation for the country involved for details on protected species. In some countries there is specific legislation covering protected species (for example, the Endangered Species Act 1971 in the USA,[10] the EU Birds Directive 2009,[11] and the Wildlife and Countryside Act 1981 in the UK[12]).

Some animals may not be disturbed or handled without a permit (e.g. birds and bats amongst others in the UK), and this extends to some microhabitats (e.g. badger setts in the UK). Rare animals and plants are often protected by law and a licence may be required to handle or disturb them, without which you could face a large fine. You should also consider the ethical aspects of your study. This is particularly important where animals, especially vertebrates, may be harmed. Under these circumstances appropriate licensing authorities should be consulted. **Note:** this handbook is a guide that does not definitively outline the legal position or interpretation of any act or regulation. In all cases of protected species, it is the responsibility of the researcher to check with the relevant bodies to understand what guidelines and regulations are in force. Legal frameworks can change with time and being unaware of the legality of your actions is not a valid excuse. Therefore, it is important to stay up to date with any legislation pertaining to your area of work.

Ethical issues

There are a number of ethical issues that should be considered before engaging on any study, even if there are no legal restrictions on the sampling regime proposed. Where any damage to a habitat or any destructive sampling or organisms is to be attempted,

10 https://www.epa.gov/laws-regulations/summary-endangered-species-act

11 http://ec.europa.eu/environment/nature/legislation/birdsdirective/index_en.htm

12 It is important to ensure that any changes to such acts are taken on board; for example, amendments to the Wildlife and Countryside Act 1981 can be found at www.legislation.gov.uk/ukpga/1981/69

there needs to be a proper evaluation of the balance of costs and benefits. For example, it is possible through handling animals and plants to vector a disease both to animals and humans and thereby increase the risk of spread. Frogs, for example, carry salmonella which can infect humans who come into contact with their faecal matter, but humans can also vector pathogenic fungi to frogs, endangering these amphibians. Large-scale depletion of populations of plants or animals should not be considered unless the data being generated are of significant value, or the destruction will occur anyway because of another impact (e.g. fogging trees to remove insects where the trees will be clear felled in the very near future – see p. 185). Similarly, where there is a risk of harming non-target species, the methods should be modified to ensure that these are unaffected. For example, when trapping for small mammals or invertebrates, there is a danger that animals such as shrews can be trapped and killed; modifications to the techniques can avoid this (see p. 266). Such problems can sometimes become apparent only once pilot studies or even the main study has begun. Under such circumstances, the methods should be modified or the study abandoned, particularly in cases where species of conservation importance are under threat. In one such example, whilst pitfall trapping for invertebrates, researchers inadvertently captured a number of great crested newts from wet grassland. Since it would be difficult to exclude the newts from the traps, trapping at this particular site was abandoned (Cullen 1995). When animals are collected alive, they should only be released into the environment from which they were taken to avoid adversely impacting other habitats. Similarly, plant material and soils should not be disposed of in areas where they might cause ecological issues (e.g. biological or heavy metal contamination). Costello et al. (2016) provide a checklist of ethical aspects that should be thought through before, during, and after implementing any field project. Ethical aspects of working on animals are covered by Reed and Jennings (2007). Where relevant, your work may need to be examined through a local ethical review process, often captured in a risk assessment. Most universities and similar organisations operate these, and further details can be found in RSPCA and LASA (2010). It is also worth reading the discussions by Minteer and Collins (2005a, 2005b, and 2008) and Parris et al. (2010) who comment on fieldwork activities.

Health and safety issues

Look after your own health and safety and that of those around you. All investigations should be assessed for any risks, including those caused by the terrain, the techniques, and any sudden changes in weather (see Barrow 2004). Any chemicals being used should be checked against appropriate regulations and risk assessments should be produced to identify safe use, disposal, and how to deal with spillage and accidents. In the UK, such regulations (COSHH – Control of Substances Hazardous to Health)[13] are covered by the Health and Safety Executive (HSE). Many organisations have their own health and safety guidelines; in the absence of these, advice is available in Nichols (1999), Winser (2004), Aldiss (2007), USHA/UCEA (2011),[14] and Daniels and Lavallee (2014). It is advisable for at least some participants in any field project to be trained in emergency first aid, especially when working in remote environments.

13 www.hse.gov.uk/coshh
14 www.usha.org.uk or https://uceastorage.blob.core.windows.net/ucea/download.cfm/docid/responsible_research_-_managing_h_s_in_research.pdf

Although all environments are potentially hazardous, you should pay particular attention when working at the coast (especially with regard to tides and hazards including quicksand), rivers (above all with regard to potential flash floods), the uplands and mountain areas (especially with regard to sudden changes in weather conditions and the risk of exposure), and in situations where there is a risk from disease transmission, poisonous and venomous plants and animals, or anti-social or violent behaviour from other people.

Try to avoid working alone in the field. If you must work alone, always carry a mobile (cell) phone and check out and back in with someone who knows your plan of work. Clothing and footwear should be suitable for the terrain and climatic conditions (warm and waterproof, or cool and sun-proof). Safety glasses and gloves should be worn to handle chemicals, and suitable gloves to protect you against thorns and infection from soil and waterborne disease. Take particular care where there is risk of disease transmission such as Weil's disease[15] (near to rivers and canals) and Lyme disease[16] (transmitted by ticks). Blooms of some algae, such as Cyanobacteria (blue-green algae), can be toxic to humans through skin contact and swallowing.[17] In the tropics, it is important to be aware of mosquito-borne diseases (including but not confined to malaria) and how to avoid being bitten both during the day and at night.[18] Be aware of any other risks, including bites from venomous creatures (e.g. snakes), other toxic species (e.g. poisonous plants, scorpions, spiders, and stonefish, which have toxic spines), and the possibility of rabies from mammals (depending on the country in which you are working). Keep your tetanus injections up to date and consult appropriate authorities (e.g. The UK National Health Service[19] and the US Department of Health and Human Services[20]) regarding advised vaccines when travelling further afield. In general, ensure you are properly equipped (see Box 1.2) and avoid risks to help to ensure problem-free project work.

Box 1.2 Suggested minimum equipment required for fieldwork

Always recommended

- First aid kit.
- Map(s) of the area.
- Paper for recording (preferably in notebook form and waterproof if possible).
- Pencils and sharpener (avoid ink if possible; even waterproof inks can run when wet).
- Mobile (cell) phone (fully charged and with spare batteries).
- Whistle.
- Compass or global positioning system (GPS) receiver.
- Watch.
- Appropriate clothing and footwear.

15 http://www.nhs.uk/Conditions/Leptospirosis/Pages/Introduction.aspx
16 http://www.nhs.uk/Conditions/Lyme-disease/Pages/Introduction.aspx
17 https://www.gov.uk/government/publications/algal-blooms-advice-for-the-public-and-landowners/algal-blooms-advice-for-the-public-and-landowners
18 https://www.fitfortravel.nhs.uk/advice/malaria/mosquito-bite-avoidance
19 http://www.nhs.uk/chq/Pages/1072.aspx?CategoryID=67
20 https://www.vaccines.gov/who_and_when/travel

- Appropriate safety equipment (e.g. gloves, safety glasses, etc.).
- Appropriate sampling equipment (nets, traps, plastic tubes, plastic bags). NB: put any samples in a double plastic bag and label each bag so that if one label does come off, the other is there for reference. Whilst some marker pens may be reasonably waterproof, because of the risk of the ink running or rubbing off, it is better to add a label to each bag using waterproof paper marked with a soft pencil).

Recommended depending on terrain, weather, and timing and extent of work

- Survival bag.
- Emergency food.
- Torch (fully charged and with spare batteries).

If working outside of your usual comfort zone, for example in locations overseas, check with researchers or organisations (e.g. The Royal Geographical Society of London[21]) who have experience of working in such countries or habitat types. The travel advice section of the UK Foreign Office website provides up to date information which can help to plan and implement risk assessments for overseas work.[22] Even within your sphere of experience, avoid complacency, since conditions may change and even small risks can be hazardous.

Once the legal, ethical, and health and safety issues have been properly examined, it is important to address these through a thorough formal risk assessment process. Most organisations have their own risk assessment proformas and a protocol for producing and checking them. Remember that you should include any hazards that you envisage could occur, together with appropriate equipment and actions (including training) that should help to mitigate the level of risk. All members of the team implementing the work should be fully briefed on the risk assessment and be comfortable that the work in which they will participate is within their skill level. Include within the risk assessment pack any additional information, such as details of important contacts, relevant COSHH material, appropriate first aid information, etc. It is important to recognise that risk assessments are dynamic, working documents that should be revisited and amended where necessary during all phases of the project implementation (especially if issues do arise). When issues do occur, they need to be recorded and any actions taken fully documented. USHA/UCEA (2011) include a list of possible hazard types that should be considered.

