

## **Chapter 2**

# **Essential Steps Before Writing a Paper**

Having completed a study and acquired all the data required to present it, you are ready to begin preparing a paper. However, before beginning to write, you have to take several critical preliminary steps: Your notes must be gathered, a suitable place for writing must be found, a selective literature review may be helpful, a target journal must be identified, linguistic limitations must be recognized, the study must be defined and delimited, and the information must be arranged. Failure to take these steps will make writing more difficult and seriously compromise the chances of publication. Therefore, this chapter outlines what needs to be done in each of these steps.

### **2.1 Gather Your Notes**

Writing a paper can be greatly facilitated by keeping good notes during the study phase. A comprehensive laboratory or field notebook (either electronic or handwritten) is a valuable resource when you finally sit down to write your paper. It is easy, during data collection, to assume that you will be able to remember why or how you did things. Unfortunately, when you begin to write up your work, several months or even years later, you are likely to find it difficult to recall details. Keeping thorough notes as you proceed, supported by other sources of information such as sketches or photographs, can alleviate a great deal of stress in the latter stages of your study. Of course, it is not only essential to keep good notes, it is also essential to use them. Thus, before sitting down to write, all the relevant notes must be gathered.

### **2.2 Find Somewhere Quiet**

Having gathered your notes, it is essential to find somewhere peaceful to write, or at least somewhere where disturbances are minimal, because writing well requires intense concentration. Ideally, one should find a quiet room and pin a 'Do not disturb' note on the door while arranging the material and writing. Supervisors can help by refraining from demanding to see how the work is progressing every few

minutes. Supervisors have many ways to make students' and post-doctoral workers' lives miserable if they wish, but constantly disturbing them when they are writing is particularly unhelpful; it reduces the quality of the output, thus impairing the chances of publication. On the other hand, an inexperienced writer will need help. Hence, setting a good balance between providing helpful advice and interfering too much is an important 'soft skill'.

## 2.3 *Selectively* Review the Literature

It is vital to review the relevant literature, to ensure that no important observations that either support your findings or contradict them have been missed. Failure to mention such references will create a poor impression and may seriously impair the chances of publication (especially if you miss publications by one or more of the referees). Thus, it is essential to search all of the relevant databases, such as CAB Abstracts (<http://cababstracts.edina.ac.uk>) and ISI Web of Knowledge<sup>SM</sup> (<http://apps.isiknowledge.com>), using all of the potentially relevant keywords. The websites of these databases themselves, and a number of books and other web sources (e.g., Harvard College Library's site; <http://hcl.harvard.edu/research/guides/citationindex/> accessed September 5, 2010), describe ways to search these databases.

However, you can only tell what fields of literature you need to scan when you know precisely what each section of a planned paper is going to cover, and thus the *kinds* of references you need to cite. Otherwise much time could be wasted reading material that is only tangentially related to your study. For instance, if you tried to read every report that has ever been written on over-fertilization of every type of soil, supporting every crop, and all the associated problems, you would never stop reading.

Furthermore, a detailed review of the literature before drafting a paper can be counter-productive, since nearly all of the papers that *could* be read will have some sections that are relevant to your study, but also many sections that have little relevance. Reading such papers can be seriously distracting because they will present many ideas that you may start to think should perhaps be mentioned, complicating rather than helping attempts to draft a clear, coherent framework. In addition, you are likely to be an expert in the subject (otherwise you would not have been able to plan and execute the work) and you should already have good knowledge of the pertinent literature. Therefore, it is generally better to draft your paper first, and *then* identify the aspects of the literature that you can *selectively* focus on. Hence, we recommend restricting any literature survey, at this stage, to at most a couple of recent reviews to refresh your memory about key aspects of the subject that may need to be covered, jotting down brief details of references that *could* be cited. However, even this is optional until the framework of the paper has been drafted.

Ways of identifying references that *need* to be cited while drafting the framework are described in detail later, but here we will mention that key steps in writing

several sections of a paper (especially the *Introduction* and *Discussion*) are identifying appropriate references and deciding where they should be placed. There are two classes of references: essential and illustrative. Essential references are those that have to be cited because they are critical for justifying your study, those for instance showing that a model you used provides robust predictions for analyses such as yours. Illustrative references are those that have been selected from a large number that could have been cited, showing (for instance) that over-fertilization of soil can cause poor root development. A systematic way in which both classes of references can be identified and allocated suitable places for citation is shown in the descriptions of procedures applied to construct sections of papers describing the case studies. However, it should be noted here that, if possible, for illustrative references it is sensible to choose those published by likely referees.

