1

Preparing to write

> Find a subject you care about and which you in your heart feel others should care about. It is this genuine caring, not your games with language, which will be the most compelling and seductive element in your style.
>
> – *Kurt Vonnegut*

Most of us were drawn to science because, like Vonnegut, we found a subject we feel deeply about, not just because we wanted to write about it. However, all scientists recognize that research must be made known if it is to have lasting value. This is how science moves forward, with the shared word illuminating each step of discovery for the sake of others that follow.

"Scientific writing" can be defined narrowly as the reporting of original research in journals or more broadly to encompass other ways that scientists share research information with one another, such as review articles, posters, and slide-based presentations. (The term "science writing" is often used for writing about science topics for the general public.) Whatever form it takes, successful scientific writing must answer basic questions and address problems raised during the dialogs that identify and define a given subject. It must be clear, concise, and follow established formats. In many ways, its language forms a dialect all its own.

What is the most efficient way to write a paper or presentation that successfully covers all this? This book exists to help you tackle the task, step by step. In this chapter, we suggest that you back up from actual writing, and start where your research does – with a question. Learn the most effective ways of compiling background information. For help defining, organizing, and planning the content, use techniques borrowed from problem-solving strategies. Choose a journal so that you have a goal and format. Finally, take charge of the whole project by using the Process Approach.

SEARCH AND RESEARCH

Any time we reach past our own knowledge and experience to seek out, investigate, and use materials beyond personal resources, research is involved. It may be the study of a subject through firsthand observation and investigation, such as carrying out a laboratory experiment, conducting a survey, or sifting through statistical data. Or it may be the examination of studies that other researchers

have made of a subject, as presented in books, articles, or scientific debates. Most often it is an amalgamation of the two, for literature research and laboratory research form a powerful combination.

The first substantial writing that many beginning scientists produce is either a prospectus or progress report on their thesis, or dissertation research, or a short journal article written jointly with their supervisor or major professor. Increasingly, a detailed prospectus, including a literature review, is being requested before research projects can begin. Likewise, in business and industry, a well-written proposal often must precede approval for research projects, and its worth can influence promotion and pay. In fact, one would be hard pressed to find any scientific profession that would not require checking sources of information about a specific subject, integrating this information with one's own ideas, and presenting thoughts, findings, and conclusions effectively.

Conducting a comprehensive literature review

Conducting a comprehensive literature review is undeniably a big job. Here are a few general points of advice to help you coordinate your work, followed by tips specific to conducting computer-based searches.

Organization is a journey, not a destination

A literature review means you'll soon be handling an avalanche of papers – at the very least, personal notes, photocopies, journal reprints, and printed copies of electronic publications. It is essential to have some system in place to deal with all the information that will be converging upon you.

What system is most effective? There is no one-size-fits-all answer. The popular press is brimming with suggestions, often coupled with explicit or implicit promises of spectacular life results if one can only become properly organized (for examples, see Aslett, 1996; Bolker, 1998). Seek out such materials if you feel you need motivation, inspiration, or novel approaches, but maintain your perspective. The secret to effective and efficient scientific writing isn't simply in getting organized. It is in wanting to get the job done and committing oneself to do it. However, having a system from the beginning and consistently staying up with it can go a long way to keeping that commitment on track.

Mind your Ps and Qs

Whether you photocopy journal articles, request reprints, or print potentially helpful information from the Internet, you will soon amass a great many facts and ideas couched in the words of others. The old advice from typesetting days, "mind your Ps and Qs," is worth remembering in this new context.

First, watch the *P*s – print materials. It will be tempting to use these copies as a substitute for taking notes. However, because of the way that writing and thinking are related to each other, it is actually more effective if you can begin to digest these written materials as you go along. Adopt a good note-taking procedure right from the start. Take many more notes than you think you need and prune them

later. Staple the notes to the print materials so they will remain together through the inevitable subsequent paper-shuffling.

