Chapter 8
Design Applications

8.1 Introduction

Our goal in this chapter is to provide guidance that will enhance the decision-making process throughout a given study. The often-stated axiom about "the best laid plans…" certainly applies to study design. The initial plan that any researcher, no matter how experienced, develops will undergo numerous changes throughout the duration of a study. As one's experience with designing and implementing studies grows, he or she anticipates and resolves more and more major difficulties before they have an opportunity to impact the smooth progression of the study. The "feeling" one gains for what will and what will not work in the laboratory or especially the field is difficult to impart; there is truly no substitute for experience. However, by following a systematic path of steps when designing and implementing a given study, even the beginning student can avoid many of the most severe pitfalls associated with research.

The guidelines we present below will help develop the thought process required to foresee potential problems that may affect a project. We present these guidelines in a checklist format to ensure an organized progression of steps. Perhaps a good analogy is that of preparing an oral presentation or writing a manuscript; without a detailed outline, it is too easy to leave out many important details. This is at least partially because the presenter or writer is quite familiar with their subject (we hope), and tends to take many issues and details for granted. The speaker, however, can always go back and fill in the details, and the writer can insert text; these are corrective measures typically impossible in a research project.

Throughout this chapter, we provide examples from our own experiences. We will focus, however, on a "theme" example, which provides continuity and allows the reader to see how one can accomplish a multifaceted study. The goal of the theme study is development of strategies to reduce nest parasitism by the brown-headed cowbird (*Molothrus ater*) on host species along the lower Colorado River (which forms the border of California and Arizona in the desert southwest of the United States). We provide a brief summary of the natural history of the cowbird in Box 8.1.

The theme example is multifaceted and includes impact assessment (influence of cowbirds on hosts), treatment effects (response of cowbirds to control measures),

Box 8.1

The brown-headed cowbird is a small blackbird, is native to this region, and has been increasing in both abundance and geographic range during the late 1800s and throughout the 1900s. This expansion is apparently due to the ability of the cowbird to occupy farm and rangelands and other disturbed locations. The cowbird does not build a nest. Rather, it lays its eggs in the nest of other species, usually small, open-cup nesting songbirds such as warblers, vireos, and sparrows. The host adults then raise the nestling cowbirds. Research has shown that parasitized nests produce fewer host young than nonparasitized nests. Although this is a natural occurrence, the population abundance of certain host species can be adversely impacted by parasitism. These effects are exacerbated when the host species is already rare and confined to environmental conditions favored by cowbirds. In the Southwest, host species that require riparian vegetation are the most severely impacted by cowbirds. This is because most (>90% along the lower Colorado River) of the riparian vegetation has been removed for river channelization and other management activities, thus concentrating both cowbirds and hosts in a small area.

problems in logistics (extremely hot location, dense vegetation), time constraints (at least one host is nearing extinction), and many other real-life situations. We added other relevant examples to emphasize salient points.

Numerous publications have developed step-by-step guides to study development (Cochran 1983; Cook and Stubbendieck 1986; Levy and Lemeshow 1999; Martin and Bateson 1993; Lehner 1996; Garton et al. 2005). We reviewed these publications and incorporated what we think are the best features of each into our own template (Fig. 8.1). Project development should follow the order presented in Fig. 8.1; below we follow this same progression. Not surprisingly, the steps we recommend for planning, implementing, and completing a given research project (Fig. 8.1) are components of the steps typically used during the process implementing a research program in natural science (Table 1.2), which includes multiple individual research projects. In previous chapters, we discussed in detail nearly all materials presented below. While presenting our examples, we refer you to previous sections of the book and reference key primary publications to facilitate review.

8.2 Sequence of Study Design, Analysis, and Publication

8.2.1 Step 1 – Questions

A study is designed around questions posed by the investigator (see Sect. 1.3). Although this statement may sound trivial, crafting questions is actually a difficult and critically important process. If a full list of questions is not produced before the

1. QUESTIONS
⇓
2. HYPOTHESES & PREDICTIONS
⇓
3. DESIGN RESEARCH
⇓
4. CHOICE OF VARIABLES
⇓
5. CHOICE OF RECORDING METHODS
⇓
6. ESTABLISH ACCEPTABLE LEVEL OF PRECISION
⇓
7. PRELIMINARY DATA COLLECTION
⇓
8. QUALITY/QUANTITY ASSURANCE
⇓⇑
9. FINAL DATA COLLECTION
⇓
10. ANALYSIS/HYPOTHESIS TESTING
⇓
11. INTERPRETATION
⇓
12. PUBLICATION

Fig. 8.1 Progression of steps recommended for planning, implementing, and completing a research project. Reproduced from Morrison et al. (2001), with kind permission from Springer Science + Business Media

study plan is developed, it is often problematic to either insert new questions into an ongoing study or answer new questions not considered until the conclusion of data collection. Thus, the first step should be listing all relevant questions that should be asked during the study. Then, these questions should be prioritized based on the importance of the answer to the study. Optimally, we then design the study to answer the aforementioned questions, in a statistically rigorous fashion, one question at a time. The investigator *must* resist the overpowering temptation to try to address multiple questions with limited resources simply because "they are of interest." Such a strategy will, at a minimum, doom a study to mediocrity. We would much rather see one question answered thoroughly than see a paper reporting the results of a series of partially or weakly addressed questions.

Guiding the selection of questions will be the availability of study locations, time, personnel, and funding. All studies will have limitations for all of these parameters, which place constraints on the number and types of questions that can be addressed.

Questions are developed using literature, expert opinion, your own experiences, intuition, and guesswork. A thorough literature review is an essential cornerstone of all studies. There is no need to reinvent the wheel; we should attempt to advance knowledge rather than simply repeating a study in yet another geographic location. Likewise, we should critically evaluate published work and not repeat biased or substandard research. We must be familiar with the past before we can advance into the future. The insights offered by people experienced in the field of interest always should be solicited, because these individuals have ideas and intuition often unavailable in their publications. You should synthesize all of these sources and develop your own ideas, blended with intuition, to devise your questions.

8.2.1.1 Theme Example

We developed the following questions by reviewing the literature, discussing research topics with cowbird researchers, considering the needs of resource agencies, and brainstorming among field personnel; no prioritization is implied by the order of questions.

1. What is the effectiveness of a single trap; i.e., the area of influence a trap exerts on parasitism and host productivity?
2. What are the movements of males/females by age within and between riparian patches?
3. What is the relative effectiveness of removing cowbird adults vs. removing their eggs and young?
4. When is the nesting cycle to remove eggs/young?
5. When is it most effective to trap adult cowbirds?
6. Is it more effective to remove adult females rather than both male and female cowbirds?
7. What is the trapping intensity of cowbirds within a year necessary to influence host productivity?
8. What is the best measure of project success (e.g., increased host reproductive success or fledgling success)?
9. What are the population abundance of cowbirds and host species?
10. Will winter cowbird trapping exert effective control of cowbirds?
11. Where do the cowbirds reside during winter?
12. Do predators significantly influence host-breeding success, thus effectively negating any positive effects of cowbird control?

Exercise: The preceding 12 questions are not arranged in any priority. As an exercise, recall the previous discussion describing the theme example. Then, on a separate sheet list what you would consider the top four questions in the priority that you would pursue them. Although there is no “correct” answer, in Box 8.2 we provide our list along with a brief explanation.

8.2.2 Step 2 – Hypotheses and Predictions

In Chaps. 1 and 2, we discussed the philosophical underpinnings of hypotheses testing, model selection, and related approaches to gaining knowledge. In his now classic paper, Romesburg (1981) argued that wildlife science was not advancing

Box 8.2

Here is our prioritization of the study questions for the theme example given in text under “Questions.”

#1 – h. What is an adequate measure of project success (e.g., reduced parasitism rates, increased reproductive success; fledgling success)?

This should be the obvious choice given the project goal of determining ways of lowering parasitism and increasing breeding success of cowbird hosts. The complicating issue here, however, is that simply lowering parasitism may not increase reproductive success. This is because other factors, most notably predation by a host of species, could negate any beneficial effects of adult or egg and young removal. Thus, while documenting a reduction in parasitism is essential, it alone is not sufficient to declare the study a success.

#2 – c. Effectiveness of removal of adults vs. removal of eggs and young

It is extremely time consuming to locate host nests. Thus, it would be easier to focus management on removal of adult cowbirds. It was deemed necessary to examine both methods because of the worry that removal of adults alone would be inadequate, and the necessity of rapidly determining an effective management strategy.

#3 – a. Effectiveness of a single trap; i.e., the area of influence a trap exerts on parasitism and host productivity

The failure to detect a significant influence of adult cowbird removal on host reproduction could be criticized as resulting from inadequate trapping effort. Therefore, it was judged to be critical to ensure that an adequate trapping effort be implemented. The decision was made to “overtrap” to avoid this type of failure; future studies could refine trapping effort if trapping was shown to be an effective procedure.