## 2.4 Identify a Target Journal

It is also essential to identify an appropriate target journal. There are several factors to consider here, including the significance of the study, the subject matter and the impact ratings of candidate journals. Assessing the significance of the study is the most difficult, since it is highly subjective. Clearly, all studies are important to the researchers involved, and they are often surprised when friends, relatives and referees fail to see their importance. However, their significance can be roughly assessed by considering the applicability, novelty and generality of the results.

If the results of a study can be applied in multiple disciplines, or major industrial processes, the interest in them will be very high. Similarly, if they include highly unexpected or novel results that are likely to create a major shift in theoretical understanding, the interest will be very high and wide. In addition, if a study has been very extensive, covering large numbers of factors, there is likely to be much greater interest than if the study has been very restricted. In such cases submission to a very highly ranked general science journal, such as *Nature* or *Science*, can be considered. In other cases submission to a journal covering your field of interest is more likely to be successful. It is easy to list possible options in this respect, and both their impact factors and the specific areas that they tend to focus on.

The next step is to identify the journal, amongst likely candidates, with the highest impact factor that routinely accepts papers of similar significance to your study. It is then *essential* to read the journal's instructions for authors *thoroughly*. Astonishingly large proportions of authors either fail to do this, or read them but fail to follow them. Editors of journals find this extremely irritating, since it means that if they accept papers by these authors, a lot of time will have to be wasted telling the authors to amend their papers in accordance with the guidelines. In practice, of course, the editor may simply decide to reject the papers and publish papers by authors who have followed the guidelines instead. Hence, failure to follow the guidelines can seriously compromise the chances of publication (and at best create unnecessary delays).

The journal's guidelines are usually available on-line and should be consulted prior to writing. Perhaps the most obvious restriction is the number of words – it is remarkable how many papers are rejected or require revision because they are too long. Usually the maximum numbers of words allowed for both the *Abstract* and the paper as a whole will be specified. It is wise to know these limits before beginning to write. Indeed, they provide a useful guide. If the maximum length of a paper's main text is 5,000 words, the journal's editor will probably expect most papers to have ca. 4,000–5,000 words. Thus, if your study can be fully covered in less than 2,000 words, there is a substantial chance that the editor will regard it as too slight, that is, as containing too little information (unless the findings are unusually important or unexpected). In such cases, submission to other journals may be more fruitful – or submission to the same journal as a *Short communication*, if it has a section for such contributions.

When you have completed, or nearly completed, a draft of a paper you may decide that the first identified journal is not the best choice, perhaps because the paper is too short, too slight or even possibly that a higher-impact journal might accept it. In such cases another target journal should be identified, and the paper should be adjusted in accordance with that journal's instructions. This is tedious, but it is far better than either sending it to a journal that is likely to reject it or sending it in an inappropriate format.

In addition, there may be limitations on the number of tables or figures and the nature of figures. For example, some publishers ask authors to cover the cost of reproducing figures in color; such expense can be avoided by ensuring that figures are clear in black and white and perhaps providing a link to a website where more detailed versions can be viewed in full color. Furthermore, journals often require either British or US English spellings and grammar to be used. Thus, it is important to use the appropriate language setting and apply your word processor's spell-checking function before submitting a paper. They may also have certain other linguistic requirements, some of which are discussed at various points in the text later and should be followed (if possible).

## 2.5 Awareness of Linguistic Limitations

Having compiled your data, found a quiet place to work, delimited your study, identified a suitable target journal and thoroughly read its *Instructions for authors*, you are ready to begin writing. However, while switching on your word processor, and throughout the rest of the process, it is important to note that if you are writing in English and you are not a native speaker or highly skilled non-native speaker, your writing style has to be adjusted accordingly. Notably:

- It is essential to write more simply than in your first language.
- Much of the advice in standard textbooks about writing papers in English is not helpful, because it tells you *what* to do, but not *how* to do it, thus it is like a

sculptor saying that to create a model of Napoleon you should form a mental image of him and then mold your material into a likeness of it. This is true, but most of us need a little more guidance.