Implementation

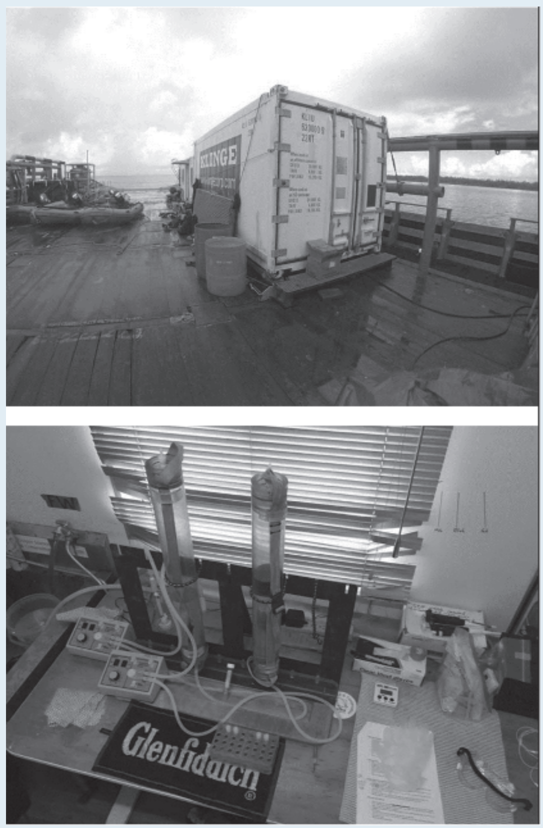
You will need to complete a thorough literature search and possibly even consult other scientists experienced in the field of your proposed research in order to appropriately define and refine the appropriate methods for your study. There will usually be a balance between the ideal solution in terms of the methods used, and logistical restrictions of time, and availability of equipment and expertise. Work through all the logistical aspects

21 https://www.rgs.org/in-the-field
22 https://www.gov.uk/foreign-travel-advice

you can, well in advance of final implementation, to reduce the possibility that there will be significant problems and delays. This is especially important where there is a large team of researchers involved, where the field sites and/or techniques are relatively novel to you, and where equipment and specimens require transporting, especially overseas. Case Study 1.2 discusses an example of how one research team dealt with the creation of appropriate facilities in difficult environmental conditions and successfully brought their samples home. Provided that your project has been well designed, and pilot studies have enabled you to refine your techniques, the implementation of the project should be straightforward. Here, we emphasise the importance of careful note taking and time management during your project. It is important that all participants are thoroughly briefed on the project, its risks, methods, lines of communication, actions to be taken in the event of any problems, etc., before implementation.

Case Study 1.2 Processing and transporting marine microbes from one of the most remote places on earth

Gareth Williams is a Reader (Associate Professor) in Marine Biology based within Bangor University's School of Ocean Sciences in the UK. Much of his work involves collecting data from very remote parts of the planet where conditions are challenging and where there are extremely limited resources on hand. This case study deals with the challenges of processing and transporting marine microbe samples collected at remote reefs in the Indian Ocean.



Photos GW

Temporary microbiological laboratory on board ship.

Model organism and research challenges faced

Marine microbes are an extremely interesting group of organisms; what they are and what they are doing can offer valuable insight into the health of marine ecosystems such as coral reefs. Here, Gareth and his team were interested in sampling marine microbes living in the water column above coral reefs within the Chagos Archipelago, a remote group of oceanic atolls and islands spanning over 200 km of latitude in the remote Indian Ocean. In fact, the coral reefs of the Chagos

Archipelago are arguably some of the most remote coral reefs on Earth and form one of the world's largest Marine Protected Areas.

After travelling from the UK to Bahrain and on to a military base on the island of Diego Garcia located within the Chagos Archipelago, the team boarded the British Indian Ocean Territory's enforcement vessel, the Grampian Frontier.

This was to be their home and science platform for the next 4 weeks and they set sail to survey 29 reefs across the Archipelago, including collecting near-reef water microbial samples. The first challenge arose when attempting to carry out sample processing on a ship designed with cold climates and oily rescues in mind, rather than tropical temperatures and sterile laboratory conditions. This presented a significant problem: how to construct a temporary microbiology laboratory where they could process, freeze, and store samples on such a vessel. The second challenge was to then determine how to transport frozen seawater samples (that contained the microbes) back from Chagos, through Bahrain and then on to the UK, without the samples melting in temperatures regularly exceeding 40 °C.

How the challenge was resolved

Fortunately, the team had managed to ship liquid nitrogen ahead of time and the majority of it had amazingly survived the long, warm ship transit from Singapore to Diego Garcia. This allowed them to flash-freeze seawater samples – a crucial part of the processing protocol. However, things only stay frozen if kept below 0 °C. The ship's cook kindly agreed to allow Gareth to store the frozen samples (and the liquid nitrogen) in a large walk-in freezer on the back deck of the ship, as long as they were properly contained and away from any food. Problem solved.

The next challenge was to construct a laboratory. The solution was to turn a deck container that had been sweltering and rotting in the heat of the tropics for the best part of a year into a workable space for microbiology – a discipline that demands consistently sterile working conditions. This required ethanol, lots of ethanol, a cloth, and some good old-fashioned elbow grease. Ethanol, fortunately, was not in short supply. Having sectioned off a portion of the container and cleaned it repeatedly, this appeared to solve the problem – provided Gareth repeated the cleaning frenzy on an almost daily basis. Asking the other scientists nicely to stay away from the area with their wet and dirty equipment was also very crucial.

Then came the end of the journey and the challenge to transport the samples back to the UK and keep them frozen when flights, transfers, and baggage handling are largely out of the team's control. Gareth had a small cooler to aid in this process, but key to its success was keeping it cool throughout the journey and at all costs trying to avoid it sitting on a blistering tropical tarmac runway. The team froze sponges in the last of the liquid nitrogen and packed them in to the cooler around the samples, while also accidently freezing part of a colleagues' shoe! They had also frozen Nalgene bottles full of water to act as ice blocks, but ones that have better longevity than your standard home freezer block (the ice is much thicker). These were also packed in the cooler. However, days prior to this, Gareth had realised the ambitious nature of the journey home and the fear of the samples defrosting. Despite extremely limited internet access, he had managed to get an email out to his post-doc with an emergency request to organise delivery of 2 kg of dry ice to the hotel in Bahrain where they had planned to spend the 9 hours transit time prior to their UK flight. Dry ice is a solid form of carbon dioxide and is around −80 °C. The two 1 kg dry ice blocks kept the samples frozen all the way to Manchester airport and the following 1.5 hour taxi journey to the safety of a freezer.

Advice for students working under challenging conditions

Always plan ahead and have back-up plans A, B, and C for fieldwork. Remember to stay safe even when pushed for time under challenging circumstances – it could have been much worse than a frozen shoe! When transporting frozen samples, invest in a good cooler with proper insulation, pack it well (i.e. leave no air spaces and put the most important samples at the bottom), and make your own ice blocks that have a large volume to surface area ratio. Finally, do not rely on luck on the day as Gareth did, but instead organise any additional supplies you may need for your sample journey to be a success well ahead of time.

Equipment and technical support

Ensure the availability of equipment before starting and obtain essential items well in advance of beginning your research project. You may need to allow adequate time to order specialist equipment or materials. If your project requires technical support, arrange this as far in advance as possible.

You need to be as familiar with your equipment as possible before commencing your fieldwork. This includes knowing how reliable it is likely to be under the conditions in which you are working and whether you need to have access to spare components or extra full items of equipment. For example, small mammals will eventually chew through the sides of an aluminium trap and, and they rather more quickly get through the sides of an equivalent plastic trap. Whilst it is possible to patch these up, this is tricky in the field, and therefore spares should be taken. Anything that runs on batteries (e.g. data loggers or light traps) need to be recharged on a regular basis and spare batteries, bulbs, etc. should be available whilst in the field. If you are using multiple pieces of equipment, then you should ensure they are comparable (e.g. different makes of bulb may provide different wavelengths and illumination in light traps, and monitoring equipment from different companies may have different levels of accuracy and resolution). Wherever possible, ensure that identical equipment is used for an individual project. Instrumentation errors may occur if users are unaware of the limits of the equipment (e.g. where attempts are made to estimate between gradations on an analogue scale). Some equipment may require regular calibration against standards of approximately similar values to the variables being measured (e.g. calibrating pH meters at pH 7 for neutral soil and water pH measurements). It is also important to take care of equipment, including protecting it against vandalism, theft, and animal damage (many a moth trap has been trampled by inquisitive cattle when placed in their pasture, and crows and magpies seem to very much enjoy pulling white pitfall cups out of the ground).

Field/laboratory notebook

Keep all your data and notes in an organised format, preferably in a hard-backed notebook (see Box 1.3) and scan these on a regular basis to retain an electronic second copy.

Have a standardised way of recording your data, including everything that might be relevant: the date, weather conditions, and notes of any important points that occur to you whilst carrying out the project. It is useful to record data in the same layout as you will on a computer spreadsheet for analysis (see p. 28). If you do use sheets of paper (similar to the one illustrated in Figure 1.3), make sure that they all go into a ring binder as soon as possible. It is very easy for single data sheets to get lost. It is worth checking to see if there is a standardised recording sheet available for use with the technique that you are employing. Examples include the recording sheets produced for the Biological Records Centre[23] and Breeding Bird Survey.[24] Make photocopies of data at frequent intervals and scan them into a computer if possible. If data loggers are being used either to note climatic variables (see Chapter 2) or to log behaviour (see Chapter 4), then make sure that you take backups of your files as soon as possible. Enter data and comments in electronic form whenever possible and create backup copies on a regular basis, including copies saved on a networked drive, internet hub, or cloud (NB: emailing copies to yourself and co-workers/supervisors can provide useful protection).