#4 – i. Population abundance of cowbirds and host species

The initial reason for concern regarding the influence of cowbird parasitism on hosts was the increase in cowbird abundance and the decrease in many host species abundance. Therefore, determining trends in cowbird and host abundance will serve, over time, to determine the degree of concern that should be placed on this management issue.

Not all questions are mutually exclusive, nor is it possible to produce a perfect ordering. However, it is critical that a short list of prioritized items be developed, and that this prioritization be based on the ultimate reason for conducting the study.

our ability to make reliable predictions because, in part, of the failure to follow rigorously the H–D method (see Sect. 1.4). However, testing hypotheses does not necessarily lessen the numerous ad hoc explanations that accompany studies that fail to establish any formal and testable structure. A decent biologist can always explain a result. Our point here is that simply stating a hypothesis in no way guarantees increased knowledge or revolutionary science (Kuhn 1962). Instead, what is required is careful and thoughtful evaluation of the predictive value of any proposed hypothesis. In designing a study, the researcher must carefully determine what is needed to advance management of the situation of interest. Questions are driven by goals, and hypotheses that follow from the questions are driven by the predictive power needed to answer the questions.

8.2.2.1 Theme Example

In our theme example, we developed specific research hypotheses directed at determining whether cowbird trapping could result in a biologically meaningful increase in host productivity.

The primary hypotheses to be tested are:

H_0: No difference in parasitism rates between treated and controls plots
H_1: Parasitism rates differ significantly between treated and control plots

Given rejection of the above two-tailed null hypothesis, the following one-tailed hypotheses were tested:

H_0: Parasitism rates on treated plots are ≤30%
H_1: Parasitism rates on treated plots are >30%

Previous research indicated those parasitism rates exceeding 25–30% caused breeding failure in hosts (Finch 1983; Laymon 1987; Robinson et al. 1993). Therefore, we determined that parasitism <30% would be an appropriate measure of success for this aspect of the study.

Next, the same set of hypotheses can be stated for examining nesting success (defined here as the number of young fledged/nest):

H_0: No difference in nesting success between treated and controls plots
H_1: Nesting success differs significantly between treated and control plots

Given rejection of the above two-tailed null hypothesis, the following one-tailed hypotheses were tested:

H_0: Nesting success on treated plots was ≥40%
H_1: Nesting success on treated plots was <40%

Establishing an appropriate measure of nest success is difficult. The literature indicates that success of >40% is a general measure of population health of many passerine nests (Martin 1992). You will find that choosing an appropriate magnitude of difference that has biological relevance is one of the most difficult aspects of study design. But without setting a biological magnitude, you are allowing your study be driven by statistics (i.e., the test statistic and associated *P*-value), and then a posteriori trying to establish biological explanations. That is, you are letting the statistics lead the study; rather, you want your study to be led by biology as supported by statistics. Again, refer to Chaps. 1 and 2 for a thorough discussion of these design issues.

8.2.3 Step 3 – Design

Design of the project entails the remaining sections of this chapter. Indeed, even the final section on publishing should be considered in designing your study (see Sect. 1.3.1). For example, a useful exercise when designing your study is to ask, “How will I explain this method when I write this up?” If you do not have absolute confidence in your answer, then a revision of methods is called for. Put more bluntly, “If you cannot explain it, how can you expect a reader to understand it?” A well-written study plan/proposal can become, in essence, the methods section of a manuscript.

Imbedded in our previous discussion of question and hypotheses development are issues related to delineation of the study population, spatial and temporal extent of study, sex and age considerations, needed generalizability of results, and so forth. Each of these separate issues must now be placed into a coherent study design. This requires that we make decisions regarding allocation of available resources (personnel, funding, logistics, and time). At this step, it is likely that any previous failure to reduce the number of questions being asked will be highlighted.

Delineation of the study population is a critical aspect of any study (see Sects. 1.3–1.5). Such delineation allows one to make definitive statements regarding how widely study results can be extrapolated. Unfortunately, few research papers make such a determination. A simple example: The willow flycatcher (*Empidonax traillii*) is a rare inhabitant of riparian and shrub vegetation throughout much of the western United States (Ehrlich et al. 1988). It is currently divided into three subspecies, one of which (*E. t. extimus*) is classified as threatened/endangered and is restricted to riparian vegetation in the arid Southwest. The other two subspecies (*E. t. brewsteri* and *E. t. adastus*) occur in higher elevation, mountain shrub (especially willow, *Salix*) and riparian vegetation; they are declining in numbers but not federally listed (Harris et al. 1987). Thus, it is unlikely that research findings for *E. t. brewsteri* and *E. t. adastus* would be applicable to *E. t. extimus*. Further, it is unlikely that results for studies of *E. t. extimus* from the lower Colorado River – a desert environment – would be applicable to populations of this subspecies occurring farther north in less arid regions. Subspecies are often divided into ecotypes that cannot be separated morphologically. Indeed, the physiological adaptations of ecotypes are little studied; differences among ecotypes (and thus populations) are probably a leading reason why research results can seldom be successfully applied to other geographic locations.

The distribution of sampling locations (e.g., plots, transects) is a critical aspect of all studies (see Chap. 4). First, the absolute area available to place the plots is usually limited by extent of the vegetation of interest, legal restrictions preventing access to land, and differences in environmental conditions. If the goal of a study is to determine effects of a treatment in riparian vegetation on a wildlife refuge, then the study is constrained by the extent of that vegetation on a defined area (as is the case in our theme example).

Second, the behavior of the animal(s) under study will constrain the placement of sampling locations. Most studies would be severely biased by movement of animals between what are designated as independent sampling locations. For example, avian ecologists often assess the abundance of breeding passerines by placing counting points 200–300 m apart. This criterion is based on the argument that most passerines have breeding territories of <100-m radius, and that it is unlikely that individuals will move between counting points during the short period of time (usually 5–10 min) that an observer is present. These arguments are suspect, however, given that breeding birds move beyond territory boundaries, and adjacent points are likely being influenced by similar factors (e.g., disturbances, relatively microclimatic conditions, predators, competitors). The placement of small mammal trapping plots, the locations used to capture individuals for radio tagging, the selection of animals for blood sampling, and the gathering of pellets for food analysis are but a few of the other sampling decisions that, are difficult to conduct with independence (see Sect. 2.3).

The time and money available to conduct a study constrains study design. It is critical that you do not try to conduct a study that overextends available resources. You do not want to place yourself in the position of using "time and money constraints" as an excuse for not achieving reliable research results (e.g., small sample sizes).

8.2.3.1 Theme Example

The optimal design for this study would have been based on at least 1 year of pretreatment data collection, followed by random assignment of treatments (adult and/or egg–young removal). However, funding agencies required a more rapid assessment of treatment effects because of pressure from regulatory agencies. Thus, the decision was made to forego use of a BACI-type design, and instead use another impact assessment approach (see Chap. 6). Specifically, pairs of plots were selected based on similarity in environmental conditions and proximity to one another. Treatments were then assigned randomly to one member of each pair. Paired plots were placed in close proximity because of the lack of appropriate (riparian) vegetation. Because of limited riparian vegetation, it was not possible to have control plots for adult cowbird removal that were within a reasonably close proximity to treated plots. This is because previous research indicated that a single cowbird trap could impact parasitism rates up to at least 0.8 km from the trap. The decision was made to treat all plots for adult removal, but only one of the pairs would be treated for egg and young removal. Plots receiving no treatments were placed upstream far outside the influence of the cowbird traps. This design carries the assumption that the upstream reference plots will be an adequate indicator of naturally occurring parasitism rates in the area. Research on parasitism rates conducted several years prior to the new study provided additional support. Thus, the design implemented for this study is nonoptimal, but typical of the constraints placed on wildlife research. Textbook examples of experimental design are often impossible to implement in the real world of field biology (although this is not an excuse for a sloppy design; see Sect. 8.3). Thus, a certain amount of ingenuity is required to design an experiment that can still test the relevant hypotheses. Of course, some hypotheses cannot be tested in the field regardless of the design (see Chap. 1).

8.2.4 Step 4 – Variable Selection

Variables must be selected that are expected to respond to the treatment being tested, or be closely linked to the relationship being investigated. Even purely descriptive, or hypothesis generating, studies should focus sampling efforts on a restricted set of measurements. A thorough literature review is an essential part of any study, including development of a list of the variables measured by previous workers (see Sect. 1.3). However, one must avoid the temptation of developing a long "shopping list" of variables to measure: This only results in lowered precision because of smaller sample sizes. Rather, a list of previous measurements should be used to develop a short list of variables found previously to be of predictive value. Measuring numerous variables often means that some will be redundant and measure the same quality (e.g., tree height and basal area). Also, lists of variables from past studies may provide an opportunity to learn from past efforts; that is, there may be better variables to measure than those measured before.