- It may be impossible for a non-native speaker to cover all the key points *and* be clear *and* concise *and* highlight the importance of the study.

If your language skills are not sufficient to address all of the major points clearly and concisely, it is essential to prioritize *clarity*. Then an English-speaking editor, friend or colleague can understand the paper and if necessary improve the English. To illustrate this point, both authors of this guide can write simple sentences in French, but if we try to write complex sentences they become incomprehensible. Editors and referees have similar problems with many papers written by non-native speakers.

For example, a paper we edited recently was full of sentences like:

Physical obstacles have been created purely historically reasons and therefore they would be disappeared, especially at nuclear loci quite quickly, but environmental obstacles will probably be persistent for much longer times due to ecological reasons.

Many referees or editors of journals may decide that such papers have to be re-written, because they are too difficult to understand, or merely reject them, because they think that making them comprehensible will require too much effort. Therefore, if you cannot write fluently in English (or any other stipulated language), it is essential to write simply, preferably in short sentences that are easy to understand. An experienced editor can then make the phrasing more elegant. For example, another paper we handled recently had many passages such as:

A linear correlation between nickel uptake and nitrate uptake was found; the nickel uptake increased with nitrate uptake. Another correlation was found with phosphate uptake; when nickel uptake decreased phosphate uptake also decreased.

This can be stated much more elegantly, as follows: *Positive linear correlations were found between nickel uptake and both nitrate and phosphate uptake*. However, although the text was too long and repetitive, it was easy to understand and edit because the sentences were simple. Thus, it was preferable in many ways to papers with confusingly complex sentences.

Linguistic limitations are further factors that may be considered when choosing a target journal since there are substantial variations in the linguistic standards of journals, and these do not always correlate strongly with their impact factors (i.e., some highly ranked journals do not require the language to be as polished as some less highly ranked journals). Therefore, if you are not a skilled writer, it may be worth identifying journals that sometimes publish papers with poor linguistic standards, especially if you need to publish a paper quickly (for instance to support a grant application). A native English-speaking friend working in your field may be able to help identify such journals.

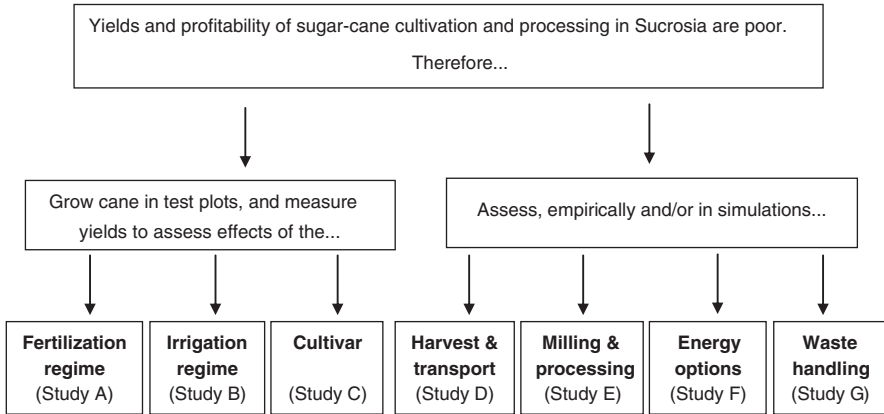
## 2.6 Defining and Delimiting ‘the Study’

The word ‘study’ can be confusing because it has several meanings, *inter alia* an investigation of certain phenomena, and a written report of such an investigation. However, for simplicity, in this guide ‘a study’ always refers to an investigation, and a manuscript describing an investigation is referred to as ‘a paper’. Clearly, before starting to write a paper describing a study, it is essential to decide what the paper is going to cover, that is, the study must be delimited. Sometimes this is easy. For instance, a study could be summarized as follows:

- *Rationale*: It is generally believed that mature bananas are yellow and bent. However, the Learned Society of Unorthodox Thinkers (a fictitious body) has recently postulated that they are in fact red and straight, and if they aren’t they certainly should be.
- *Objectives*: To test the general belief and the Learned Society’s conflicting hypothesis.
- *What was done*: Two thousand mature bananas were acquired and examined.
- *Findings*: All of the mature bananas examined were yellow and bent.
- *Implications*: The results indicate that bananas are generally yellow and bent, supporting the traditional belief (although it is possible that some are red and straight, since the survey was far from comprehensive). Whether they *should* be red and straight requires further investigation.

