Preface

Criticism and testing are of the essence of our work. This means that science is a fundamentally social activity, which implies that it depends on good communication. In the practice of science we are aware of this, and that is why it is right for our journals to insist on clarity and intelligibility. . . .

—Hermann Bondi

◇ ◇ ◇

Good scientific writing is not a matter of life and death; it is much more serious than that.

The goal of scientific research is publication. Scientists, starting as graduate students, are measured primarily not by their dexterity in laboratory manipulations, not by their innate knowledge of either broad or narrow scientific subjects, and certainly not by their wit or charm; they are measured, and become known (or remain unknown) by their publications.

A scientific experiment, no matter how spectacular the results, is not completed until the results are published. In fact, the cornerstone of the philosophy of science is based on the fundamental assumption that original research *must* be published; only thus can new scientific knowledge be authenticated and then added to the existing database that we call scientific knowledge.

It is not necessary for the plumber to write about pipes, nor is it necessary for the lawyer to write about cases (except *brief* writing), but the research scientist, perhaps uniquely among the trades and professions, must provide a written document showing what he or she did, why it was done, how it was done, and what was learned from it. The key word is *reproducibility*. That is what makes science and scientific writing unique.

Thus the scientist must not only "do" science but must "write" science. Bad writing can and often does prevent or delay the publication of good science. Unfortunately, the education of scientists is often so overwhelm-

ingly committed to the technical aspects of science that the communication arts are neglected or ignored. In short, many good scientists are poor writers. Certainly, many scientists do not like to write. As Charles Darwin said, "a naturalist's life would be a happy one if he had only to observe and never to write" (quoted by Trelease, 1958).

Most of today's scientists did not have the chance to undertake a formal course in scientific writing. As graduate students, they learned to imitate the style and approach of their professors and previous authors. Some scientists became good writers anyway. Many, however, learned only to imitate the prose and style of the authors before them—with all their attendant defects— thus establishing a system of error in perpetuity.

The purpose of this book is to help scientists and students of the sciences in all disciplines to prepare manuscripts that will have a high probability of being accepted for publication and of being completely understood when they are published. Because the requirements of journals vary widely from discipline to discipline, and even within the same discipline, it is not possible to offer recommendations that are universally acceptable. In this book, I present certain basic principles that are accepted in most disciplines.

Contents

Thus, with appropriate humility, I will try to tell you a few things that may be of use in writing scientific papers.

In the Preface to the First Edition, I stated that I would "view the book as a success if it provides you with the information needed to write effective scientific papers and if it makes me rich and famous." Having since achieved neither fame nor fortune, I nonetheless continue to hope that this book is "a success" for you, the reader.

Finally, I hope that those of you who have used earlier editions of this book will notice improvements in this edition. One thing I'm sure of: I'm not as big a fool as I used to be; I've been on a diet.

Chapter 1
What Is Scientific Writing?

State your facts as simply as possible, even boldly. No one wants flowers of eloquence or literary ornaments in a research article.

—R. B. McKerrow

◇ ◇ ◇

THE NEED FOR CLARITY

The key characteristic of scientific writing is clarity. Successful scientific experimentation is the result of a clear mind attacking a clearly stated problem and producing clearly stated conclusions. Ideally, clarity should be a characteristic of any type of communication; however, when something is being said *for the first time,* clarity is essential. Most scientific papers, those published in our primary research journals, are accepted for publication precisely because they *do* contribute *new* knowledge. Hence, we should demand absolute clarity in scientific writing.

RECEIVING THE SIGNALS

Most people have no doubt heard this question: If a tree falls in the forest and there is no one there to hear it fall, does it make a sound? The correct answer is no. Sound is more than "pressure waves," and indeed there can be no sound without a hearer.

And, similarly, scientific communication is a two-way process. Just as a signal of any kind is useless unless it is perceived, a published scientific paper (signal) is useless unless it is both received *and* understood by its intended audience. Thus, we can restate the axiom of science as being: A scientific experiment is not complete until the results have been published *and understood.* Publication is no more than "pressure waves" unless the published paper is understood. Too many scientific papers fall silently in the woods.

UNDERSTANDING THE SIGNALS

Scientific writing is the transmission of a clear signal to a recipient. The words of the signal should be as clear and simple and well ordered as possible. In scientific writing, there is little need for ornamentation. The flowery literary embellishments—the metaphors, the similes, the idiomatic expressions—are very likely to cause confusion and should seldom be used in writing research papers.

Science is simply too important to be communicated in anything other than words of certain meaning. And that clear, certain meaning should pertain not just to peers of the author, but also to students just embarking on their careers, to scientists reading outside their own narrow discipline, and *especially* to those readers (the majority of readers today) whose native language is other than English.