For example, since multivariate statistical tools became readily available and easy to use, there has been a proliferation of studies that collected data on a massive number of variables. It is not unusual to see studies that gathered data on 20–30 variables or more, or read analyses of all possible subsets of ten or more variables (which results in millions of comparisons); this tendency is especially apparent in studies of wildlife–habitat relationships (see Morrison et al. 2006 for review). Preliminary data collection (see Sect. 8.2.7) can aid in developing a short list of variables. Most studies, including those using multivariate methods, identify only a few variables that have the majority of predictive power. There is ample evidence in the literature to justify concentrating on a minimum number of predictive or response variables.

Each variable measured increases the time spent on each independent sample, time that could be better spent on additional independent samples. Further, researchers often tend to use rapid measurement techniques when confronted with a long list of variables. For example, many workers visually estimate vegetation variables (e.g., shrub cover, tree height) even though research indicates that visual methods are inferior to more direct measurements (e.g., using a line intercept for shrubs and a clinometer for tree height; see Block et al. 1987). Thus, there is ample reason to concentrate on a few, carefully selected variables.

This is also the time to identify potential covariates (see Chap. 3). Analysis of covariance is an indirect, or statistical, means of controlling variability due to experimental error that increases precision and removes potential sources of bias. As a reminder, statistical control is achieved by measuring one or more concomitant variates in addition to the variate of primary interest (i.e., the response variable). Measurements on the covariates are made for adjusting the measurements on the primary variate. For example, in conducting an experiment on food preference, previous experience will likely influence results. The variate in this experiment may be a measure of the time taken to make a decision; the covariate may be a measure associated with the degree of experience at the start of the trials. Thus, covariates must be designed into the study, and should not be an afterthought. Covariates may be used along with direct experimental design control. Care must be taken in the use of covariates or misleading results can result (see Chap. 3 (see also Winer et al. 1991, Chap. 10). Step 5 (Sect. 8.2.5) incorporates further discussion of variable selection.

8.2.4.1 Theme Example

One of our research interests is the movement of males and females by age within and between riparian patches. The process of selecting variates to measure begins with prioritization of the information needed to address project goals. Where and when the birds are located and their activity (e.g., foraging, nest searching) is important because this provides information useful in locating traps and searching for host nests that have been parasitized. The sex composition of the birds (such as in flocks) is important because it provides information on (1) how the sex ratio may be changing with trapping and (2) where females may be concentrating their activities. The age composition also is important because it provides a measure of experimental success (i.e., how many cowbirds are being produced). Finally, the type of vegetation used by cowbirds provides data on foraging and nest-searching preferences.

An important aspect of the cowbird study is the effectiveness of the removal of adults in controlling parasitism. Confounding this analysis could be the location of a host nest within a study plot. This is because research has shown that cowbirds seem to prefer edge locations (e.g., Laymon 1987; Robinson et al. 1993; Morrison et al. 1999), with a decreasing parasitism rate as you move farther into the interior of a woodland. Thus, the variate may be a measure of host nesting success or parasitism rate, with the covariate being the distance of the nest from the plot edge.

8.2.5 Step 5 – Recording Methods

There are usually many methods available for recording data on a specific variable. For example, birds can be counted using fixed area plots, variable distance measures, and transects; vegetation can be measured using points, plots of various sizes and shapes, and line transects; and so forth (see Chapter 4). However, many of these decisions will be guided by the objectives of the study and, thus, the specific types of data and precision associated with the data needed. For example, if study objectives require that nesting success is determined, then an intensive, site-specific counting method (such as spot mapping) might be most appropriate. This same method, however, would not necessarily be appropriate for a study of bird abundance along an elevational gradient. Thus, there is no "best" method; rather, there are methods that are most appropriate for the objectives of any study.

It is also important that all proposed methods be thoroughly reviewed for potential biases and degree of precision attainable (see Sect. 2.6.4). Unfortunately, there is often little guidance regarding how to handle bias in publications reporting implementation of a method. It is usually necessary to go back to the original publication that reported on development of the method, or to search for the few papers that have analyzed bias. Within a single category of methods, such as "vegetation sampling" or "bird counting," there are numerous submethods that carry different levels of bias. For example, measuring tree height carries a much higher level of bias than does measuring tree dbh (Block et al. 1987). Similarly, the variable circular plot counting method for birds requires estimation of the distance from observer to bird and often determination of species identification by sound alone. Precision of the distance estimate declines with distance to the bird, and can be highly variable among observers (Kepler and Scott 1981; Alldredge et al. 2007).

Designing data forms is not a trivial matter and includes four distinct steps (Levy and Lemeshow 1999):

1. Specify the information to be collected
2. Select the data collection strategy
3. Order the recording of data
4. Structure the recording

We developed item 1 in Sect. 8.2.4, and item 2 earlier in this section. Below we develop the remaining two items.

8.2.5.1 Order the Recording

Once the information to be recorded is determined and the collection strategy established, the next step is to group the items in some meaningful and efficient manner, and order the items within the groups. Developing an efficient recording order greatly simplifies data collection, thus reducing observer frustration and increasing the quantity and quality of the data. For our theme example, some items refer to an individual bird, some to the flock the birds may be in, some to the foraging substrate, some to the foraging site, and so on. Items within each group are then ordered in a manner that will maximize data collection. For example, items related to the foraging individual include:

- Age
- Sex
- Plant species the bird is in or under
- Substrate the bird is directing foraging upon
- The behavior of the bird
- Rate of foraging

8.2.5.2 Structuring the Recording

This is the "action plan" for data collection. This plan includes three specific parts:

1. Sampling protocol
2. Data form
3. Variable key

In Box 8.3 we present an example of these three steps that was developed for a study of the distribution, abundance, and habitat affiliations of the red-legged frog (*Rana aurora draytonii*).

Box 8.3

The material given below outlines an example of the three primary products necessary to organize properly a data-collecting strategy (1) sampling protocol, (2) variable key, and (3) data form.

1. Sampling Protocol

This protocol is designed to quantify the habitat use of the red-legged frog (RLF) at three hierarchical scales. It is focused on aquatic environments, although the species is known to use upland sites during certain periods of the year. The hierarchy used is based on Welsh and Lind (1995). Terrestrial sightings are accommodated in this protocol in a general fashion, although more specific information would need to be added at the microhabitat scale than is provided herein.

The landscape scale describes the general geographic relationship of each reach sampled. Additional information that could be recorded in separate notes includes distance from artificial water containments; the distance to, and type of, human developments; and so forth. The appropriate measure of the animal associated with this broad scale would be presence or absence of the species. The macrohabitat scale describes individual segments or plots stratified within each reach by general features of the environment. The appropriate measure of the animal associated with this midscale is abundance by life stage. The microhabitat scale describes the specific location of an egg mass, tadpole, or adult.

All data should be recorded on the accompanying data form using the indicated codes. Any changes to the codes, or addition to the codes, must be indicated on a master code sheet and updated as soon as possible.

The following table summarizes the hierarchical arrangement of sampling used to determine the distribution, abundance, and habitat affinities of red-legged frogs, although the general format is applicable to a variety of species (see Welsh and Lind 1996). Variables should be placed under one classification only.

(continued)

- I. *Landscape Scale* (measured for the reach)
  - A. Geographic relationships
    1. UTM coordinates
    2. Elevation (m)
    3. Slope (%)
    4. Aspect (degrees)
- II. *Macrohabitat Scale* (measured for the segment or plot)
  - A. Water quality (average values in segment or plot)
    1. Depth (cm)
    2. Flow (cm s^{-1})
    3. Temp. (°C)
    4. Salinity (ppt)
  - B. Site description
    1. Size
       - a. Width
       - b. Length (m), if applicable
    2. Predominant condition
       - a. Stream
         - (1) Pool
         - (2) Glide
         - (3) Riffle
       - b. Seep
       - c. Marsh
       - d. Upland
         - (1) Grassland
         - (2) Shrubland
         - (3) Woodland
    3. Sediment
       - a. Coarse
       - b. Medium
       - c. Fine
  - C. Predominant aquatic vegetation
    1. Type 1 (open water with no vegetation)
    2. Type 2 (algae, flotsam, ditchgrass, low herbs and grass)
    3. Type 3 (tall, vertical reed-like plants [cattail, rush, sedge])
    4. Type 4 (live and dead tangles of woody roots and branches [willow, cotton wood, blackberry])
  - D. Stream covering
    1. % stream covered by overhanging foliage

(continued)

Box 8.3 (continued)

E. Red-legged frogs
  1. Egg mass present
  2. Tadpole present
  3. Adult present

F. Predators and competitors
  1. Predatory fish present
  2. Bullfrogs present (adults or tadpoles)

G. Other herpetofauna

III. Microhabitat Scale (measured for the animal location or subsegment/subplot)

A. Number of red-legged frogs
  1. Egg mass
  2. Tadpole
  3. Mature:
    a. Small (<60 mm snout–urostyle length)
    b. Large (≥60 mm)

B. Animal location:
  1. Land (>50 cm from water edge)
  2. Bank (within 50 cm of water)
  3. Shore (within 50 cm of land)
  4. Water (>50 cm from land)
    a. Depth (cm)

C. Specific animal location
  1. Site description
    a. In litter
    b. Under bank
    c. Under woody debris
  2. Distance from water (m)

D. Predominant aquatic vegetation (if applicable)
  1. Cover (%) by species
  2. Vigor of II.C.1. (live, declining, dead)
  3. Height (cm)

E. Predominant terrestrial vegetation (if applicable)
  1. Cover (%) by species
  2. Vigor of II.C.1. (live, declining, dead)
  3. Height (cm)

F. Aquatic substrates
  1. Cover (%) coarse sediment (rocks, boulders)
  2. Cover (%) medium sediment (gravel, pebble)
  3. Cover (%) fine sediment (sand, silt)

2. Variable Key

The following is a measurement protocol and key for the data form given below. Each variable on the data form is described on this key. In field sampling, changes made to the sampling procedures would be entered on a master key so that data forms can be revised and procedures documented (which is especially important in studies using multiple observers).