Box 1.3 Keeping a field notebook

Use a field notebook to write down data, ideas, observations, tentative conclusions, and hypotheses as you do your fieldwork to create an immediate and faithful history of your research. Produce comprehensive, clearly organised notes as a reference and so that you can reconstruct the research timeline and follow the development of your thoughts and ideas. Although you may use other collection sheets (e.g. pre-printed data collection forms to ensure data are collected consistently in different locations and at different times), your field notebook should provide the context for data collection and help resolve ambiguities or inconsistencies when preparing for analysis. After data analysis, reference to your notebook may generate further hypotheses and suggest further lines of enquiry.

Select an A5 or A6 hardback notebook with a spiral binding and wide-ruled lines, ideally on waterproof paper. Use a clutch-type propelling pencil with a moderately soft lead (HB or B). If you do not use waterproof paper, then encase your notebook in a plastic bag large enough to cover your hand and the notebook when writing. In very wet conditions, write on an A4 sheet of white plastic with a thick soft pencil (use kitchen cleaner to erase your notes after transcription).

What should be recorded?

The first page should include contact details in case of loss, the subject of your research, and the start and end dates of the period covered by that notebook. Include any conventions used, e.g. 'All times are recorded as local time'. Number the pages and ideally add a contents table to make searching for information easier. Write on the right-hand page only so the left-hand page can be used for ideas generated by reading about similar observations or relevant research papers. Leave a few lines between observations for

23 www.brc.ac.uk/record-cards
24 http://www.bto.org/bbs/take_part/download-forms.htm

comments to be inserted later (e.g. 'No bark damage here 23 June, see p. 39'). Add a 2 cm margin to write the time, location (e.g. from a GPS reading), or other identifying labels. Create lists of codes, acronyms, specialist terminology, etc. at the back and include any emergency numbers (e.g. those of field buddies). Other useful notes about equipment (how to use, limitations of instruments, etc.) and any numerical information you might require in the field (simple formulae for calculations, random numbers, etc.) can also be added here.

Before starting work each day, write down the date, weather, general location, nature of the habitat, and purpose of the day's work. Write down any changes in weather or habitat that occur during the day, e.g. 'At 15.00 hours snow began to fall and visibility was reduced to 20 m'. When observing behaviour, note the sampling method, how animals were chosen for observation, and the recording method (e.g. whether you noted all occurrences or used a time-sampled method). If animals or start times are chosen at random, note how this was done.

Note the type and model number of any equipment (e.g. GPS receiver type Garmin 12). Some instruments need calibrating at intervals, so record the time of calibration and any raw data and subsequent calculations so that any arithmetic errors can be identified and corrected later. Use your notebook to create rough species accumulation curves, etc. so you can tell when you should stop collecting data (see p. 31). Although notes should be made at the time observations are made, it can be difficult to observe and write at the same time, but if you do rely on memory, you should note this. Write exactly what you see or hear, e.g. when describing behaviour do not ascribe a function to it in the guise of a description (i.e. do not write that a goose was 'vigilant' when you mean that the bird was in a standing posture with an elongated neck and raised head).

Sketches enhance any photographs you take of your study sites and you will have a sketch available in your notebook the next time you visit the area. Sketches can be added to subsequently (annotating any changes with the date of the amendment). The value of sketches can be increased by explanatory labels. A careful sketch can aid species identification and will help to jog your memory when you encounter a species in the future; such sketches are more valuable if labelled with the diagnostic feature(s) *you* use (e.g. 'two spots on forewing' or 'sepals bent backwards'). Landscapes change over time and maps may not reflect this. In some cases, no map of a suitable scale may be available, and a sketch map can be made using compass and tape, or by pacing out distances using a pedometer. This may be adequate to note the locations of those animals or plants of interest.

It is also useful to record any notes and actions from supervisory or team meetings, both as a reminder and to ensure that any designated actions have been completed as planned.

Pilot studies

Although you might worry that using a pilot study will delay the start of you collecting the data you need for your project, in reality it may save you time and further problems down the line. Trying out your project over a small scale (including in time), will enable you to better plan the logistics of implementing your project. Not only do you get a chance to ensure you can realistically complete the project, you will also gain an insight into the variability and values of data you are collecting. Such insights are particularly useful when assessing the sample size required (p. 30). Pilot studies enable you to become familiar with your techniques so that the data collected will be the result of skilled implementation, rather than early data being the result of less well managed techniques that are new to you. Any changes that you include to your final design as a result of the experiences from your pilot studies should be documented in your final methods section of your report.

Time management

Conducting a piece of research within a relatively short timespan is a demanding process and will require careful time management. You will need to allow time to check the feasibility of the research project and should also add sufficient time for method training or familiarisation. Ensure that you leave time spare to allow for the almost inevitable problems associated with both fieldwork and laboratory analysis. Check out your methods by first running a pilot to help identify any pitfalls and inadequacies. It is easy to underestimate how long it takes to analyse, interpret, and write up a research report. Figure 1.2 illustrates a timetable in the form of a Gantt chart for a research project expected to last 33 weeks – around a full academic year: many research projects are substantially shorter than this. Note how the time allocated to actually collecting the data is relatively short and overlaps with the ongoing literature review and writing process. Also, you are strongly urged to begin writing the report before completing your data collection; it should be possible at this stage for you to write the methods and introduction since the former will describe what you have been doing, whilst the

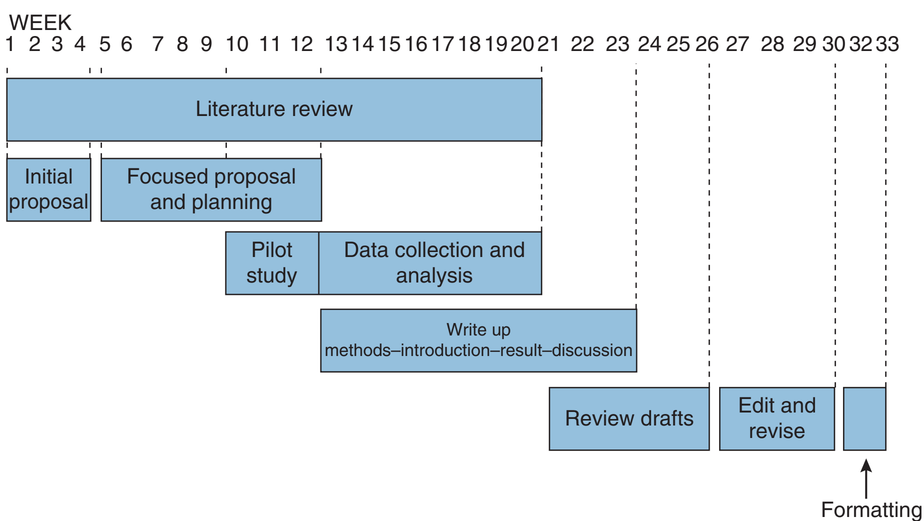


Figure 1.2 Example timescales for a medium-term research project. Note that health and safety, ethical, and legal issues should be examined within the initial proposal and a refined risk assessment produced during the focussed proposal and planning phase. This risk assessment should be revisited regularly throughout the project.

Table 1.1 Example timescales for a short research project.

Day	Morning	Afternoon	Evening
1	Select topic, identify aims and hypotheses, create programme of study, select sites and techniques, identify major resources required, complete risk and ethical assessments	Test, and become experienced in using, techniques and equipment, implement a pilot study	Evaluate pilot study, amend programme of study
2	Implement amended programme and collect data		Enter data into spreadsheet, write methods
3	Collect data		Enter data into spreadsheet, write introduction
4	Collect data		Enter data into spreadsheet, edit introduction and methods
5	Collect data		Enter data into spreadsheet, plan results tables and figures
6	Collect data		Analyse data
7	Write results section	Write discussion	Complete report

latter reflects your background reading of the literature. Table 1.1 shows an example timescale for a research project lasting a week (e.g. on a field course). Note that even with such a short timescale, sufficient time needs to be devoted to planning the project to ensure the best chance of success (see Box 1.4).

Box 1.4 Some tips on time management

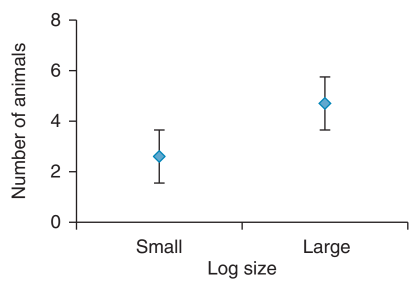
- Be realistic about what you can achieve in the time you have available and work within your strengths and weaknesses.
- Plan your long-term goals.
- Have a weekly plan, with realistic and achievable targets, and update this on a regular basis to reflect your progress.
- Identify not only the key phases of your research project, but also other areas that will take up your time (both in terms of study and general living) to ensure that your research timescales are realistic and that your aims are achievable.
- Prioritise your work into that you have got to do, that you ought to do, and that you would like to do but may not have time.
- Make good use of your time: a trip by train may be an ideal opportunity to read references or edit your manuscript.