Second, watch the *Q*s – quoted material. To avoid unintentional plagiarism, always write notes in your own words. Indicate their source. If you must quote directly, use extreme care to identify quoted material either with quotation marks or with the letter *Q*.

Use many different search strategies

Trace information in all directions through time and space. Each search strategy has different strengths and weaknesses, and will uncover a somewhat different set of information.

Later in this chapter, we will discuss computerized searches in some detail. However, the idea of networking pre-dates computer-based searches. For example, a time-honored search strategy called the Ancestry Approach starts by acquiring a research report and examining its references to find other relevant references. Through reiteration, researchers work their way back through the literature until either the important concepts disappear or the studies become so old they can be judged obsolete.

A more recent set of searching tools employs the Descendency Approach. Citation indexes identify a publication's offspring – those more recent books and journal articles that reference the earlier work.

Make it easy to relocate relevant material

Write the full journal source on each photocopy or computer printout, if the source is not printed somewhere on the page. For material obtained from online sources, list the author, if available; title, document, file, or website; date of the material; name of the database or other online source; date you accessed the source; and the full electronic address or Uniform Resource Locator (URL).

It is particularly easy to forget how one actually located online material. To minimize this problem, it is a good idea to set up an electronic bookmark that identifies a location you may want to revisit. Over time these bookmarks will accumulate into a customized list that makes it easy to locate and return to particular sites.

Use email as a timesaving resource

Being transmitted in machine-readable form, email text can be printed, revised, and sent back, or even incorporated directly into another computer file without being retyped. These abilities can be used to your advantage in many ways. References, abstracts, and even entire articles located in a particular database can be directed to your personal email address. There you can download them, then print or add them directly to your computerized literature retrieval system.

Conducting computer-based searches

The way in which we obtain information is changing rapidly. A decade or two ago, most literature searching was done manually. Computerized literature

databases were searchable only through a mainframe, searching software was difficult to use, and online searching was expensive and limited in scope. Specially trained librarians did most of the searching, and researchers paid telecommunication charges for reaching the mainframe and were charged for each record received.

Today, in many fields, a literature search that once took six months to a year can often be done in less than ten minutes, and with far more thorough results. Thousands of specialized databases exist around the world. Database software has become increasingly user-friendly. Research libraries and even moderately sized community libraries buy site licenses to various indexes, and offer their clients free searching of CD-ROMs and mainframe-mounted indexes. The Internet offers direct access to both new and old sources of information.

The upshot of this revolution is that you need to know how to conduct a literature search yourself. Whether you consider this a blessing or a curse depends on your approach to the task and your knowledge of available resources.

Compilations are there to help – use them!

Research bibliographies, research registers, reference databases, and citation indexes are compilations constructed for the explicit purpose of providing relatively comprehensive lists of published information related to a topic. They can be some of your most valuable literature searching sources.

Each of these databases has limitations, however. Some contain only published research; others, only unpublished research. As with searching the Internet, one searches the database by specifying keywords; any mismatch between the seeker and the indexer is likely to result in missed articles. There can be a long time lag before references appear in an electronic index, because after publication the work must be identified and catalogued into the reference database. Thus, physically browsing for newly appearing information is still advisable. Furthermore, despite their claims, none of the online databases access all relevant journals on a topic. Use multiple sources.

Consult research bibliographies and research registers

Research bibliographies can be a great help and time-saver. They generally take the form of nonevaluative listings of books and articles relevant to a particular topic area, but it is even possible to find bibliographies of bibliographies. Research bibliographies are often maintained by single scientists or groups of individuals, rather than by a formal organization.

Prevalent in the medical sciences, research registers are databases of studies focusing on a common feature, such as subject matter, funding source, or design. Prospective research registers are unique in attempting to include not only completed research, but also research that is in the planning stage or is still under way. Some research registers are more comprehensive than others; whenever possible, determine how long a register has been in existence and how the research included in the register got to be there.