Many kinds of writing are designed for entertainment. Scientific writing has a different purpose: to communicate new scientific findings. Scientific writing should be as clear and simple as possible.

LANGUAGE OF A SCIENTIFIC PAPER

In addition to organization, the second principal ingredient of a scientific paper should be appropriate language. In this book, I keep emphasizing proper use of English, because most scientists have trouble in this area. All scientists must learn to use the English language with precision. A book (Day, 1992) wholly concerned with English for scientists is now available.

If scientifically determined knowledge is at least as important as any other knowledge, it must be communicated effectively, clearly, in

must therefore be literate. David B. Truman, when he was Dean of Columbia College, said it well: "In the complexities of contemporary existence the specialist who is trained but uneducated, technically skilled but culturally incompetent, is a menace."

Although the ultimate result of scientific research is publication, it has always amazed me that so many scientists neglect the responsibilities involved. A scientist will spend months or years of hard work to secure data, and then unconcernedly let much of their value be lost because of lack of interest in the communication process. The same scientist who will overcome tremendous obstacles to carry out a measurement to the fourth decimal place will be in deep slumber while a secretary is casually changing micrograms per milliliter to milligrams per milliliter and while the compositor slips in an occasional pounds per barrel.

English need not be difficult. In scientific writing, we say: "The best English is that which gives the sense in the fewest short words" (a dictum printed for some years in the Instructions to Authors of the *Journal of Bacteriology*). Literary devices, metaphors and the like, divert attention from the substance to the style. They should be used rarely in scientific writing.

Chapter 2
Origins of Scientific Writing

For what good science tries to eliminate, good art seeks to provoke—mystery, which is lethal to the one, and vital to the other.

—John Fowles

◇ ◇ ◇

THE EARLY HISTORY

Human beings have been able to communicate for thousands of years. Yet scientific communication as we know it today is relatively new. The first journals were published only 300 years ago, and the IMRAD (Introduction, Methods, Results, and Discussion) organization of scientific papers has developed within the past 100 years.

Knowledge, scientific or otherwise, could not be effectively communicated until appropriate mechanisms of communication became available. Prehistoric people could communicate orally, of course, but each new generation started from essentially the same baseline because, without written records to refer to, knowledge was lost almost as rapidly as it was found.

Cave paintings and inscriptions carved onto rocks were among the first human attempts to leave records for succeeding generations. In a sense, today we are lucky that our early ancestors chose such media because some of these early "messages" have survived, whereas messages on less-durable materials would have been lost. (Perhaps many have been.) On the other hand, communication via such media was [illegible]

the U.S. Postal Service would have today if the medium of correspondence were 100-lb rocks. They have enough troubles with ½-oz letters.

The earliest book we know of is a Chaldean account of the Flood. This story was inscribed on a clay tablet in about 4000 B.C., antedating Genesis by some 2,000 years (Tuchman, 1980).

A medium of communication that was lightweight and portable was needed. The first successful medium was papyrus (sheets made from the papyrus plant and glued together to form a roll sometimes 20 to 40 ft long, fastened to a wooden roller), which came into use about 2000 B.C. In 190 B.C., parchment (made from animal skins) came into use. The Greeks assembled large libraries in Ephesus and Pergamum (in what is now Turkey) and in Alexandria. According to Plutarch, the library in Pergamum contained 200,000 volumes in 40 B.C. (Tuchman, 1980).

In 105 A.D. the Chinese invented paper, the modern medium of communication. However, because there was no effective way of duplicating communications, scholarly knowledge could not be widely disseminated.

Perhaps the greatest single invention in the intellectual history of the human race was the printing press. Although movable type was invented in China in about 1100 A.D. (Tuchman, 1980), the Western World gives credit to Gutenberg, who printed his 42-line Bible from movable type on a printing press in 1455 A.D. Gutenberg's invention was effectively and immediately put to use throughout Europe. By the year 1500, thousands of copies of hundreds of books (called "incunabula") were printed.

The first scientific journals appeared in 1665, when coincidentally two different journals commenced publication, the *Journal des Sçavans* in France and the *Philosophical Transactions of the Royal Society of London* in England. Since that time, journals have served as the primary means of communication in the sciences. Currently, some 70,000 scientific and technical journals are published throughout the world (King et al., 1981).

THE IMRAD STORY

The early journals published papers that we call "descriptive." Typically, a scientist would report that "First, I saw this, and then I saw that" or "First, I did this, and then I did that." Often the observations were in simple chronological order.

This descriptive style was appropriate for the kind of science then being reported. In fact, this straightforward style of reporting is still used today in "letters" journals, in case reports in medicine, in geological surveys, etc.

By the second half of the 19th Century, science was beginning to move fast and in increasingly sophisticated ways. Especially because of the work of Louis Pasteur, who confirmed the germ theory of disease and who developed pure-culture methods of studying microorganisms, both science and the reporting of science made great advances.