Statistical considerations in project design

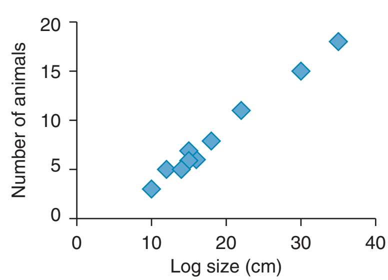
Since research is about asking questions, you need to design your project so that it will answer them effectively, without allowing your design to introduce ambiguous results, or results that are open to other interpretations. This is where the planning phase starts to define what you are going to measure and how. If, for example, we investigate the types of birds found inhabiting a woodland patch, then we have a choice of ways in which we record the data. We might note how many individual birds there are, or the numbers of each feeding type (insect feeders, seed feeders, etc.), or how many individuals there are in each species. These measurements enable us to obtain a picture of the birds found in a woodland patch. If we monitor birds only in a single woodland patch, we could worry that our chosen woodland is unusual in some way and therefore not representative of woodland patches in general. We could therefore examine a series of patches and obtain data for 10 or more of these. Now, if we wish to describe how many birds were found in all of these woodlands, we require some sort of descriptive statistic to summarise the information across 10 or more patches. Descriptive techniques include estimates of the average values per sampling unit (e.g. per site), population estimates and densities, methods of describing distributions (i.e. whether organisms are distributed randomly, evenly or in aggregations), and measures of community richness including diversity and evenness indices. These techniques are discussed in more detail in Chapter 5.

Most projects go beyond a simple description of particular species and sites in an attempt to make comparisons or generalisations that can hopefully have wider applicability. For example, if we decide to investigate whether the number of animals found under decaying logs on a woodland floor is influenced by the size of the log, we might approach this in one of three basic ways:

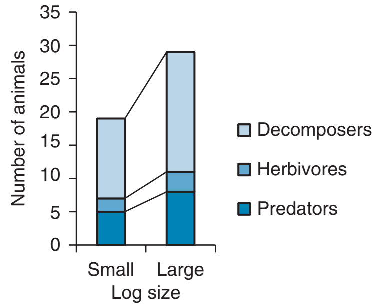
1. by looking at possible differences between samples; for example, if the logs were easily divided into two classes (large and small, i.e. <20 cm and ≥20 cm), we could compare the numbers of animals found under each size class;



2. by looking at possible relationships between variables; for example, we might have a wide range of sizes of logs and decide to examine whether the number of animals varies in some systematic way (either increasing or decreasing) as log size increases;



3. by looking at possible associations between frequency distributions; for example, we could compare the frequency of predators, herbivores, decomposers, etc. from under each of two size classes of logs (i.e. <20 cm and ≥20 cm).



From this simple example it can be seen that how we ask the question has an impact on how we design our study. The three different ways of looking at this study (listed above) also illustrate three broad (basic) types of statistical questions: differences, relationships, and associations between frequency distributions. We will examine each type of question in a little more detail later in this chapter (p. 36), whilst the analytical techniques needed to answer these questions are described in Chapter 5. Other questions that might be asked include: looking at the similarity of sites based on their species composition (i.e. are the animals found under logs from different tree species similar in species composition to each other?); or predicting the presence or numbers of a species from a knowledge of the environmental conditions (e.g. is the presence of wood ants' nests predictable if we know the woodland type, topography, microclimate, etc.?).

By careful design, we strive to ensure that our study does not produce ambiguous results. For example, in a comparison of the invertebrate diversity between urban ponds and rural ponds we could aim to include the size of each pond studied into the survey design. If we did not manage this, and found that the rural ponds surveyed happened to be both larger and contain more invertebrates, it would not be clear whether the results were due to rural ponds being more diverse or whether it was simply an effect of pond size. The correct experimental design would be to either standardise on a given pond size for both environments, or make sure that the full range of pond sizes was included in both environments. Pond size would then be measured, recorded, and built into the subsequent analysis: pond size is then an example of a 'covariate'. Other factors that would have to be standardised, or at least recognised as covariates, in this particular study would be the quality of the water, the pH, age of pond, and so on.

The goal of the study may be to get a deeper understanding of the system by gathering a wide range of variables. In the pond survey example, this might mean that in addition to pond size, we should take various measures of water quality and chemistry (nutrient status, oxygen content, pH, etc.) and the numbers of each species of plant and animal. For a given number of ponds, there may be a large number of variables giving rise to a complex datasheet. In this example, each pond would have its own row in a spreadsheet, and each variable (e.g. size, pH, number of species) would be a column. In order to examine and make sense of such a complex data set, we would need to move into the realm of multivariate analysis (see Chapter 5).

Designing and setting up experiments and surveys

There are two main approaches to collecting field data: experiments and surveys. An experiment involves the manipulation of a system, whilst a survey depends on observations being taken without manipulation. For example, if we were interested in how many invertebrates could be found under logs of varying size, we could either survey a woodland floor finding as many logs as possible and recording both the number of invertebrates and the size of the log, or we could devise an experiment where we placed logs of differing sizes on a woodland floor and after a period of time examined the number of invertebrates underneath them. The advantage of the experimental approach would be that we could standardise all aspects of the logs except for size; for example, age, the degree of decay, the type of wood, and the distance between logs. All of these factors may influence the invertebrates found and confuse any relationship with log size. However, with a survey we would get an impression of what was happening in a real-life situation (i.e. under logs that had been naturally deposited). Moreover, we may decide that the experimental approach is damaging to the environment; here, artificially placing logs in a natural system. In addition, for practical reasons, we might decide that the colonisation of newly introduced logs by invertebrates would take longer than the time available for the project to be completed. In most environmental research programmes, surveys are useful for generating ideas about important factors, but because of the additional complexity in real situations, surveys cannot identify cause and effect. Because experiments strip away the additional complexity, they are more useful in identifying cause and effect, but less likely to be applicable to real-life situations.

When designing experiments, it is important that as many factors as possible are kept constant. So, for example, if we are interested in identifying whether an increase in insecticide concentration will lead to a decrease in aphid infestation of a crop, then the same amount of water (assuming this is the solvent or carrier for the pesticide) should be used for each application (irrespective of the concentration applied) so that we are testing the amount of pesticide added, rather than the amount of water added. In addition, it is important, where possible, to include a control treatment. In this example we would use a water only treatment to see if the addition of any water had an impact. If we did not do this and found a reduction in aphid numbers with any application of pesticide, we would be unable to tell whether this was due to the pesticide or the fluid added.

Choosing sampling methods

The choice of sampling method will usually be dependent upon the habitat type and organisms being studied (see Chapters 2–4). However, all sampling techniques have limitations, and there are some general principles that are applicable to most sampling methods, for example:

- Some techniques may be suitable for a limited range of habitats, or be biased in favour of active rather than sedentary animals, or collect only a subset of the population being examined (e.g. males rather than females, or those migrating rather than those resident). It is therefore very important that limitations are known and accounted for during the design of the research to avoid later problems in interpretation.

- Maximising the number of replicates or survey points to increase a study's power is desirable. However, this is often constrained by fieldworker, equipment, species, or habitat factors. For example, a common misconception is that behavioural studies in the wild, particularly with large mammals, will yield sufficient data for robust analysis. However, often such data are of poor quality or lacking entirely, ironically because large animals are often hard to observe. Under such circumstances, the observer may have to either abandon the study or report using descriptive or qualitative methods. We cannot emphasis enough the importance of estimating how much time it can take on average to get one data point in order to derive the time needed to complete the whole field study component in sufficient detail for statistical analysis.
- Many techniques are not directly comparable with each other, and even using the same technique, but under different conditions (e.g. between habitats with very different vegetation layers, between night time and daylight collections, at different times of the year) may not produce comparable data.
- Limitations of the equipment being used may mean that monitoring environmental variables is restricted if, for example, differences between areas are smaller than the accuracy of the equipment allows.
- Resource issues may determine the methods available for use: the cost of equipment, necessity for training, ease of relocation of apparatus between sites, and health and safety issues could all limit the choice of methods.

Types of data

In order to design an appropriate experiment or a survey, you need to think about the type of data you wish to collect. The pieces of information that are recorded (e.g. height of tree, number of birds, density of plants per unit area) are termed 'variables', and may be in the form of one of three types of data. The simplest type is categorical or nominal data where each value is identified as one of several distinct categories (e.g. male or female animals; purple, red, or yellow flowers; grasses, ferns, herbaceous plants, shrubs, or trees). Where we can place the categories in some kind of logical order, so that the data are able to be ranked, this is called ordinal data (e.g. large, medium sized, or small ponds; above the high tide line, mid shore, and below the low tide line on a rocky shore). The most detailed type of data are those measurements that not only can be placed in a logical order, but where there is also a known interval between adjacent items in the sequence (e.g. the number of deer in a herd; the temperature in the centre of patches of plants; the depths of a series of ponds). There are two types of measurement data: interval data and ratio data (see Box 1.5). In most cases, the analysis of interval and ratio data uses the same statistical techniques and so in this text we will tend to combine them and refer to them as interval/ratio data or measurement data.

Box 1.5 Differences between interval and ratio data

Interval data have no true zero so that negative values are possible (as in temperature measured on the Celsius scale where 0 °C refers to the freezing point of water rather than the lowest possible temperature) and where measurements cannot be multiplied or divided to give meaningful answers (as in dates).