I. Landscape Scale

A. Take measurements from a central, characteristic location in the reach.

II. Macrohabitat Scale

Each reach will be divided into sections or plots based on some readily identifiable landscape feature, or some feature that is easy to relocate (e.g., bridge, road, house).

A. Take the mean of measurements needed to characterize the segment (1–3 measurements depending on size of segment)

B. 1. a. Average width of predominant condition given in 2 below
b. Length for nonstream condition
2. Predominant condition: record single predominant condition in the segment
3. Sediment: coarse (rocks, boulders); medium (gravel, pebble); fine (sand, silt)

C. Record the indicated "type" of predominant aquatic vegetation

D. Record the percentage of the stream within the segment that is obscured (shaded) by live foliage of any plant species

E. The goal is to note if any life stage of RLF is present; approximate numbers can be recorded if time allows (code 1 = 1 or a few egg masses or individuals; code 2 = 5–10; code 3 = >10).

F. The goal is to note the presence of any predator or competitor; approximate numbers can be recorded as for E above

III. Microhabital Scale

Individual animals that are selected for detailed analysis, or all animals within subsegments or subplots, can be analyzed for site-specific conditions.

A. The goal is to accurately count all RLF by life stage if working within a subsegment or subplot; or record the specific life stage if sampling microhabitat for an individual.

B. Record the location of the life stage by the categories provided (1–4); also record water depth (4.a.) if applicable.

C. Within the location recorded under B above, record the specific site as indicated (l.a. to c.); also include the distance from water if applicable (2.).

D. 1. Record the estimated percentage cover of the four predominant aquatic plant species within the subsegment/subplot, or within a 1-m radius of the individual RLF; also record the vigor (D.2.) and height (D.3.) of each plant species.

(continued)

Box 8.3 (continued)

E. As for D above, record for terrestrial vegetation. Both D and E can be used if the subsegment or subplot for the individual includes both aquatic and terrestrial vegetation (e.g., if animal is on a bank).

F. Record the approximate percentage cover for each of the indicated substrates.

3. Data Form

RED-LEGGED FROG SURVEY
DATE ___ ___ ___ ___ ___ TIME ___ ___ ___ ___ ___
DRAINAGE___ ___ LOG ___ ___ ___ OBS ___ ___; ___ ___ ___
I. LANDSCAPE:
UTM ___ ___ ___. ___ ___ ___; ___ ___ ___. ___ ___ ___
ELE ___ ___ ___ ___ SLOPE ___ ___ ASPECT ___ ___ ___
II. MACROHABITAT: SITE ___ ___ ___ ___
A. WATER QUALITY: DEPTH ___ ___ ___
FLOW ___ ___ TEMP ___ ___
B. SITE: SIZE: WIDTH ___ ___ ___ LENGTH ___ ___ ___ CONDITION ___ ___ SEDIMENT ___ ___
C. VEG TYPE ___ D. COVER ___ ___ ___ E. RLF ___ ___ ___ F. PRED ___ ___ ___
G. HERPS: ___ ___; ___ ___; ___ ___; ___ ___
III. MICROHABITAT: SITE ___ ___ ___
A. RLF: EGG ___ ___ ___ TAD ___ ___ ___ AD: SM ___ ___ ___ LG ___ ___ ___
B. LOC ___ ___ If LOC = ‘W’, record DEPTH ___ ___ ___
C. SPECIFIC ___ ___ ___ DIST ___ ___ ___
D. AQ VEG: COVER ___ ___ ___ VIGOR ___ HT ___ ___ ___
E. TERR VEG: COVER ___ ___ ___ VIGOR ___ HT ___ ___ ___
F. SUB: CO ___ ___ ME ___ ___ FI ___ ___
NOTES:

8.2.5.3 Theme Example

Suppose we have indeed decided to record details on foraging birds, their movements, and related topics. Referring back to the section on Step 1 (Sect. 8.2.1) – Questions, we might ask include:

1. Estimate the use of foraging areas
2. Determine if use varies within and between seasons and years, and age and sex
3. Determine the energy expenditures in various crop types

From these three objectives, we might extract the following inventory list of information needs:

1. Characteristics of the bird: age, sex, foraging location, foraging rates
2. Characteristics of the foraging location: the species, size, and health of the plants used for foraging
3. Characteristics of the foraging region: plant species composition, plant density by size and health categories; slope, aspect, distance to water
4. Characteristics of the environment: weather, both regional and foraging site; wind, solar radiation

Note that each of the inventory items is needed to meet the objectives. For example, weather conditions and foraging rates are needed, in part, to calculate energy expenditures.

Data Collection Strategy: There are numerous means of recording foraging behavior. The method chosen must allow analyses that meet objectives. For example, following a single individual for long periods vs. recording many individual behaviors simultaneously; collecting data during inclement weather vs. fair weather; recording data on breeding vs. nonbreeding animals. These choices determine the types of statistical analyses that are appropriate.

8.2.6 Step 6 – Precision and Error

The critical importance of determination of adequate sample sizes and power analyses was developed in Chaps. 2.6.6 and 2.6.7, respectively (see also Thompson et al. 1998, Chap. 6). A priori power analyses are often difficult because of the lack of sufficient data upon which to base estimates of variance. Nevertheless, power calculations remove much of the arbitrariness from study design (power calculations are performed again at various times during the study, as discussed in Sect. 8.2.7). In addition, power calculations give a good indication if you are trying to collect data on too many variables, thus allowing you to refocus sampling efforts on the priority variables (see Steidl et al. [1997] for a good example of application of power analysis to wildlife research).

8.2.6.1 Theme Example

To determine the number of paired plots (treatment vs. control) needed to rigorously test the null hypothesis of no treatment effect, data on nesting biology collected during several previous years of study were available. Although the data were not collected from plots, they were collected within the same general areas. Thus, the data were collected in similar environmental conditions as would be encountered during the new study. If these data had not been available, then information from the literature on studies in similar conditions and with similar species would have been utilized.

For example, the data available on nesting success of Bell's vireos, a species highly susceptible to cowbird parasitism, indicated that nesting success of parasitized nests was 0.8 ± 1.3 (SD) fledglings/nest. Using this value, effect sizes of 0.5, 1.0, and 1.5 fledglings/nest were calculated with power of 80% and $\alpha = 0.1$. Assuming equal variances between plots, and using calculations for one-sided tests (because primary interest was in increased nest success with cowbird control), sample sizes (number of plots) were calculated as >>20 (power = ~50% for n = 20), 15, and 7 for the three effect sizes. Thus, we concluded that time and personnel availability allowed us to identify an effect equal to 1.5 additional fledgling per nest following treatment at 80% power.

An additional important and critical step in development of the study plan is independent peer review (see Sect. 1.3.1). Included should be review by experts in the field and experts in technical aspects of the study, particularly study design and statistical analyses. Review prior to actually collecting the data can help avoid, although not eliminate, many problems and wasted effort. In this example, experts from management agencies helped confirm that the questions being asked were relevant to their needs, several statisticians were consulted on the procedures being used, and several individuals studying bird ecology and cowbirds reviewed the design.

8.2.7 Step 7 – Preliminary Analyses

All studies should begin with a preliminary phase during which observers are trained to become competent in all sampling procedures. The development of rigid sampling protocols, as developed above (see Sect. 8.2.5) improves the chances that observers will record data in a similar manner. Training of observers should include:

1. *Testing of visual and aural acuity*. Much wildlife research involves the ability to see and hear well. For example, birds can produce calls that are near the limits of human hearing ability. Slight ear damage, however, can go unnoticed, but result in an inability to hear high-frequency calls. Hearing tests for personnel who will be counting birds using sight and sound should be conducted (e.g., Ramsey and Scott 1981).
2. *Standardization of recording methods*. Unfortunately, there is seldom a correct value with which we can compare samples. Most of our data represent indices of some "true" but unknown value (e.g., indices representing animal density or vegetation cover). Further, field sampling usually requires that the observer interpret a behavior or makes estimate of animal counts or plant cover. Thus,

there is opportunity for variation among observers, which in turn introduces variance into the data set. Training observers to ensure that they collect data in a standardized and consistent manner is thus essential in reducing variance. Studies such as those conducted by Block et al. (1987) and papers in Ralph and Scott (1981, pp. 326–391) are examples of the value of observer training.