At this time, methodology became all-important. To quiet his critics, many of whom were fanatic believers in the theory of spontaneous generation, Pasteur found it necessary to describe his experiments in exquisite detail. Because reasonably competent peers could reproduce Pasteur's experiments, the principle of *reproducibility of experiments* became a fundamental tenet of the philosophy of science, and a segregated methods section led the way toward the highly structured IMRAD format.

Because I have been close to the science of microbiology for many years, it is possible that I overemphasize the importance of this branch of science. Nonetheless, I truly believe that the conquest of infectious disease has been the greatest advance in the history of science. I further believe that a brief retelling of this story may illustrate science and the reporting of science. Those who believe that atomic energy, or molecular biology, is the "greatest advance" might still appreciate the paradigm of modern science provided by the infectious disease story.

The work of Pasteur was followed, in the early 1900s, by the work of Paul Ehrlich and, in the 1930s, by the work of Gerhard Domagk (sulfa drugs). World War II prompted the development of penicillin (first described by Alexander Fleming in 1929). Streptomycin was reported in 1944, and soon after World War II the mad but wonderful search for "miracle drugs" produced the tetracyclines and dozens of other effective antibiotics. Thus, these developments led to the virtual elimination of the scourges of tuberculosis, septicemia, diphtheria, the plagues, typhoid, and (through vaccination) smallpox and infantile paralysis (polio).

As these miracles were pouring out of our medical research laboratories after World War II, it was logical that investment in research

would greatly increase. This positive inducement to support science was soon (in 1957) joined by a negative factor when the Soviets flew Sputnik around our planet. In the following years, whether from hope of more "miracles" or fear of the Soviets, the U.S. government (and others) poured additional billions of dollars into scientific research.

Money produced science. And science produced papers. Mountains of them. The result was powerful pressure on the existing (and the many new) journals. Journal editors, in self-defense if for no other reason, began to demand that manuscripts be tightly written and well organized. Journal space became too precious to waste on verbosity or redundancy. The IMRAD format, which had been slowly progressing since the latter part of the 19th Century, now came into almost universal use in research journals. Some editors espoused IMRAD because they became convinced that it was the simplest and most logical way to communicate research results. Other editors, perhaps not convinced by the simplistic logic of IMRAD, nonetheless hopped on the bandwagon because the rigidity of IMRAD did indeed save space (and expense) in the journals and because IMRAD made life easier for editors and referees (also known as reviewers) by "indexing" the major parts of a manuscript.

The logic of IMRAD can be defined in question form: What question (problem) was studied? The answer is the Introduction. How was the problem studied? The answer is the Methods. What were the findings? The answer is the Results. What do these findings mean? The answer is the Discussion.

It now seems clear to us that the simple logic of IMRAD does help the author to organize and write the manuscript, and IMRAD provides an easy road map for editors, referees, and ultimately readers to follow in reading the paper.

Chapter 3
What Is a Scientific Paper?

Without publication, science is dead.
—Gerard Piel

◇ ◇ ◇

DEFINITION OF A SCIENTIFIC PAPER

A scientific paper is a written and published report describing original research results. That short definition must be qualified, however, by noting that a scientific paper must be written in a certain way and it must be published in a certain way, as defined by three centuries of developing tradition, editorial practice, scientific ethics, and the interplay of printing and publishing procedures.

To properly define "scientific paper," we must define the mechanism that creates a scientific paper, namely, valid (i.e., primary) publication. Abstracts, theses, conference reports, and many other types of literature are published, but such publications do not normally meet the test of valid publication. Further, even if a scientific paper meets all of the other tests (discussed below), it is not validly published if it is published in the wrong place. That is, a relatively poor research report, but one that meets the tests, is validly published if accepted and published in the right place (a primary journal or other primary publication); a superbly prepared research report is not validly published if published in the wrong place. Most of the government report literature and conference literature, as well as institutional bulletins and other ephemeral publications, do not qualify as primary literature.

Many people have struggled with the definition of primary publication (valid publication), from which is derived the definition of a scientific paper. The Council of Biology Editors (CBE), an authoritative professional organization (in biology, at least) dealing with such problems, arrived at the following definition (Council of Biology Editors, 1968):

> An acceptable primary scientific publication must be the first disclosure containing sufficient information to enable peers (1) to assess observations, (2) to repeat experiments, and (3) to evaluate intellectual processes; moreover, it must be susceptible to sensory perception, essentially permanent, available to the scientific community without restriction, and available for regular screening by one or more of the major recognized secondary services (e.g., currently, Biological Abstracts, Chemical Abstracts, Index Medicus, Excerpta Medica, Bibliography of Agriculture, etc., in the United States and similar services in other countries).