In this case, delimiting the study is very straightforward. It is also often straightforward in other cases where one or two simple hypotheses are postulated and tested. However, it is not generally quite so easy, because most investigations are much more complex. For example, in PhD projects multiple phenomena are often investigated, which could be reported in (say) three long papers, or larger numbers of short papers. Clearly, in such cases it is essential to decide which parts of the project are going to be covered in a particular paper. Fortunately, researchers usually have intimate knowledge of the scope (and linguistic style) of papers published in journals covering their fields of interest, and this can provide a good guide for deciding how much information should be included in each paper and thus dividing the project into a series of studies.

In addition, the elements of a larger project can usually be grouped into a set of reasonably discrete investigations, which greatly facilitates the delimitation of studies. For example, let us consider the following hypothetical project. The small, fictitious country Sucrosia has a near-ideal location and conditions for producing sugar from sugarcane, hence sugarcane is cultivated in large areas of the country, after which it is milled and the resulting sugar is refined for export. Some of the waste biomass (‘bagasse’) is also used for cogenerating energy. However, the yields and profitability are generally low by international standards. There are grounds for thinking that the poor yields are partly due to over-fertilization. Thus, this possibility clearly needs to be addressed, but many other variables also need to be considered, including the irrigation strategy applied, the cultivars used, harvesting operations



**Fig. 2.1** Flow chart delimiting studies within the hypothetical Sucrosia sugarcane project

and scheduling, field-to-mill transport, cut-to-milling delays, the milling and other processing equipment, the use of steam and power, energy cogeneration options and waste treatment. In such cases, as illustrated in Fig. 2.1, flow charts may be very useful for visualizing the work to be done, delimiting studies within the project (and assigning human and other resources to them), tracking progress and (eventually) writing sections of papers and/or reports.

Sub-dividing a project in this manner can also provide a very convenient means for constructing sub-titles of sections of interim and final reports, by simply slightly re-wording the summarizing terms (shown in bold) for each of the delimited studies (e.g., *Optimization of the fertilization regime*, *Optimization of the irrigation regime* and *Milling and processing strategies*). Of course, each of these studies may be quite complex, so further, more detailed flow charts for each of the studies may be useful, as illustrated for two complex hypothetical case studies in Chapter 4. Such detailed flow charts can provide templates for writing sections of papers (Chapter 4), and can be useful for identifying references that need to be cited in each section, as described later. It should also be noted that there will be substantial overlap between some sections of papers describing these studies, for instance, plots at the same study sites will probably be used for the fertilization, irrigation and cultivar selection studies, so the descriptions of the sites, and the criteria used to select them, will be the same in Studies A, B and C. This is convenient, because these aspects of the studies need to be described in detail only once, and after (say) writing a paper on optimization of the fertilization regime, papers on optimization of the irrigation regime and cultivar selection can refer to information in the first paper.

Having divided a project into discrete studies that can be described in papers of appropriate length, a related problem is deciding where to start from, that is, what aspect of each study to describe first. Some authors recommend starting by describing *what was done*, that is, the *Materials & Methods* section of a scientific paper in traditional format, or the *findings*, that is, the *Results* section (Malmsfors et al. 2004,



Gustavii 2008, Booth 1993). Starting with *what was done* has some merits, since it is the only aspect that is certain (e.g., there may be uncertainty about what to include in an *Introduction* or *Discussion*, but provided good records have been kept, there should be little doubt about *what was done*). Similarly, the researchers will have clear knowledge about their results (although their *implications* may be disputed). Furthermore, having written either of these sections, the resulting text can provide a framework on which to base the rest of the paper, to ensure that all sections are consistent.