Initial field sampling should include tests of data collection procedures; often called *pretesting period*. Such pretesting allows for redesign of data forms and sampling protocols. Pretesting sampling should cover as much of the range of conditions that will be encountered during the study. Some, but seldom all, of the data collected during pretesting might be suitable for inclusion with the final data set.

Levy and Lemeshow (1999) differentiated between the pretest and the pilot survey or pilot study, with the latter being described as a full-scale dress rehearsal. The *pilot study* includes data collection, data processing, and data analyses, and thus allows thorough evaluation of all aspects of the study including initial sample size and power analyses. Thus, a pilot study is often done with a much larger sample than a pretest. Such studies are especially useful when initiating longer term studies.

8.2.7.1 Theme Example

An essential part of the theme study was determining the abundance of cowbirds and potential host species. Because the majority of birds encountered during a formal bird count are heard but not seen, observers must possess advanced identification skills and excellent hearing capabilities. Selecting experienced personnel eases training of observers. Although a talented but inexperienced individual can usually learn to identify by song the majority of birds in an area within 1–2 mo, it takes many years of intensive study to be able to differentiate species by call notes, including especially rare or transient species. Thus, observers should not be asked to accomplish more than their skill level allows.

Even experienced observers must have ample time to learn the local avifauna; this usually involves 3–4 weeks of review in the field and the use of song recordings. Regardless of the level of experience, all observers must standardize the recording of data. In the theme study, observers went through the following steps:

1. During the first month of employment, new employees used tapes to learn songs and call notes of local birds. This included testing each other through the use of tapes. Additionally, they worked together in the field, making positive visual identifications of all birds heard and learning flight and other behaviors.
2. While learning songs and calls, observers practiced distance estimation in the field. Observers worked together, taking turns pointing out objects, with each observer privately recording their estimate of the distance; the distance was then measured. This allows observers to achieve both accurate and precise measurements.

3. Practice bird counts are an essential part of proper training. When any activity is restricted in time, such as a 5-min point count, inexperienced observers become confused, panic, and fail to record reliable data. Having the ability to identify birds (i.e., being a good bird watcher) does not mean a person can conduct an adequate count. Thus, realistic practice sessions are necessary to determine the capability of even good bird-watchers. In the theme study, observers, accompanied by the experienced project leader, conducted "blind" point counts during which each person independently recorded what they saw and heard, as well as an estimate of the distance to the bird. These practice counts were then immediately compared with discrepancies among observers discussed and resolved.

8.2.8 Step 8 – Quality Assurance

The purpose of *quality assurance* (also called quality assurance/quality control, or QA/QC) is to ensure that the execution of the plan is in accordance with the study design. It is important to the successful completion of the study that a formal program of QA/QC is instituted on both the data collection and data processing components (see Levy and Lemeshow 1999).

A principal method of assuring quality control is through resampling a subset of each data set. For example, a different team of observers can resample vegetation plots. Unfortunately, in field biology there are often no absolutely correct answers, especially when visual estimates are involved (e.g., visual estimations of canopy closure). As quantified by Block et al. (1987), sampling errors and observer bias can be minimized through by using relatively rigorous and repeatable measurement techniques whenever possible. For example, measuring dbh with a tape or using a clinometer to estimate tree height rather than through visual estimation. Despite these difficulties, quantifying the degree of interobserver error can only enhance a research project.

In addition to the use of well-trained observers, several other steps can be taken to assure data quality:

1. When conducting data recording in pairs, observers should repeat values back to one another.
2. Another observer at the close of each recording session should proof all data forms. Although this proofing cannot determine if most data are "correct," this procedure can identify obvious recording errors (e.g., 60-m tall oak trees do not exist), eliminate illegible entries (e.g., is that a "0" or a "6"?), and reduce blank entries. Because some of these errors must be corrected by recall of an observer, this proofing must take place as soon after recording as possible.
3. Place each field technician in charge of some aspect of the study. This increases the sense of responsibility and allows the project manager to identify relatively weak or strong personnel (i.e., mistakes are not "anonymous").

8.2.8.1 Theme Example

The quality of data was enhanced and controlled throughout the study duration by incorporating activities such as:

- *Banding*. Repeating of band numbers and colors, sex, age, and all other measurements between observer and recorder. This serves as a field check of data recording, and keeps field technicians alert.
- *Band resightings*. Color bands are difficult to see in the field, and many colors are difficult to differentiate because of bird movements and shadows (e.g., dark blue vs. purple bands). Observers can practice by placing bands on twigs at various distances and heights and under different lighting conditions.
- *Bird counts*. Regular testing of species identification by sight and sound, distance estimation, and numbers counted. Technicians do not always improve their identification abilities as the study proceeds. Numerous factors, such as fatigue, forgetfulness, deteriorating attitudes, etc., can jeopardize data quality.
- *Foraging behavior*. Regular testing of assignment of behaviors to categories. As described above for bird counts, many factors can negatively impact data quality. Additionally, technicians need to communicate continually any changes in interpretation of behaviors, additions and deletions to the variable list, and the like. This becomes especially important in studies using distinct field teams operating in different geographic locations.
- Data proofing: as described above, it is important that one proofs data after every sampling session.

8.2.9 Step 9 – Data Collection

There should be a constant feedback between data collection and QA/QC. Probably one of the principal weaknesses of most studies is a failure to apply QA/QC on a continuing basis. The prevalent but seldom acknowledged problem of *observer drift* can affect all studies regardless of the precautions taken during observer selection and training. The QA/QC should be ongoing; including analysis of sample sizes (are too many or too few data being collected?). Electronic data loggers can be useful in many applications. People can collect data in the field and the data loggers can be programmed to accept only valid codes. Data can then be directly downloaded into a computerized database for proofing and storage. The database then can be queried and analyses made in the statistical program of choice. Previously, data loggers often were beyond the financial reach of many projects. This is no longer true as prices have dropped precipitously and technology advanced considerably.

A weakness of most studies is failure to enter, proof, and analyze data on a continuing basis. Most researchers will quickly note that they have scarcely enough time to collect the data let alone enter and analyze even a portion of it. This is, of course, a fallacious response because a properly designed study would allocate adequate resources to ensure that oversampling or undersampling is not occurring, and that high-quality data are being gathered. Data analysis in multiyear studies should not be postponed until several years' data are obtained. Allocating resources to such analyses will often necessitate a reduction in the scope of the study. Yet, we argue that it is best to ensure that high-quality data be collected for priority objectives; secondary objectives should be abandoned. This statement relates back to our discussion of the choice of questions (Sect. 8.2.1) and the design of the research (Sect. 8.2.3).

8.2.9.1 Applications: Managing Study Plots and Data

When initiated, most studies have a relatively short (a few months to a few years) time frame. As such, maintaining the sites usually only requires occasional replacing of flagging or other markers. It is difficult to anticipate, however, potential future uses of data and the location from where they were gathered (i.e., they might become monitoring studies). In southeastern Arizona, for example, few anticipated the dramatic spreading in geographic area by the exotic lovegrasses (*Eragrostis* spp.) when they were introduced as cattle forage. However, several studies on the impacts of grazing on wildlife included sampling of grass cover and species composition, thereby establishing baseline information on lovegrasses "by accident." By permanently marking such plots, researchers can now return, relocate the plots, and continue the originally unplanned monitoring of these species. Unexpected fires, chemical spills, urbanization, and any other planned or unplanned impact will undoubtedly impact all lands areas sometime in the future. Although it is unlikely that an adequate study design will be available serendipitously, it is likely that some useful comparisons can be made using "old" data (e.g., to determine sample size for a planned experiment). All that is required to ensure that all studies can be used in the future is thoughtful marking and referencing of study plots. All management agencies, universities, and private foundations should establish a central record-keeping protocol. Unfortunately, even dedicated research areas often fail to do so; no agency or university we are aware of has established such a protocol for *all* research efforts.

It is difficult to imagine the amount of data that have been collected over time. Only a tiny fraction of these data resides in any permanent databases. As such, except for the distillation presented in research papers, these data are essentially lost to scientists and managers of the future. Here again, it is indeed rare that any organization requires that data collected by their scientists be deposited and maintained in any centralized database. In fact, few maintain any type of catalog that at least references the location of the data, contents of the data records, and other pertinent information.

Perhaps one of the biggest scientific achievements of our time would be the centralization, or at least central cataloging, of data previously collected (that which is not lost). An achievement unlikely to occur. Each research management organization can, however, initiate a systematic plan for managing both study areas and data in the future.