Ratio data are measurements that have an absolute zero point that is the lowest possible value (as in temperature measured on the Kelvin scale where zero Kelvin is absolute zero) and so negative values are not possible (e.g. you cannot have – 6 foxes). With ratio data, all basic mathematical operations can be performed to give meaningful answers. For example, you can derive a ratio of water lost from soil following drying out as follows (where the original mass = 20 g, and dried mass = 16.5 g):

$$\text{the proportion lost on drying} = \frac{20 - 16.5}{20}$$
$$= 0.175 \text{ which equates to } 17.5\% \text{ of the original mass}$$

Note that we can readily reduce measurement data to ordinal or categorical, but not the other way around. Thus, if we count the numbers of invertebrates of different species on a particular type of plant, we could subsequently express this in order of dominance from abundant through to rare (an ordinal scale), or indicate the presence or absence of different species (categories). However, if we merely record presence and absence of species, we cannot subsequently calculate the numbers of individuals. Thus, if in doubt, it is safest to collect the information at the highest resolution possible.

It is good practice to use a standardised data recording sheet (ideally in your field notebook) that is as similar as possible to the way in which data will be entered into a computer for analysis to avoid data transcription errors in moving from paper to a computer spreadsheet. In our example (Figure 1.3), we have two types of variables: fixed and measured. It is easier to deal with these in order so that fixed variables come first, followed by measured variables. Fixed variables are those determined by the research design and do not vary during the investigation (record number, site, date, and time). Hence, these can be added to the recording sheet early in its production. Measured variables, on the other hand, are those factors recorded during the investigation the values of which will vary depending on the site, date, time, etc. (numbers of wrens, blackbirds, etc.). Sometimes, derived variables are also required (i.e. variables produced from measured data, e.g. the proportions that each species forms of the whole catch). Such derived variables can be added to the right of the measured data once the latter have been entered on a computer spreadsheet, since the required computations are usually easily carried out using spreadsheet functions. In most cases, data will be recorded as numerical values. Where categories (e.g. site) occur, codes or names can be used, although some computer programs will not accept letter codes, so you may need to allocate numeric

Data recordingsheet: *12*

Comments and notes:

Weather mild, sunny at first then clouded over a little around 9 am

Ramblers walked by at about 8:30 am – didn't make much noise (no dogs!)

Record number	Site	Date	Time	Number of wrens	Number of blackbirds	Etc.
0101	*Black Wood*	*02/05/19*	*08:00*	*4*	*0*	
0102	*Black Wood*	*02/05/19*	*08:15*	*3*	*2*	
0103	*Black Wood*	*02/05/19*	*08:30*	*2*	*2*	

Figure 1.3 Example of a section of a data recording sheet for an investigation into the distribution of woodland birds.

codes to such variables. You should make sure that any paper copies of results sheets are photocopied or scanned as soon as possible after completion, and that electronic copies are properly backed up.

Sampling designs

When implementing a project, it is rarely possible to collect information on all the animals or plants present. Usually we need to use a sample that we hope to be representative of the situation as a whole. The total number of data points that could theoretically be gathered is known as the population (this is a statistical population rather than the actual population of animals or plants – see Box 1.6); the actual number of data points is termed the sample size. Larger samples are usually more representative of populations, although this depends on the variability of the system being studied (small samples may be reliable representations of populations with low variability). Those elements of a system that are calculated (e.g. the mean number of plants, such as plantains, per square metre in a meadow) are termed statistics and are estimates of the true attributes of a statistical population (called parameters – see Box 1.6). So, if we counted all the plantains in the entire meadow, we would be able to calculate the actual mean value per square metre (a parameter). Since it is usually impractical to count all individual plantains, in reality we usually count plantains in a subset of the meadow (i.e. take a sample), and calculate the mean numbers per square metre using this sample in the expectation that it will be representative of the whole site (a statistic). This sort of situation occurs in many types of survey. For example, market researchers obtain opinions from large groups (samples) of people and use these to indicate the attitudes of the population as a whole.

Box 1.6 Terms used in sampling theory

See also the Glossary of statistical terms in Appendix 1.

- A **population** is a collection of individuals, normally defined by a given area at a given time. For example, scientists refer to the decline in the world population of Atlantic cod in the last century or the annual harvest of Northeast Atlantic cod. These are both true populations. The size of a population is rarely measured directly but usually estimated from samples.
- A **sample** is a term that can be used ambiguously, but is a subset drawn from a population, which usually includes a quantity. For example, 100 individual fish taken from the Northeast Atlantic cod population and measured in order to get an estimate of body size. Another example would be taking 50 small areas from a meadow (each 1 square metre in size) in order to count the number of plantains within them.
- A **parameter** is a population metric that is estimated from a variable (e.g. the mean body size of Northeast Atlantic cod, or the mean number of plantains per square metre of a meadow) and can be used to summarise data. Importantly, statistical tests aim to estimate parameters from a population in order to test for differences, relationships, associations, etc.
- A **variable** is a measurement that may change from sampling unit to sampling unit (e.g. the body size of Northeast Atlantic cod taken from a sample, or the number of plantains in a square metre of a meadow) and can be used to summarise collected data (e.g. by taking the mean).

The decision over which samples to take requires some care, and at this point it is worth discussing why replication is important. Since environmental systems are usually intrinsically variable (i.e. physical, chemical, and biological factors differ spatially and temporally), the larger the sample, then the more representative it will be of the population (i.e. the more of the natural variation will be covered). However, the larger the sample, the more time and effort it will take to collect it. There are methods to calculate the optimum sample size; however, these rely on knowledge of the variability of the system. This is rarely known in advance, although a small pilot study may give some indication. If it is known or suspected that there is substantial variability, then a large sample should be taken. In most ecological surveys, a large sample would include over 50 observations. However, where the population is likely to be very large and variation is expected to be great, even larger sample sizes may be required. Otherwise it is best to aim for as large a sample as possible after taking into account constraints including the size of the workforce, the time available, and how much material is present in the system under investigation. Sometimes, previous studies on similar topics can be used as a guide to what a reasonable sample size might be (i.e. from the literature or from a pilot study). Where several levels of a number of variables are to be analysed (e.g. male and female animals of each of three different age groups: young, mature, and old), then it is important to take sufficient replicates of each subgroup (e.g. young males, mature males, etc.) to be able to account for within-group variability. This will inevitably have an impact on the required sample size and is another reason why the intended statistical analyses should be considered at an early stage of project planning. Box 1.7 shows the factors that should be considered when determining the sample size.

Box 1.7 Aspects to be considered when determining the sample size

A larger sample size is needed when there is:

- high variability – use a pilot study or consult similar investigations in the literature to get a feel for the likely variability;
- a small difference or relationship or association to be detected – it is worth recognising that very small differences may not be important ecologically (e.g. a native plant may have more insect species than an introduced one, but if this difference is by only one or two common insects, it is unlikely to be of conservation importance);
- a requirement to subdivide the data for analysis (e.g. separate analysis of males and females would require similar appropriate sample sizes of both males and females).

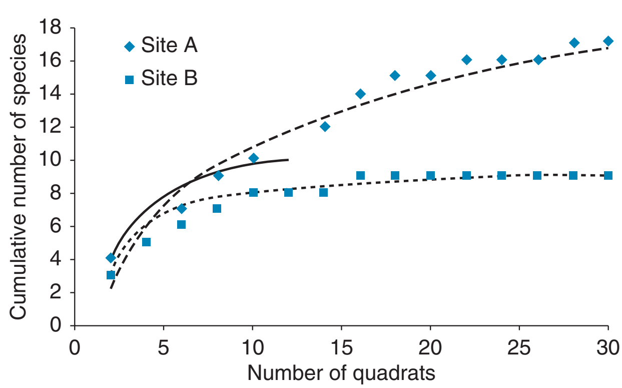
See Krebs (1999), van Belle (2002), and various online calculators[25] for further details of the different calculations that can be used to estimate sample sizes, depending on the intended statistical analysis technique to be used.

25 For example, Rollin Brant at the University of British Columbia (http://www.stat.ubc.ca/~rollin/stats/ssize/index.html) and the Australian National Statistical Service (http://abs.gov.au/websitedbs/D3310114.nsf/home/Sample+Size+Calculator). In order to access sample size calculations in R see http://www.statmethods.net/stats/power.html

In surveys of community structure, it may be important to know that the majority of species in an area have been recorded at least once in your sample. In this case, species accumulation curves may help. At its simplest, this involves plotting the accumulated number of species against increasing sampling effort. Sampling effort is the number of sampling units (quadrats, pitfall traps, animals handled, hours of observations, sites surveyed, etc.). Box 1.8 illustrates the use of species accumulation curves in quadrat sampling (see Chapter 3). There are a variety of methods of modelling species accumulation curves (see Colwell et al. 2004 and Magurran 2004 for further information) and many standard software packages include routines for this (e.g. those obtained from Pisces Conservation).[26]

Box 1.8 Species accumulation curves for two sites

By plotting the cumulative number of species found against the number of quadrats examined, it can be seen that as the number of quadrats used increases, the number of species also increases. At the point at which the curve levels off towards the horizontal (the asymptote), we may assume that we have obtained the maximum number of species and can stop sampling. For site A (dashed line, diamonds), we may not yet have reached the total number of species, even after 30 quadrats, and should consider increasing the sampling effort. For site B (dotted line, squares), it appears that we have reached about the maximum number of species that we can expect to get. In fact, we probably reached this number at round about 16 or so quadrats. This difference between sites A and B might reflect not only a difference in the number of species found there, but also a difference in heterogeneity of the site, with site A being less homogeneous than site B. Note that had we looked at the data for site A after 12 quadrats (solid line, diamonds), we might have assumed that we had reached the maximum number of species as the curve levels off. This highlights the importance of collecting past the initial point of curve levelling to check that it truly does reflect the asymptote.