However, this raises two problems. First, it is essential to know exactly what to include in these sections, which can only be decided after delimiting the study. Second, it is essential to describe the *Materials & Methods* (and *Results*) in a logical order. Generally, the optimal order is chronological, for reasons discussed later (although other approaches to organizing material *can* be used, see Matthews and Matthews 2008). However, investigators might only remember to analyze certain control samples that should have been analyzed in early stages of an investigation toward its end. In such cases, they would seem foolish if they presented what they did in the *true* chronological order, stating at the end of the *Materials & Methods* section *We then analyzed the controls, which we had previously forgotten to do*. Instead, it would be far better to state at an earlier point that *Both the extracts and controls were analyzed*. Similarly, a substance that has taken months to purify may be dropped, scraped off the floor, re-purified and then analyzed. In such cases we would not recommend stating *The substance was dropped, scraped off the floor, re-purified and then analyzed*. Instead, we would write, simply *The purified substance was then analyzed*. Thus, the *Materials & Methods* section should present *what was done*, or rather what would have been done if everything had been done correctly the first time in an ideal order, which may not coincide completely with the order in which everything mentioned was actually done.

For these reasons, a framework (which should be clear, simple and consistent) is required *before* starting to write this or any other section. Fortunately, such a framework can be constructed, for any study, by briefly describing the *rationale*, *objective(s)*, *what was done*, the *findings* and the *implications* of the study. The way in which such a framework can be used to compose each section of a paper is described in detail in following parts of this guide, but before doing so we should define these terms, recognizing that a scientific investigation is rarely a smooth progression from an initial rationale, through formulation of a set of testable hypotheses, to experiments that have been perfectly designed and executed, yielding perfectly analyzed and interpreted results. Thus, here:

- *Rationale* refers to the context or background of the study, as understood at the time of writing, which may not fully coincide with the initial rationale. For instance, the initial rationale may have been partly based on a misunderstanding of a previous author's work. If so, we would not recommend writing that *certain hypotheses were tested because we misunderstood Smith's conclusions*, but instead adjust the rationale. Note, this is quite different from cases in which a



well-grounded hypothesis formulated from a sound initial rationale was tested and found to be false, for which there is no need to adjust the rationale.

- *Objectives* refers to the specific goals of the study as understood (with hindsight) at the time of writing, which again may not fully coincide with the initial objectives, since the goals may shift during the course of the study; some may be added, some may be dropped and others may change. Thus, for instance, it would usually be pointless for an author to describe the context of a hypothesis that he/she initially planned to test, but did not because there was insufficient time, except perhaps in the conclusion, if possible future analyses are mentioned.
- *What was done* refers to the experiments and analyses that were performed, in the order that they should ideally have been performed.
- *Findings* refers to the results from those experiments and analyses, and the conclusions that can be directly drawn from them.
- *Implications* refers to conclusions that can be indirectly drawn or inferred from the *findings*, for example, whether a tested substance could be viably used in a proposed application, with referenced comparisons to previously published findings.

Initially, statements describing these aspects can be very short. Indeed, writing short statements describing each of the aspects is essential for composing key sections of a paper (especially the *Title*, *Abstract* and *Conclusion*). In addition, a fuller, much more detailed framework can be very useful for checking that all aspects of the paper are consistent, in other words that:

- the *rationale* provides sufficient context to justify everything that was done
- the description of *what was done* details all the materials used, treatments applied and experiments for which results will be mentioned
- results of all experiments mentioned in the *what was done* section are covered
- all of the main *findings* are discussed and
- appropriate references have been added at appropriate places.

How such a framework can be drafted and applied in practice are the main concerns of the rest of this guide.

## References

- Booth V (1993) Communicating in science: writing a scientific paper and speaking at scientific meetings. Cambridge University Press, Cambridge, UK
- Gustavii B (2008) How to write and illustrate a scientific paper. Cambridge University Press, New York, NY
- Malmsfors B, Garnsworthy P, Grossman M (2004) Writing and presenting scientific papers. Nottingham University Press, Nottingham
- Matthews JR, Matthews RW (2008) Successful scientific writing: a step-by-step guide for the biological and medical sciences. Cambridge University Press, New York, NY

A Scientific Approach to Scientific Writing

Blackwell, J.; Martin, J.

2011, X, 114 p. 3 illus., Softcover

ISBN: 978-1-4419-9787-6