Locate and use reference databases and abstracting services

Reference databases (Table 1.1) are particularly fruitful sources of information. Maintained by both private and public organizations, these services focus on a specific kind of document (such as theses and dissertations) or field (such as agriculture or medicine). At present, most include only titles and abstracts, but full-text databases are becoming more prevalent and probably will be the norm in the future.

Table 1.1. *Examples of helpful literature abstracting and indexing databases available to biological and medical researchers. All are available in both traditional formats and online from various vendors*

Database	Description
Agricola	Covers all major areas of agricultural sciences.
Agricultural and Environmental Biotechnology Abstracts	Especially useful for genetic engineering and its agricultural implications.
Bioengineering Abstracts	Covers biomedical and genetic engineering and related fields.
BIOSIS (Biological Abstracts)	Widely used for literature in biology, agriculture, and biomedicine. Includes five different indexes – author, genus, biosystematic grouping from phylum through family, concept, and subject. Records prior to 1993 are formatted and indexed different from records since that time.
Biological and Agricultural Index	Particularly useful for environmental and conservation sciences, agriculture, veterinary medicine, and related areas of applied biology.
Books in Print	Covers in-print, out-of-print, and forthcoming books from North American publishers.
CAB Abstracts	Excellent coverage for agriculture, veterinary medicine, and biology.
Cambridge Scientific Abstracts	Source of several particularly useful databases.
CINAHL (Cumulative Index to Nursing and Allied Health)	Particularly strong coverage of the nursing and allied health professions literature.
Current Contents	Indexes recent articles in a variety of life sciences by reproducing the tables of contents of numerous journals. Authors' addresses enable contact to request a copy of the paper if the journal is unavailable. Includes abstracts.
Dissertation Abstracts	Provides complete abstracts of dissertations from U.S., Canadian, British, and other countries, plus select coverage of masters theses.

(*cont.*)

Table 1.1. (*cont.*)

Database	Description
General Science Index	Helpful place to start when working with a broad topic. Includes both papers in selected technical journals and nontechnical overviews, many of which are written by scientists who have also published technical papers on the same topic. (Locate the latter by searching by author names in more specialized databases.)
Journal Citation Reports (an annual volume of *Science Citation Index*)	Lists indexed journals grouped by subject field. Ranks journals by their relative "impact factors," including number of citations of a journal's papers in other publications during a given calendar year and other statistics.
Medical and Pharmaceutical Biotechnology Abstracts	Covers human health, molecular biology, and biotechnology.
PubMed (*MEDLINE*)	The online counterpart to *Index Medicus*, and one of a group of databases (MEDLARS = Medical Literature Analysis and Retrieval System) produced by the U.S. National Library of Medicine. Includes all the medical and health sciences; unsurpassed for preclinical and clinical medicine.
Science Citation Index	Widely used to locate other authors who have mentioned a paper relevant to one's topic.
Web of Knowledge	Incorporates various searchable databases (including *Science Citation Index*) from Thomson Scientific (formerly Institute of Scientific Information [ISI]).
Zoological Record	The most comprehensive index to zoological literature.

All major research libraries subscribe to numerous reference databases and have reference librarians to help first-time users. Many databases are available in more than one medium or format. The older media (print, microfilm, microfiche, and more recently CD-ROMs) require physically visiting the library. Using online reference databases can save considerable time and ensure a high degree of accuracy. Furthermore, online reference databases are sometimes updated more frequently than their CD-ROM or print equivalents. Some databases are accessible only through licensed sites, such as a university library.