26 http://www.pisces-conservation.com/software.html

Since we generally take a sample in order to make a valid estimate of a parameter of the population (e.g. the number of species, the mean temperature, the proportion of predators), a central requirement is that the individuals sampled are independent of each other. It is important to recognise, and avoid or if not account for, situations where the individuals sampled are linked in some way as a result of the sampling design. For example, we might compare the number of spangle galls found on leaves chosen at random on oak trees growing in clumps, with those on isolated oak trees. If we found over 20 trees in separate clumps, but only 10 isolated trees, we might be tempted to take double the measurements from each of the individual isolated trees. However, this would mean that individual data points from isolated trees were linked by virtue of the tree on which they were growing and shared many different attributes with each other. Such data would not be independent of each other (known as pseudoreplicates) and hence may cause problems in interpretation since we would be unsure whether any differences between clumped and isolated trees were due to the multiple measurements from some trees. It would be better to use unbalanced sample sizes (i.e. 20 clumped and 10 isolated trees) than use non-independent data. Similarly, we should not take data from more than one tree in any clump since these are likely to be more similar to each other than to those in other clumps. From a statistical analysis point of view, few tests require equal sample sizes and, even where this is a problem, it would be preferable to reduce the number of trees from clumps that were measured. Note that we may wish to take account of some of the variation between leaves on each tree by taking several (perhaps 10) leaves per tree and using a mean value to represent each tree. There are also statistical tests that allow for multiple measurements per tree, but these usually require the same number of samples per sampling unit – see repeated measures analysis in Chapter 5.

If we survey a pond in order to look at the animals and their relationships with several physical, chemical, and/or biological factors, then no matter how many replicates we take, we are merely describing what happens in a single entity (i.e. this one pond). Such a study does not tell us anything about pond ecology in general, and the use of such replicates is termed pseudoreplication and should be avoided (Hurlbert 1984; van Belle 2002). In order to broaden our approach and gain more of an understanding of ponds in general, we would need to study a large number of separate ponds. Thus, studies of single sites or small parts of sites may not reveal information applicable to the wider ecological context.

In some situations, the data collected are linked to each other by design. For example, we might be interested in comparisons of matched data (e.g. examining the animals found on cabbages before and after the application of fertiliser or pesticide, or the numbers of mayfly larvae found above and below storm drain outflows into a series of streams). These designs can be perfectly sound, but because the data are matched (by cabbage or by stream) we require a slightly different approach to the resulting analysis (see Chapter 5).

When designing your sampling strategy, it is important to consider the variability and whether the timing or order of sampling might bias the result by measuring only part of the potential variation. For example, sampling the insects present on thistle flower heads will be biased if all the data are collected in the early morning, since this will miss any animals that are active later in the day. If two areas are being compared, sampling one site early and one site later will introduce another variable into the comparison: we would not just be looking at the two sites, but also at two times of day. Since it would be impossible to separate the two variables, it would be difficult to draw conclusions from such a survey design. In this example we would say that the findings were 'biased by time of day'. It is in managing some of this variability that experiments come into their own, because they standardise as far as possible the conditions under which the subjects are examined, thus removing bias. It is much easier to design an experiment where only one factor (also known as the treatment) is manipulated, whilst all others remain constant. However, if we wished to survey a real-life situation (as opposed to examining a rather more artificial experimental design) then we would take into account the time of day. We could do this by designing our survey so that we alternated the measurements or observations that we took from our two sites, sampling first one then the other, then back to the first, and so on, to get a spread of measurements for each site over the day. Alternatively, we could sample on successive days, reversing the order in which we sampled the sites on each day, or justify the need to obtain additional fieldwork assistance to make a balanced study easier to implement.

There are several sampling layouts that help us to avoid bias. One commonly used approach is random sampling. Here, a random sequence is used to determine the order in which to sample plants, or the coordinates to sample experimental plots or survey sites. Hence, if we wanted to randomly sample 1 m × 1 m quadrats in a field, random coordinates can be used to position the sampling sites (Figure 1.4a) using pairs of random numbers generated using a calculator or computer, or obtained from a table (see Table 1.2). This works by using pairs of numbers as sampling coordinates, so if we have coordinates of 23 and 85 in a sampling grid that is 10 m by 10 m, we would place our quadrats 2.3 m along the base and 8.5 m up the vertical axis. In our example above, of insects on thistle flowers, random sampling may also be used to determine which site is visited first: here sites would be allocated number codes that are then selected randomly from the table.

Table 1.2 Random numbers.
Coordinates can be extracted simply by taking pairs of random numbers in sequence from the table (e.g. 23, 85 – shaded values – provides the position within a sampling area where we would take the first measurement of a series).[27]

23	85	56	84	92	4
62	51	27	74	83	84
56	32	87	75	95	5
87	7	20	30	25	12
99	86	29	41	29	39
31	73	30	73	27	97
24	38	91	16	17	66
94	59	12	17	37	39
41	67	25	42	2	84
32	67	48	99	74	3
68	1	59	20	25	7

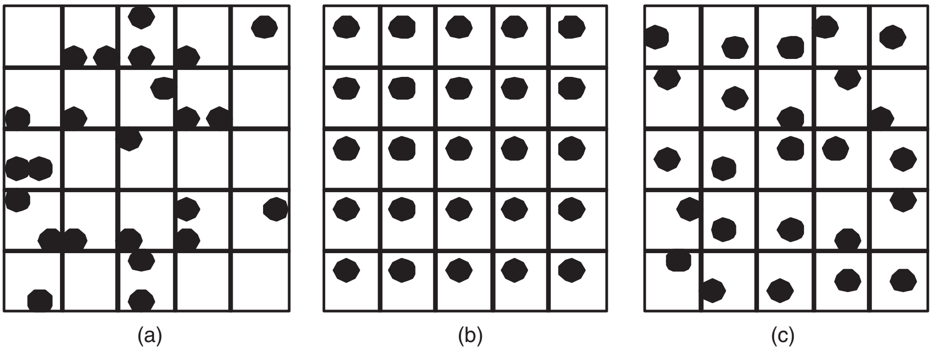


Figure 1.4 Examples of sampling designs. (a) Random sampling; (b) systematic sampling; (c) stratified random sampling.

Although random sampling is often appropriate for selecting sampling points, where there is a great deal of variation across a sampling unit such as a site, by chance the coverage may not include all of the heterogeneity present. For example, in Figure 1.4a, the two squares in the lower right of the sampling site have no sampling points. If the site was reasonably homogeneous, then this would not be a problem. However, if these small squares represented the only damp area within the site (covering around 8% of the total area), then this particular habitat variation would have been missed altogether. An alternative strategy would be to use systematic sampling (Figure 1.4b). This is an objective method of spreading the sampling points across the entire area, thus dealing with any spatial heterogeneity. So, to systematically sample the insects on trees, we might collect from every tenth tree in a plantation.

Usually, systematic sampling would provide us with random individuals – unless for some reason every tenth individual is more likely to share certain characteristics. Suppose we used systematic sampling to examine the distribution of ants' nests in a grassland. We could place 2 m × 2 m quadrats evenly 10 m apart across the site and then count the number of nests within each quadrat. However, if ants' nests are in competition with each other, they are likely to be spaced out. If this spacing happens to be at about 10 m distances, we would either overestimate the number of nests if our sequence of samples included the nests, or underestimate if we just missed including nests in each quadrat. It would be better in this situation to use a mixture of random and systematic sampling (called stratified random sampling – Figure 1.4c) where the area was divided into blocks (say of 10 m × 10 m) and then the 2 m × 2 m quadrats were placed randomly within each of these. This type of sampling design can also be applied to temporal situations by, for example, dividing the day into blocks of 4 hours and allocating the order of the sites to be sampled within each block using different random numbers.

More sophisticated methods of laying out sampling plots (or allocating sampling periods) may be useful for planning experiments. We could lay out a series of treatments in rows so that we have replicates of each treatment (Figure 1.5a). Although this would be relatively easy to manage (adding fertiliser, dealing with particular cutting regimes, etc.) since each treatment is clustered together, there may be variability within the plot that

27 FCStats provides a program to produce random numbers for any size of plot: www.wiley.com\go\wheater\practicalfieldecology

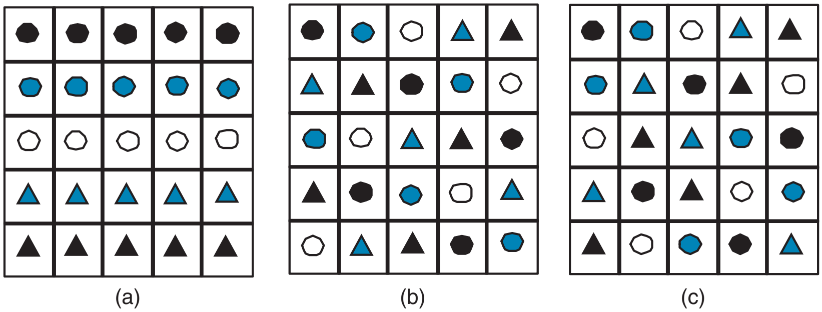


Figure 1.5 Experimental layouts for five different treatments. (a) Clustered design; (b) stratified design; (c) Latin square design. Each treatment is represented by a different symbol.

masks the impacts of the treatment themselves. An alternative is to ensure that each row and each column of the plot has one of each treatment (see Figure 1.5b). It is even better if these treatments can be distributed randomly, whilst still maintaining an even spread across the rows and columns using a Latin square design (see Figure 1.5c). Variations on this theme have been proposed, including ones based on the patterns used in the Suduku game (Sarkar and Sinhar 2015).