8.2.9.2 Theme Example

Our example here is brief because the process initiated on the study was primarily continuation, on a regular basis, of the routine established under Step 8, QA/QC (Sect. 8.2.8). Regular testing of observer data gathering was conducted on a bimonthly basis. Additionally, all study points and plots were located both on topographic maps of the study area, and by the use of Global Positioning System (GPS) technology. These data were included in the final report of the study, thus ensuring that the sponsoring agencies had a formal record of each sampling location that could be cross-referenced with the original field data.

8.2.10 Step 10 – Analyses

Testing of hypotheses and evaluating conceptual models are, of course, central features of most wildlife studies (see Sect. 1.3). However, several steps should be taken before formally testing hypotheses or evaluating models. These include:

1. *Calculating descriptive statistics*. As a biologist, the researcher should be familiar with not only the mean and variance, but also the form of the distribution and the range of values the data take on. The message: do not rush into throwing your data into some statistical "black box" without first understanding the nature of the data set. Of interest should be the distribution (e.g., normal, Poisson, bimodal) and identification of outliers.
2. *Sample size analyses*. Hopefully, sample size analyses will have been an ongoing process in the study. If not, then this is the time to determine if you did, indeed, gather adequate samples. Many researchers have found, when attempting to apply multivariate statistics, that they can only use several of the 10, 20, or even 100 variables they collected because of limited sample sizes.
3. *Testing assumptions*. Most researchers understand the need to examine their data for adherence to test assumptions associated with specific statistical algorithms, such as equality of variances between groups, and normality of data for each variable for a t test for example. In biological analyses, these assumptions are seldom met, thus rendering the formal use of standard parametric tests inappropriate. To counter violation of assumptions, a host of data transformation techniques is available (e.g., log, square root). However, we offer two primary cautions. First, remember that we are dealing with biological phenomena that may not correspond to a normal statistical distribution; in fact, they usually do not. Thus, it is usually far more biologically relevant to find a statistical technique that fits the data, rather than trying to force your biological data to fit a statistical distribution.

For example, nonlinear regression techniques are available, thus there is little biological justification for trying to linearize a biological distribution so you can apply the more widely understood and simple linear regression procedures. Second, if transformations are applied, they must be successful in forcing the data into a form that meets assumptions of the test. Most often, researchers simply state that the data were transformed, but no mention is made of the resulting distributions (i.e., were the data normalized or linearized?). Nonparametric and nonlinear methods are often an appropriate alternative to forcing data (especially when sample sizes are small) to meet parametric assumptions.

Once a specific α-value is set and the analysis performed, the P-value associated with the test is either "significantly different from α" or "not significantly different from α." Technically, a P-value cannot be "highly significantly different" or "not significant but tended to show a difference." Many research articles provide specific α-values for testing a hypothesis, yet discuss test results as if α were a floating value. Insightful papers by Cherry (1998) and Johnson (1999) address the issue of null hypothesis significance testing, and question the historical concentration on P-values (see Sects. 1.4.1 and 2.6.1). Although we agree with their primary theses, our point here is that, once a specific analysis has been designed and implemented, changing your evaluation criteria is not generally acceptable. The salient point is that the P-value generated from the formal test must be compared directly with the α-level set during study development; the null hypothesis is either rejected or not rejected. Borderline cases can certainly be discussed (e.g., a $P = 0.071$ with an α of 5%). But, because there should have been a very good reason for setting the α-level in the first place, the fact remains that, in this example, the hypothesis was not rejected. You have no basis for developing management recommendations as if it had been rejected. This is why development of a rigorous design, which includes clear development of the expected magnitude of biological effect (the effect size), followed by thorough monitoring of sampling methods and sample sizes must be accomplished throughout the study. Statistical techniques such as power analysis and model testing (where appropriate) help ensure that rigorous results were indeed obtained. Many articles and books have been written on hypothesis testing, so we will not repeat those details here (Winer et al. 1991; Hilborn and Mangel 1997).

8.2.10.1 Theme Example

Field sampling indicated that removing cowbird adults from the vicinity of nesting hosts, in combination with removal of cowbird nestlings and addling cowbird eggs caused a 54% increase in the number of young fledged (2.0 ± 0.8 young/nest on treatment plots vs. 1.3 ± 1.1 young/nest on nontreated plots; note that these data are preliminary and should not be cited in the context of cowbird–host ecology and management). According to our a priori calculations (see Sect. 8.2.6), we could not detect an effect size of <1 with 80% power and $P = 0.1$ given our sample sizes. However, these results for treatment effects did indicate that our findings were not significantly different ($P = 0.184$) and had an associated power of about 61%. Thus, the results we obtained were what we should have expected given our initial planning based on power analysis.

8.2.11 Step 11 – Interpretation

Conclusions must be drawn with reference to the population of inference only. A properly designed study will clearly identify the population of inference; few studies provide this statement. Rather, most studies are localized geographically, but seek to extend their results to a much wider geographic area in hopes of extending the utility of their results. Statements such as "if our results are applicable to a wider area..." are inappropriate. Without testing such an assumption, the researcher is helping to mislead the resource manager who needs to use this type of data. If the resource need identified before the study covers a "wider" geographic area, then the study should have been designed to cover such an area.

Estimates of population characteristics. Biologists rely on statistics to the point of often abandoning biological common sense when reporting and evaluating research results (see Sect. 1.5.3). We have become fixated on the mean, often ignoring the distribution of the data about that value. However, the mean value often does not occur, or at least seldom occurs, in nature. Populations are composed of individual animals that use various resource axes; they use one or more points along the available resource axis or continuum; an individual does not use the mean per se. For example, a bird foraging for insects on a tree trunk may move slowly up a tree, starting at 2 m in height and ending at 10 m in height. Simplistically this places the bird at a mean height of 6 m, a height at which it may have spent only 10–12% of its foraging times. Likewise, a bimodal distribution will have a mean value, yet from a biological standpoint it is the bimodality that should be of interest. Most of the parametric statistical analyses we use to help examine our data rely on comparisons of mean values, and consider the distribution about the mean only with respect to how they impact assumptions of the test.

Thus, when interpreting research results, biologists should examine both the distribution and the mean (or median or modal) values. Graphical representations of data, even if never published (because of journal space limitations), will be most instructive in data interpretation (e.g., the bimodality mentioned above).

Comparisons of estimated population characteristics with a published constant. A standard feature of most research reports is comparison of results with published data of a similar nature. This is, of course, appropriate because it places your results in context of published standards, and helps build a picture of how generalizable the situation may be. Such comparisons, however, should begin with a thorough analysis of the original papers under review. To be frank, just because a paper is published in a peer-reviewed outlet does not imply it is without weakness or even error. Thus, rather than simply wondering why "Smith and Jones" failed to reject the hypothesis that you have now rejected, you should examine Smith and Jones to see if (1) sample sizes were analyzed, (2) power analysis was conducted, (3) α was adhered to, and (4) their interpretation of results is consistent with accepted statistical procedures (e.g., assumptions tested?).

Comparisons of complementary groups with respect to the level of an estimated characteristic. "The proportion of adult males 3–4 years of age who had a parasite load was greater among residents of urban areas (47%) than among residents of non-urban areas (22%)." For this type of statement to be justified by data, the two estimates should meet minimum specifications of reliability (e.g., each estimate should have a CV of less than 25%), and they should differ from each other statistically.

8.2.12 Step 12 – Reporting

Publishing results in peer-reviewed outlets, or at least in outlets that are readily available in university libraries (e.g., certain government publications), is an essential part of the research process (see Sect. 1.3.1). Regardless of how well you think your study was designed, it must pass the test of independent peer review. The graduate thesis or dissertation, a final report to an agency, in-house agency papers, and the like, are not publications per se. These outlets are termed *gray literature* because they usually are not readily available, and typically do not usually receive independent peer review. Both the editor and the referees must be completely independent from your project for your work to be considered truly peer reviewed. We hasten to add that publication in a peer-reviewed outlet does not confer rigor to your study, nor should the papers published in such outlets be accepted without question. Rather, the reader can at least be assured that some degree of independent review was involved in the publication process.

Further, there are several tiers of journals available for publication. Although the division is not distinct, journals can be categorized roughly as follows:

- *First tier*. Publishes only the highest quality science; work that advances our understanding of the natural world through rigorously conducted studies. *Science, Nature*.
- *Second tier*. Similar to the first tier except not as broad in scope, especially work from ecology that has broad application to animals and/or geographic regions. *Ecology, American Naturalist*.
- *Third tier*. Journals that emphasize natural history study and application of research results, especially as related to management of wildlife populations: application usually to broad geographic areas. *Journal of Wildlife Management, Ecological Applications, Conservation Biology*, the "-ology" journals (e.g., *Condor, Auk, Journal of Mammalogy, Journal of Herpetology*).
- *Fourth tier*. As for third tier except of more regional application; work not necessarily of lower quality than second or third tier. Many regional journals, *Southwestern Naturalist, Western North American Naturalist, American Midland Naturalist*.