Individual vendors and reference database publishers provide detailed and readily available instructions on database searching. Learn the shortcuts that make can make your life easier. For example, database software usually has the capacity to format reference citations in a variety of ways, representative of the formats most commonly found in the scientific literature. Some database software programs also can be integrated with many word processing programs to format references automatically within a document and insert them during typescript preparation. Become familiar with the most widely used formats in

Table 1.2. *Preliminary questions to ask about research design*

The basic question	Examples of ways in which it might be assessed
1. Do I know what I'm doing?	Have I drawn up a plan (a protocol) for what I intend to do? Do the proposed studies cover all the criticisms likely to be made? Are the statistical methods valid?
2. Do my proposed experiments meet accepted ethical standards?	If my experiments involve human beings or animals, do they meet accepted standards? Could my work adversely affect the environment or the place where I am doing field work?
3. What practical and political considerations need to be addressed?	Is publication of my work likely to break any official secrecy regulations? Could publication invalidate a later application for a patent? Are collecting or other permits required?
4. How will I record the work as it proceeds?	How will I record what I read? How will I record what I do? How will I ensure that my records are complete? How will I ensure that I can access the records again when I or others need them?

your discipline and select the most up-to-date and versatile tools available. Take the time to master them.

Consult citation indexes and Dissertation Abstracts

Citation indexes are a unique kind of reference database that identifies and groups together all newly published articles that have referenced (cited) the same earlier publication. Citation indexes limit entries to references in published research, both journals and books, but are quite exhaustive within these categories.

Academia houses a great deal of potentially valuable but largely unpublished material in the form of doctoral dissertations and masters' theses. Although many reference databases contain abstracts of dissertations, *Dissertation Abstracts* focuses exclusively on them. Both the printed and the computerized versions include records dating back to 1861. Increasingly, the full text can be purchased and printed. Alternatively, you may need to use interlibrary loan services to obtain a photocopy from the university at which the dissertation research was conducted.

Learn to use keyword search terms and apply Boolean logic

Most literature retrieval services are really matchmakers (Table 1.2). They have some provision for searching a subject by way of keywords – brief terms chosen (usually by a study's author) to describe the major topics included in the document. To find the document, one must specify the same keyword that the author has chosen (or a part of it; see "wildcard characters" later in this chapter).

Language gets much of its meaning through context, however. As a result, typing in keywords during an Internet search without specifying their context or relationships can lead to strange, frustrating, or humorous results. To improve the outcome, use a special system called Boolean logic to specify the relationships between search terms.

Boolean logic is named for George Boole, a mathematician who lived in the middle 1800s. It really is just a highbrowed way of describing three logical choices:

I want this one AND that one
I want this one OR that one
I want this one but NOT that one

Search tools let you apply Boolean logic in various ways. A common variation allows you to choose from a menu of options that describe the Boolean logic, such as "all of these words," "any of these words," and "must not contain."

Suppose you wish to undertake a comparative study of types of skin cancer. By specifying `carcinoma AND melanoma`, you would retrieve all the hits (entries computer-matched to your search) in which both types of cancers appear in the same document, but none that mention only one. For a comprehensive search on both kinds of skin cancer, you would specify `carcinoma OR melanoma`. Either or both terms would appear in each document that is retrieved. Alternatively, perhaps you want more information on skin cancers, but know that because of its potential deadliness, there will be hundreds of entries on malignant melanoma. To narrow the results, you could specify `carcinoma NOT melanoma`. Any document about skin cancer that mentioned melanoma would be omitted from the list of retrievals.

With another system called Implied Boolean, you use "logical operators" – a plus sign in place of `AND` and a minus sign in place of `NOT`. The signs abut the front of the word, with no space between them. Precede this with other search terms you want to have it coupled with. For example, type `plastic facial +surgery` to get results for facial surgery and plastic surgery but not for the words plastic or facial alone. Use a minus sign in front of a word to ensure that a word does not appear in hits. For example, `poisoning -food` would yield information on poisoning without including entries on food poisoning.

Plan an effective search strategy

For efficient use of time and energy, carefully define the scope of your literature review right at the beginning. How extensive do you want it to be? Do you want to get a broad list that includes records even slightly related to your topic, or just a few most relevant ones? To what extent do you need to rely upon informal channels versus formal ones?