Planning statistical analysis

Although at this stage we will not discuss in detail the ways in which data are analysed, it is important to at least have sight of the likely methods that may be used (see Chapter 5 for a more detailed coverage of statistical analysis). This is because different statistical methods are required to deal with different research questions. For example, if we were interested in trends over time, a regression model would seem appropriate; but if we were motivated to look for differences between treatments, then an ANOVA might be our test of choice. It should be noted that most statistical techniques require data to be gathered in a particular manner and so a lack of care at this stage could result in data being collected that cannot answer the question posed. It could also mean that some statistical methods cannot be used because the data do not meet the minimum number of observations that are needed to obtain meaningful results. This might also apply to those statistical approaches that require balanced designs (i.e. the same number of data points in each factor measured). In this section, we will discuss some of the major types of analyses that you could employ to answer certain commonly asked questions. As always, it is worth looking at the literature to see what types of analysis have been used in similar studies to the one you propose to do. There are several major groups of analysis based on the broad types of approach required.

Describing data

We need a variety of techniques to describe the data that we collect. This might be as a data exploratory technique (to check the data to see how variable a data set is, or what sort of distribution we get, etc.), to understand some aspects of the data (e.g. how diverse communities are), and for communication purposes (to be able to discuss the results, orally and in writing, with other people).

Here, simple plotting of measured variables on frequency histograms (or tables), cross-tabulation of one (nominal) variable against another, and examining the range of the data (from the minimum to the maximum) may help to check for errors and ensure that we can chose the correct type of test for subsequent analysis. Extracting statistics, such as diversity indices to describe species richness, or evenness to describe how equal the proportion of species is within a community, can be important to assess what sort of community we have. Likewise, the estimation of population size or density might also be important. Similarly, we can use the average value of a variable to describe the magnitude of the majority of the data points (usually in conjunction with some measure of how variable the values are and how many data points were collected). These descriptive statistical techniques are reviewed later and more detail can be found in Wheater and Cook (2000, 2015).

Asking questions about data

If we wish to ask specific questions of the data, then we are in the realm of inferential statistics. These usually involve the testing of hypotheses. It is standard practice to set up a null hypothesis alongside the questions to be asked. The null hypothesis tests the chance of there being no significant difference between samples (or relationship between variables, or association between categories of variables). So if we wish to know whether there is a difference between two samples (e.g. comparing the number of birds found in deciduous woodlands with the number found in coniferous woodlands),

Table 1.3 Common statistical tests.
Note that in each case, there are possible questions (and analyses) dealing with more than two samples and/or variables – see Chapter 5 for further details.

Example question	Null hypothesis	Type of test	Data required
Is there a difference between the number of birds found in deciduous woodlands and coniferous woodlands?	There is no significant difference between the number of birds in deciduous and coniferous woodlands.	Difference tests, e.g. a *t* test or a Mann–Whitney *U* test (p. 305).	Two variables: one nominal describing the woodland type and one based on either measurements (i.e. actual numbers) or on a ranked scale that describes the number of birds.
Is there a relationship between the number of birds and the size of the woodland?	There is no significant relationship between the number of birds and the size of the woodland.	Relationship tests, e.g. correlation analysis (p. 307).	Two variables: one (either measured or ranked) that describes the number of birds and one (either measured or ranked) that describes the size of the woodland.
Is there an association between whether birds are resident or not and whether the woodlands are deciduous or coniferous?	There is no significant association between the frequency of residency and the frequency of woodland type.	Frequency analysis, e.g. a Chi-square test (p. 312).	Two variables: one nominal describing the residency status of the birds and one nominal describing the woodland type.

then we actually test the null hypothesis that: *there is no significant difference between the number of birds in deciduous and coniferous woodlands.* Note that we are looking at 'significant' differences. These are differences that are unlikely to have resulted from random variation in the individual woodlands sampled. For this we need a method that tests the null hypothesis that there is no significant difference in the sample averages. In addition to difference tests between samples, there are also relationship tests between variables, and tests designed to examine associations between categories of variables. Table 1.3 summarises some commonly used, relatively simple, statistical approaches to these research questions.

Since there are various questions that we might ask as part of an investigation, it is important to be clear about possible analysis methods in advance of any sampling. The choice of test depends not only on the question being asked, but also on the data types being used. Where data are ranked, but not measured (i.e. ordinal data – p. 27) then a suite of tests called nonparametric tests may be used. The alternative (using parametric tests) is more robust and generally preferred, but requires data to be on a measurement scale (i.e. interval/ratio data). Therefore, it is usually an advantage to obtain measurement data rather than to rank data wherever possible. Even where measurements are taken, parametric tests may not be the most appropriate. This is because most parametric tests require the data to conform to a type of distribution called a normal distribution. Briefly, this is determined by examining histograms of the data (with the variable of interest plotted on the x axis and the frequency of its occurrence on the y axis) to see whether they have a symmetrical pattern (see Figure 1.6). For further details about the shape of distributions, and of which test to use, see Chapter 5. There are also different tests depending whether the data are matched or unmatched (p. 305).

To illustrate some of the considerations in project design and data collection, we start with a research question that sounds relatively simple on the face of it: is there a relationship between the size of trees and the number of squirrels' dreys in the canopy of the trees? Ideally, we would want to measure the canopy height with some degree of accuracy. This would enable us to work out whether the relationship exists using a parametric statistical technique called Pearson's product moment correlation analysis (p. 308). However, it may be difficult even to see the tops of very tall trees and those obscured by other trees. Thus, we may estimate tree height, perhaps into several groupings. We can of course rank these data, but this means that we need an alternative approach for analysis that is suitable for ordinal data. This is Spearman's rank correlation coefficient analysis, which is not quite as powerful as the Pearson's method. The power of the test is its ability to detect a true relationship (or difference, or association) if one exists. If we knew that any such relationship was likely to be fairly weak, then the less powerful technique might not reveal it and we could be wasting our time in not measuring the trees relatively accurately to obtain measurement data and thus employ the more powerful test. Alternatively, if we are only interested in revealing strong relationships, then using ranked size classes to indicate tree height may be acceptable. The other complexities in this apparently simple question include ensuring that all other aspects are as constant as possible (e.g. species of tree, surrounding landscape, density of the squirrel colony, etc.).

Predictive analysis

We may wish to collect data to set up a model that enables us to predict the outcome in a hypothetical situation, one of the simplest of which is known as a linear regression

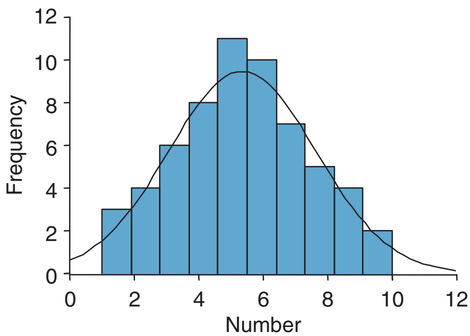


Figure 1.6 **Data set approximating to a normal distribution.**

model. Thus, if we are interested in looking at a possible relationship between woodland size and the number of birds and knew that this was likely to produce a significant linear relationship, then we may wish to use this fact to calculate the expected number of birds found in any woodland. This could be used theoretically or in conservation management to check that we have the sort of bird biodiversity that we expect from other data. Here, it is important to note that any such prediction should only be made if the woodland area in which we are interested lies between the minimum and maximum value of the data set we used to establish the model. We first need to establish which variable is the dependent and which is the independent variable: that is, which is likely to be affected (the dependent or response variable – plotted on the y axis of a scatterplot) by the other (independent variable – plotted on the x axis of a scatterplot). Here, obviously, the number of birds (the dependent variable) is more likely to be dependent on the size of the woodland (the independent variable) than vice versa. We can think of this as woodland size driving the size of the bird count. We can extend the technique to cover the case where there are a number of independent variables (e.g. woodland area, habitat diversity, area of associated green space, distance to nearest waterbody) that might influence or drive bird numbers.

Multivariate analysis

Where the question to be asked is a complicated one involving a number of dependent and/or independent variables, then multivariate analyses may be appropriate. The choice of analysis depends on whether the dependent variable is a category, or a ranked or measured variable, and on whether the independent variables are categories, ranked, or measured (or even a mixture). Whilst most such analyses only have one dependent variable, there may be multiple independent variables. For example, we may want to know whether the number of birds differs in different types of woodland when we take into account the woodland size (measured variable), woodland type (nominal variable), distance to the nearest neighbouring woodland (measured variable), age of woodland (measured variable), and the land use type surrounding the woodland (nominal variable). Here, we could enter all of the data into one analysis that would take into account the interrelationships between each variable and produce a model describing the relative importance of each variable on the number of birds (this particular example could be analysed using a generalized linear model – p. 319). Such techniques are powerful but require a full understanding of the data or data set and its attributes and may be quite complex to interpret.