- *Fifth tier*. Similar to fourth tier except accepts papers of very local interest, including distribution lists and anecdotal notes. *Transactions of the Western Section of the Wildlife Society, Bulletin of the Texas Ornithology Society, Texas Journal of Science* (or similar).

We again hasten to add that there are numerous seminal papers even in fifth tier journals. Virtually any paper might help expand our knowledge of the natural world. Within reason, we should strive to get the results of all research efforts published. Some journals emphasize publication of articles that have broad spatial applicability or address fundamental ecological principles. In contrast, other journals publish articles that include those of relatively local (regional) and species-specific interest. Contrary to what some people – including journal editors – apparently think, the spatial or fundamental applicability of an article primarily concerns the service that a journal is providing to its readers (or members of the supporting society) rather than the quality of the article per se. Thus, publishing in a "lower tier" journal does not mean your work is not of as high a quality as that published in a high tier outlet.

The publication process is frustrating. Manuscripts are often rejected because you failed to explain adequately your procedures, expanded application of your results far beyond an appropriate level, provided unnecessary detail, used inappropriate statistics, or wrote in a confusing manner. In many cases, the comments from the editor and referees will appear easy to address, and you will be confused and frustrated over the rejection of your manuscript. All journals have budgets, and must limit the number of printed pages. The editor prioritizes your manuscript relative to the other submissions and considers how much effort must be expended in handling revisions of your manuscript. Your manuscript might be accepted by a lower tier journal, even though it may have received even more critical reviews than your original submission to a higher tier outlet. This often occurs because the editor has decided that your paper ranks high relative to other submissions, because the editor or an associate editor has the time to handle several revisions of your manuscript, and/or because the journal budget is adequate to allow publication.

The response you receive from an editor is usually one of the following:

1. Accept as is. Extraordinarily rare, but does occur.
2. Tentatively accepted with minor revision. This occurs more often, but is still relatively rare. The "tentative" is added because you still have a few details to attend to (probably of minor clarification and editorial in nature).
3. Potentially acceptable but decision will be based on revision. This is usually the way a nonrejected (see below) manuscript is handled. If the editor considers your manuscript of potential interest, but cannot make a decision because of the amount of revision necessary, he will probably send your revision back to the referees for further comment. The editor expects you to try to meet the criticisms raised by the referees. It is in your best interest to try to accommodate the referees, be objective, and try to see their point. A rational, detailed cover letter should explain to the editor how you addressed referees' concerns and suggestions, and why you did not follow certain specific suggestions. This is especially important if your revision is being sent back to the same referees!

4. Rejected but would consider a resubmission. In this case, the editor feels that your manuscript has potential, but the revisions necessary are too extensive to warrant consideration of a revision. The revisions needed usually involve different analyses and substantial shortening of the manuscript. In essence, your revision would be so extensive that you will be creating a much different manuscript. The revision will be treated as a new submission and will go through a new review process (probably using different referees).
5. Rejected and suggests submission to a regional journal. The editor is telling you that you aimed too high and should try a lower tier outlet. Use the reviews in a positive manner and follow the editor's suggestion. Your work likely does not have broad appeal but will interest a more local audience.
6. Returned without review. The editors of some journals might return your manuscript without submitting it to review. He or she has made the decision that your manuscript is inappropriate for this journal. Returning without review happens much more frequently in European ecology journals and in the first and second tier North American journals. Although it is understandable that an editor wants to keep the workload on the referees to a minimum, we think it is usually best to allow the referees to make an independent judgment of all submissions; otherwise, the editor functions as a one-person peer review.

Although the reviews you receive are independent from your study, they are not necessarily unbiased. We all have biases based on previous experience (see Sect. 1.2.4). Further, at times, personal resentment might sneak into the review (although we think most people clearly separate personal from professional views). Nevertheless, it is not uncommon for your initial submission to be rejected. It is not unusual for the rejection rate on the 1–3 tier journals to exceed 60%. Although disappointing, you should always take the reviews as constructive criticism, make a revision, and submit to another, possibly lower-tier, journal. Your "duty" as a scientist is to see that your results are published; you have no direct control over the review process or the outlet your work achieves.

8.3 Trouble Shooting: Making the Best of a Bad Situation

Once you have completed your field or lab study there is nothing you can change about your study design per se. That is, if you sampled birds in three treated and three untreated, 1 ha plots; you cannot suddenly have data on birds from 500 m long transects or from larger plots. But, let us say you have determined (perhaps through a hypercritical peer review) that you had insufficient samples to determine treatment effects; you needed more than three treated plots. Is your study ruined? If you are a student, what about completing your thesis? In the sections that follow, we discuss some general approaches to getting something meaningful out of the data you collected even when your design was inappropriate, or when a catastrophe struck your study.

8.3.1 Post-study Adjustments of Design and Data Handling

As noted in Sect. 8.3, you cannot change your study design after the study is completed. You can, however, legitimately change the way you group and analyze your data, which in essence changes your design. For example, say you have collected data in a series of eight plots with the intent of determining the impacts of thinning hardwoods on bird abundance. Because there were no pretreatment data, you were forced to use plots that had been thinned over a period of 4 years; plot size varied with size of the hardwood stands. These treatment plots were distributed over a minor elevation gradient (say, 200 m), and were located on slopes of varying slope and aspect. At once you will recognize several key problems with this design with regard to treatment effects: the treatments are confounded by differences in age, elevation, slope, and aspect. Perhaps analysis of covariance can assist with some of these confounding issues, but problems with sample size and plot size remain. An alternative to trying to move forward with a study of treatment effects would be to turn the study around and look at the response of birds to variations in a gradient of hardwood density and not the treatments per se. Yes, problems with confounding variables remain, but you will be held to a different standard during the peer review process by using a gradient approach rather than an experimental "treatment effects" approach. You will still be able to talk about how birds respond to hardwood density, and make a few statements about how "my results might be applicable to the design of hardwood treatments." Thus, you have a posteriori changed your method of data processing and analysis (from a two-group treatment vs. no treatment to a gradient approach).

Most field biologists recognize that the season of study and the age and sex of the study animals can have a profound influence on study results. As such, data are often recorded in a manner to separate effects of season (time), age, and sex from the desired response variable (e.g., foraging rate). However, dividing your sample into many categories effectively lowers the sample size. That is, if a priori power analysis indicates that you need a sample of 35 individuals (we discuss the issue of independence below) to meet project goals, this usually means a sample of 35 *each* of, say, adult males, subadult males, adult females, and subadult females. Thus, your sample has just been increased to about 140 and probably 140 per season. A standard rule of thumb is to always collect data in reasonable categories because you can always go back and lump, but often you cannot go back and split (i.e., if you did not record age you cannot divide your data into age categories). Thus, while lumping data certainly lessens your ability to tease out biologically meaningful relationships, it does remain an option when sample sizes are too small in your desired categories. You do run the risk of obscuring, say, age specific activities: e.g., the adults show a positive reaction to something you are measuring and the subadults a negative reaction, which expresses itself in your data as "no reaction." But, here we are talking about ways to make the best of a bad situation; we are not able to put insight into data that were collected in a manner that could obscure meaningful relationships.

Another rather common way to change a design is to add or remove nesting or blocking within an experimental framework (see Chap. 3 for discussion of experimental designs). We have often encountered situations where a referee has suggested that data could be more effectively analyzed by blocking across some environmental feature; for example, analyzing data on rodent abundance by elevational blocks or blocks based on vegetation type. Although seen more rarely than blocking, placing data into a nested organization might also enhance an analysis; for example, analyzing young within a burrow that are nested within a colony that is nested within a region. A posteriori *adding* blocking or nesting means that you managed to have an adequate sample size for such procedures. In most cases, you probably will lack sufficient data for blocking or nesting; instead, you move or *remove* blocks or nesting criteria. A related example would be the removal of paired plots into a nonpaired analysis (see Chap. 3).

The issue of independence of data (samples) is one of the central foundations of study design and statistics. Unfortunately, for many applications there are no absolute criteria upon which "independence" can be based. Repeatedly drawing blood samples from the same individual and calling the samples "independent" is likely a clear cut case of nonindependence (or pseudoreplication; see Sect. 2.3).

8.3.2 Adjusting Initial Goals to Actual Outcome

Related to and an integral part of our discussion and examples in Sect. 8.3.1 is the altering of your study goals in light of design inadequacies (or insufficient or inappropriate samples). Situations do arise that are outside of your reasonable control, such as natural or human-caused catastrophes. Fires and floods could virtually obliterate a study area. Perhaps serendipity will have left you with an ideal "treatment effects" study (i.e., pre- and postfire). More likely, you will be left with ashes. Given that a graduate student does not want to stay on for another 2–3 years to complete a different study, about all one might be left with in such circumstances is a brief "before the fire happened" look at the ecology of the animals that were under study. Alternatively, the study can be altered to select different study sites, abandon the initial study goals, and expand to look at the basic ecology of the target animal(s) over a wider geographic area.