Then, be prepared for a bit of trial-and-error. Identify a limited number of concepts that may be useful to describe the research question at hand, and choose terms and accompanying logic that seem to define them. Precision is imperative. Searching for instances of a broad term like `ecology` would be akin to drinking

from a fire hose, summoning thousands of hits. The list that is returned often will display the total number of items found, but only show them in batches.

Run a computerized search using your initial set of terms, and look over a sample of the records it retrieves. Are they mostly relevant? If not, revise your search. To increase the number of records, expand the lists of terms connected by OR. To retrieve fewer records, narrow the search by adding terms or concepts connected by AND or (very carefully) by NOT logic. Most databases will let you define a time period or subject area for your search; many Internet searches still will not. Another useful capability of some online databases is the option of using an index tree or thesaurus. The vocabulary is arranged hierarchically, allowing the searcher to scroll through the list and select topics to broaden or narrow search parameters as desired.

When you are satisfied with the records obtained from one information channel, but feel you do not have everything that you need or want, begin all over again with another. The results will probably be different.

Handle search results wisely

Exercise care and vigilance when entering reference citations into your personal database. While it can be tempting to add bibliographic references directly into your personal database from the literature cited sections of review articles and other publications, avoid doing so. Never incorporate a reference into your database until you have actually verified its accuracy and appropriateness.

Using the Internet wisely and well

As a successful writer, you will find yourself using the Internet repeatedly. This vast interconnected system of smaller public and private networks lets users communicate around the globe, finding and sharing information, offering commercial services, and opening vast information resources.

Remember two things, however. First, the Internet is an ever-changing entity. Printed material pointing to specific sites is sometimes outdated before it is even published, and finding something useful once doesn't mean you will be able to locate it again. The secret to dealing with this vast, chaotically organized resource and its instability is learning to understand how it works and how to use specialized tools designed to facilitate your scientific writing efforts.

Second, the Internet has no gatekeepers. Material can be, and is, posted by anyone who cares to do so. This form of publishing lets everyone have a voice, and it provides for a wealth of information. However, just because something appears – even on a really fantastically professional looking page – doesn't mean that the information necessarily is credible.

Evaluate Web entries carefully

As Gurak (2000) points out, one or more of the following characteristics indicate a credible site:

Exercise 1.1. Search strategy and Boolean logic

This is an exercise in thinking logically, not in finding answers on the Web. Here are ten publication titles. Use them to answer the questions below.

A. Trap-Nesting Wasps And Bees: Life Histories, Nests, And Associates
B. Behavior Of Three Florida Solitary Wasps
C. Winged Warriors: Insects In The Garden
D. A Cluster Of Bees
E. The Wasps Of The Genus *Pisonopsis* Fox
F. Beeswax, Twine, and Time: The Art of Candlemaking
G. Cowfly Tigers: An Account Of The Bembicine Wasps Of British Guyana
H. Honeybees Attacked At Their Hive Entrance By *Philanthus* Wasps
I. A Life History of Stinging Insects
J. A Comparative Study Of The Nesting Habits Of Solitary Bees And Wasps

Write the number(s) corresponding to the title(s) that would be retrieved for each of the following Boolean statements.

1. Wasps AND Bees
2. Wasps NOT Bees
3. Bees NOT Wasps
4. Wasps OR Bees

5. When using terms in a subject directory, you will usually get only relevant titles. When using terms in a search engine, you should expect a mixture of relevant and irrelevant titles. If you were searching for Wasps OR Bees using a search engine, which of the above titles would probably be retrieved but have little to do with them?

6. When you search one of the better subject directories, you search not only titles but annotations written by a staff person. Which of the above titles would be missed by a search engine using the keywords Wasps OR Bees, but might contain relevant information that would be retrieved by a good subject directory?

- It is an online version of a reputable published source, such as a newspaper, major media source, or an academic or professional journal
- It includes a list of works cited
- It is affiliated with a reputable educational or research institution
- The authors of the site are identified, with information about how to contact them