Examining patterns and structure in communities

Ecological data sets can be very complex and difficult to visualise. For example, a data set might include many variables collected as measurements (including counts), as ranks (e.g. scores of abundance), or in a binary form (e.g. presence or absence data). Chapter 5 introduces a number of techniques for visualising complex data sets to enable the use of a range of different types of data. Variables with large numbers of observations of zero (as can occur when surveying relatively rare species), cases where data are heavily skewed, or situations where variables are measured on scales of greatly differing magnitude, may require data transformation before using these techniques (p. 285).

As an example, we might collect information about woodlands on the basis of their size, age, distance to the nearest neighbouring woodland, etc. Since some of these variables will be related to each other, we might wish to find out the underlying pattern of interrelationships within the data and hence identify a number of unrelated factors that can be used instead of our large number of variables. This is a data reduction exercise, reducing the number of variables we have measured into a smaller number of unrelated factors that take into account the interrelationships between the variables.

Alternatively, we might wish to look at the range of species found in each of several woodlands and see which woodlands have similar species types. This is a similarity or clustering analysis and, depending on the technique used to calculate the similarities, data are normally recorded as a matrix (of species by woodlands) that contains either measurements (e.g. counts), ranks (e.g. ranked abundance), or binary data (e.g. species presence or absence). A similar technique to clustering enables us to visualise patterns in either the individuals (in this example, the woodlands) and/or the variables (here, the types of species). This is known as ordination and there are a number of different methods available depending on the algorithm (i.e. statistical formula). Such methods can utilise data comprising measurements, ranks, or binary information.

Summary

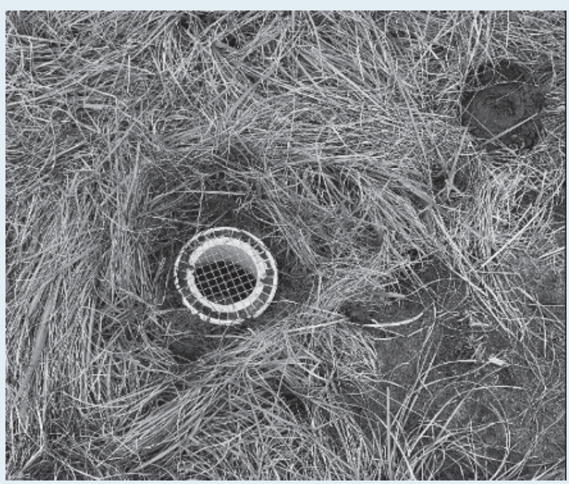
This chapter has identified the range of aspects that should be carefully considered when planning your project. Case Study 1.3 describes how one researcher approached her work, using the current literature to develop her techniques and adopting appropriate protocols for working overseas, in potentially hazardous environments. Ensure that you take sufficient time in the planning phase of your research project to cover all of the component parts. This includes health and safety and legal issues as well as making sure that your aims and objectives are focused and that any methods employed are appropriate to gather and analyse data. At each stage, consider the details of the implementation, whether this is in the practicalities of sampling or data management. Box 1.9 gives some general guidelines that should be ticked off in advance of implementing your project.

Case Study 1.3 Monitoring dung beetle richness in East Africa

***Phanaeus vindex* in feral hog dung.**

Dr Roisin Stanbrook is a Post-doctoral fellow at the University of Central Florida. She uses dung beetles to assess the impact of habitat modification on dung beetle communities, and models how changes in mammalian diversity in protected areas in East Africa affect dung beetle community composition. She has long held a passion for the little things that run our world, and for the need to increase awareness of insect conservation. This case study examines the different aspects of using baited pitfall trapping, which is most commonly employed method to determine dung beetle diversity, and the challenges encountered when working in protected areas in East Africa.

Model organism and research challenges faced



A gridded pitfall trap baited with cow dung.

Scarabaeinae, or true dung beetles, comprise one of the most species-rich groups of coprophagous insects with almost 7000 species found globally. Dung beetles have a widespread distribution, which is centred around the species-rich tropics, with diversity gradually declining in higher latitudes. This ubiquity means that dung beetles are often used as indicators of habitat quality and barometers of habitat change due to species' associations with habitat types, and reductions in diversity when habitats are negatively affected by disturbance. One of the benefits of using dung beetles as an eco-indicator is the ease with which large data sets can be gathered in a short time-frame, and consequentially used to form a rapid impact assessment (Bicknell et al. 2014). However, there are many challenges one may encounter when used baited pitfall traps.

One major issue with pitfall trapping is obtaining an accurate relative sample of dung beetle diversity. Without adequate sampling effort, dominant species are often over-represented and rare species may be missed entirely. Traps frequently generate large numbers of common species over a short time period, but longer trapping periods of up to one month may be needed in species-rich tropical locations where dung beetle niches are smaller. Factors such as the quantity of moonlight, or seasonal shifts in temperature could determine if rare species are detected or not. Another common issue when using baited pitfall traps is trap disturbance by other animals. Once traps are set, they are left unattended until the observer returns. This allows inquisitive species such as hyenas (*Crocuta crocuta*), or animals such as mongooses (e.g. *Mungos mungo*) which like to feed on captured beetles, ample time to disrupt trapping effort and consume valuable data.

Photos RAS

Collection of dung beetles from a baited pitfall trap ready to be sorted for identification.

Once trap data have been collected, the samples need to be sorted and identified to species or morphospecies. This typically requires a high-powered microscope and access to many intricately written identification keys. Identifying dung beetles to species level requires time and tenacity, and access to existing collections to verify identifications can be beneficial. This can be challenging when collecting species internationally, but can be circumnavigated by using collections belonging local or national museums.

How the challenge was resolved

A dung beetle project's success will largely depend on the amount and length of time pitfall traps are deployed. It is recommended that a pilot study be conducted in the same location, and during the same season, for at least one entire trapping effort. Traps should be set and collected daily for a minimum of 8 days. A species accumulation curve should then be constructed. Species accumulation curves are easy to construct and are readily interpretable to give the researcher an indication of the minimum sampling effort required for an adequate inventory, which avoids incurring extra time and fieldwork costs.

Standard methods for pitfall traps usually employ dung suspended over a bucket which is placed into the ground with the lip of the bucket flush with the soil surface. One method of avoiding trap loss by other animals is to use pitfall traps with built-in mesh enclosures. The trap is constructed from a one-gallon sturdy container, such as a paint can with the lid placed on the container but the centre of the lid removed and replaced with strong chicken wire. This design means that even if the trap is disturbed, captured beetles remain inside the container and the sample will not be lost.

Sorting and identifying dung beetles to species can be tricky. Forging strong collaborations with scientists and museums in-country allows access to holotype specimens found in their collections, and if export for further analyses are required these collaborations will ensure that the export process conforms with both national and international law. Roisin worked closely with National Museums Kenya in Nairobi and deposited specimens in their collection to leave a permanent record of Kenyan dung beetle diversity for future dung beetle enthusiasts.

Advice for students wanting to study dung beetles

Dung baited pitfall traps are a relatively easy and enjoyable method of assessing habitat modification and its associated effects. There are, however, health risks associated with using animal dung as bait. The handling of dung and other contaminants expose scientists to an increased risk of gastrointestinal infection, and care must be taken to minimise exposure and employ strict hygiene routines. Regular handwashing in conjunction with the use of disposable gloves are often the minimum requirements to avoid infection.

Dung beetle species richness is highest in the tropical forests and grasslands of Africa. These habitats also contain the world's highest density large wild animal populations, and an increased awareness of working in such locations is necessary to collect meaningful data safely. Studying fantastic insects such as dung beetles can be incredibly worthwhile, especially as there are still numerous new species found every year (e.g. Roggero et al. 2017).

Box 1.9 Checklist for field research planning

☑

- ☐ Determine the question and formulate the aims, objectives, and hypotheses.
- ☐ Determine whether a manipulative experiment or observational survey would be the most appropriate method to use.
- ☐ Decide whether you are looking primarily at the presence or absence of a species, relative abundance (e.g. counts of organisms), absolute abundance (e.g. population sizes or densities), community structures, behavioural responses, etc.
- ☐ Determine the statistical analyses that are likely to be employed.
- ☐ Select the appropriate sampling technique, taking into account the intended statistical analysis.
- ☐ Decide on the taxonomic level for identification, or the appropriate ecological grouping you will use (see Chapter 2).
- ☐ Assess what types of data will be collected and produce a standardised recording sheet.
- ☐ Work out the sampling design, including sample sizes.
- ☐ Select appropriate sample sites.
- ☐ Determine the site characteristics that will be monitored and choose appropriate techniques for doing this.
- ☐ Obtain permission to use the sites and check any legal restrictions.
- ☐ Carefully consider the ethical implications of the study.
- ☐ Risk assess all the work to be carried out.
- ☐ Employ a pilot study and amend your protocol if necessary.