As described above (Sect. 8.2.12), there are numerous opportunities to locate a journal that will welcome your manuscript. It might be that your study has been compromised in some way that prevents you from generating the type of ecological or management conclusions that had been intended when you began the work. Thus, you will likely save yourself, as well as referees and a journal editor, a lot of time by matching your manuscript with an appropriate journal (as described under Sect. 8.2.12 regarding tiers of journals). Again, your duty as a scientist is to get your work published. We are certainly not going to argue that a paper in *Science* or *Nature* would get you more accolades than a paper in the *Southwestern Naturalist*. Nevertheless, the fact remains that most of our work will be species and/or site specific. We recommend that you seek the advice from well-published individuals in selecting a journal for submission of your work.

Another option to consider when things have not gone as you anticipated with your study is conversion of your focus to that of a pilot study. If you are unable to draw reasonable conclusions based on your data, then a focus on hypothesis generating rather than hypothesis testing can be a reasonable approach. For example, say that your intended study goal was to make recommendations for management of a rare salamander based on marking and tracking individuals. But, because of various difficulties in first locating and then tracking the animals (e.g., marks could not be read), you failed to gather an adequate sample of individuals to address your initial goal when time and money expired. A reasonable approach, then, would be to focus your study as more methodological and report on solutions to these difficulties. Presentation of your ecological (e.g., habitat use) data would be appropriate, but only as preliminary findings; management recommendations would not likely be appropriate. While serving as co-Editors of *The Journal of Wildlife Management*, Block and Morrison frequently recommended that manuscripts be sent to a regional natural history journal because meaningful management recommendations could not be developed from the study results.

Related to focusing on a pilot study is focusing your work as a *case study*. There are situations in which your biological study was, perhaps, too localized or too brief in duration to warrant a full research article. For example, all of us have been contracted to conduct rather short-term (i.e., one season) assessments of the distribution of endangered species on a wildlife management area, military base, or a site proposed for development. Additionally, any study that results in a small data set collected over a short timeframe might be appropriate for addressing as a case study. While these are valuable data for the issue at hand, they usually have little interest to the general scientific readership of a journal. However, focusing on the *issue* underlying the reason for the study rather than the data collected is a viable way to pursue publication. For example, data collected on the distribution of endangered species on a military base that is slated for closure and potential economic development could serve as the basis for an article on the role of military bases in conservation; the story is the vehicle for carrying the data. Likewise, data collected on a few water catchments could be used to review and discuss the issue of adding water to the environment.

8.3.3 Leadership

The fundamental resource necessary for success in any scientific study is leadership. Regardless of the rigor of the design and the qualifications of your assistants, you must be able to train, encourage, critique – and accept criticism and suggestions, and overall guide your study throughout its duration. Leadership skills are required to develop and guide successful research teams. We have all read job advertisements in which a requirement reads something like “a proven ability to work well with others...”. Employers are looking for people they can work with.

Yet, seldom has a study continued to completion in which no interpersonal problems have arisen. Thus, the skills necessary to properly select, train, and then guide a research team are an essential component of a successful study.

We cannot detail the steps needed to become a leader or how to successfully develop and manage research teams in this book. There are many resources, including books and workshops, which seek to develop leadership skills in project leaders. Probably of more importance, fundamentally, are workshops and other programs that seek to help you understand what drives you as an individual and what causes you to behave and react in the manner that you do under stressful situations.

8.4 Summary

In this chapter, we provide a step-by-step guide to conceptualizing, designing, implementing, analyzing, and publishing a wildlife study. Thoroughly developing the sequence of concept to publication means not only that you have a well thought out plan, but also provides a process to foresee potential problems that may strike a project. The essential steps to a research study are outlined in Fig. 8.1. These steps are the single-study application of the steps typically used in successful natural science research programs (Table 1.2).

We also discuss some general approaches to getting something meaningful out of the data you collected even when your design was inappropriate, or when a catastrophe struck your study. Although you cannot change your study design after the study is completed, you can legitimately change the way you group and analyze your data, which in essence changes your design. We also describe other changes in how data are handled after collection that might lessen your ability to tease out biologically meaningful relationships (assuming the study had been appropriately designed and nothing went wrong), but retain enough valuable information to warrant analysis and publication. There are journals that will welcome your manuscript even if your study was compromised in some way, preventing you from generating the type of ecological or management conclusions that you intended at the onset. We close the chapter with a brief reminder on the central role that leadership plays in developing and guiding a successful research team.

References

Alldredge, M. W., T. R. Simons, and K. H. Pollock. 2007. Factors affecting aural detections of songbirds. Ecol. Appl. 17: 948–955.

Block, W. M., K. A. With, and M. L. Morrison. 1987. On measuring bird habitat: Influence of observer variability and sample size. Condor 89: 241–251.

Cherry, S. 1998. Statistical tests in publications of The Wildlife Society. Wildl. Soc. Bull. 26: 947–953.

Cochran, W. G. 1983. Planning and Analysis of Observational Studies. Wiley, New York, NY.

Cook, C. W., and J. Stubbendieck. 1986. Range Research: Basic Problems and Techniques. Society for Range Management, Denver, CO.

Ehrlich, P. R., D. S. Dobkin, and D. Wheye. 1988. The Birder's Handbook: A Field Guide to the Natural History of North American Birds. Simon and Schuster, New York, NY.

Finch, D. M. 1983. Brood parasitism of the Abert's towhee: Timing, frequency, and effects. Condor 85: 355–359.

Garton, E. O., J. T. Ratti, and J. H. Giudice. 2005. Research and experimental design, in C. E. Braun, Ed. Techniques for Wildlife Investigations and Management, pp. 43–71, 6th Edition. The Wildlife Society, Bethesda, MD.

Harris, J. H., S. D. Sanders, and M. A. Flett. 1987. Willow flycatcher surveys in the Sierra Nevada. West. Birds 18: 27–36.

Hilborn, R., and M. Mangel. 1997. The Ecological Detective: Confronting Models with Data. Monographs in Population Biology 28. Princeton University Press, Princeton, NJ.

Johnson, D. H. 1999. The insignificance of statistical significance testing. J. Wild. Manage. 63: 763–772.

Kepler, C. B., and J. M. Scott. 1981. Reducing bird count variability by training observers. Stud. Avian Biol. 6: 366–371.

Kuhn, T. S. 1962. The Structure of Scientific Revolutions. University of Chicago Press, Chicago, IL.

Laymon, S. A. 1987. Brown-headed cowbirds in California: Historical perspectives and management opportunities in riparian habitats. West. Birds 18: 63–70.

Lehner, P. N. 1996. Handbook of Ethological Methods, 2nd Edition. Cambridge University Press, Cambridge.

Levy, P. S., and S. Lemeshow. 1999. Sampling of Populations: Methods and Applications, 3rd Edition. Wiley, New York, NY.

Martin, T. E. 1992. Breeding productivity considerations: What are the appropriate habitat features for management? in J. M. Hagan and D. W. Johnston, Eds. Ecology and Conservation of Neotropical Migrant Landbirds, pp. 455–473. Smithsonian Institution Press, Washington, DC.

Martin, P., and P. Bateson. 1993. Measuring Behavior: An Introductory Guide, 2nd Edition. Cambridge University Press, Cambridge.

Morrison, M. L., L. S. Hall, S. K. Robinson, S. I. Rothstein, D. C. Hahn, and J. D. Rich. 1999. Research and management of the brown-headed cowbird in western landscapes. Stud. Avian Biol. 18.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 2006. Wildlife–Habitat Relationships: Concepts and Applications, 3rd Edition. Island Press, Washington, DC.

Ralph, C. J., and J. M. Scott, Eds. 1981. Estimating numbers of terrestrial birds. Stud. Avian Biol. 6.

Ramsey, F. L., and J. M. Scott. 1981. Tests of hearing ability. Stud. Avian Biol. 6: 341–345.

Robinson, S. K., J. A. Gryzbowski, S. I. Rothstein, M. C. Brittingham, L. J. Petit, and F. R. Thompson. 1993. Management implications of cowbird parasitism on neotropical migrant songbirds, in D. M. Finch and P. W. Stangel, Eds. Status and Management of Neotropical Migratory Birds, pp. 93–102. USDA Forest Service Gen. Tech. Rpt. RM-229. Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.

Romesburg, H. C. 1981. Wildlife science: Gaining reliable knowledge. J. Wildl. Manage. 45: 293–313.

Steidl, R. J., J. P. Hayes, and E. Schauber. 1997. Statistical power analysis in wildlife research. J. Wildl. Manage. 61: 270–279.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring Vertebrate Populations. Academic, San Diego, CA.

Welsh Jr., H. H., and A. J. Lind. 1995. Habitat correlates of Del Norte Salamander, *Plethodon elongatus* (Caudata: Plethodontidae), in northwestern California. J. Herpetol. 29: 198–210.

Winer, B. J., D. R. Brown, and K. M. Michels. 1991. Statistical Principles in Experimental Design, 3rd Edition. McGraw-Hill, New York, NY.