At first reading, this definition may seem excessively complex, or at least verbose. But those of us who had a hand in drafting it weighed each word carefully, and we doubt that an acceptable definition could be provided in appreciably fewer words. Because it is important that students, authors, editors, and all others concerned understand what a scientific paper is and what it is not, it may be helpful to work through this definition to see what it really means.

"An acceptable primary scientific publication" must be "the first disclosure." Certainly, first disclosure of new research data often takes place via oral presentation at a scientific meeting. But the thrust of the CBE statement is that disclosure is more than disgorgement by the author; effective first disclosure is accomplished *only* when the disclosure takes a form that allows the peers of the author (either now or in the future) to fully comprehend and use that which is disclosed.

Thus, sufficient information must be presented so that potential users of the data can (1) assess observations, (2) repeat experiments, and (3) evaluate intellectual processes. (Are the author's conclusions justified by the data?) Then, the disclosure must be "susceptible to sensory perception." This may seem an awkward phrase, because in normal practice it simply means published; however, this definition provides for disclosure not just in terms of visual materials (printed journals, microfilm, microfiche) but also perhaps in nonprint, nonvisual forms. For example, "publication" in the form of audio cassettes, if that publication met the other tests provided in the definition, would constitute effective publication. And, certainly, the new "electronic journals" (such as *The Online Journal of Current Clinical Trials*, which commenced "publication" in 1992) meet the definition of valid publication.

Regardless of the form of publication, that form must be essentially permanent, must be made available to the scientific community without restriction, and must be made available to the information retrieval services (*Biological Abstracts, Chemical Abstracts, Index Medicus,* etc.). Thus, publications such as newsletters, corporate publications, and controlled-circulation journals, many of which are of value for their news or other features, cannot serve as repositories for scientific knowledge.

To restate the CBE definition in simpler but not more accurate terms, primary publication is (1) the first publication of original research results, (2) in a form whereby peers of the author can repeat the experiments and test the conclusions, and (3) in a journal or other source document readily available within the scientific community. To understand this definition, however, we must add an important caveat. The part of the definition that refers to "peers of the author" is accepted as meaning prepublication peer review. Thus, by definition, scientific papers are published in peer-reviewed publications.

I have belabored this question of definition for two reasons. First, the entire community of science has long labored with an inefficient, costly system of scientific communication precisely because it (authors, editors, publishers) has been unable or unwilling to define primary publication. As a result, much of the literature is buried in meeting abstracts, obscure conference reports, government documents, or books or journals of minuscule circulation. Other papers, in the same or slightly altered form, are published more than once; occasionally, this is due to the lack of definition as to which conference reports, books, and compilations are (or should be) primary publications and which are not. Redundancy and confusion result. Second, a scientific paper is, by definition, a particular kind of document containing certain specified kinds of information in a prescribed (IMRAD) order. If the graduate student or the budding scientist (and even some of those scientists who have already published many papers) can fully grasp the significance of this definition, the writing task should be a good deal easier. Confusion

results from an amorphous task. The easy task is the one in which you know exactly what must be done and in exactly what order it must be done.

ORGANIZATION OF A SCIENTIFIC PAPER

A scientific paper is a paper organized to meet the needs of valid publication. It is, or should be, highly stylized, with distinctive and clearly evident component parts. The most common labeling of the component parts, in the basic sciences, is Introduction, Methods, Results, and Discussion (hence, the acronym IMRAD). Actually, the heading "Materials and Methods" may be more common than the simpler "Methods," but it is the latter form that was fixed in the acronym.

I have taught and recommended the IMRAD approach for many years. Until recently, however, there have been several somewhat different systems of organization that were preferred by some journals and some editors. The tendency toward uniformity has increased since the IMRAD system was prescribed as a standard by the American National Standards Institute, first in 1972 and again in 1979 (American National Standards Institute, 1979). A recent variation in IMRAD has been introduced by *Cell* and several other journals. In this variation, methods appear last rather than second. Perhaps we should call this IRDAM.

The basic IMRAD order is so eminently logical that, increasingly, it is used for many other types of expository writing. Whether one is writing an article about chemistry, archeology, economics, or crime in the streets, the IMRAD format is often the best choice.

This is generally true for papers reporting laboratory studies. There are, of course, exceptions. As examples, reports of field studies in the earth sciences and clinical case reports in the medical sciences do not readily lend themselves to this kind of organization. However, even in these "descriptive" papers, the same logical progression from problem to solution is often appropriate.

Occasionally, the organization of even laboratory papers must be different. If a number of methods were used to achieve directly related results, it might be desirable to combine the Materials and Methods and the Results into an integrated "Experimental" section. Rarely, the results might be so complex or provide such contrasts that immediate discussion seems necessary, and a combined Results and Discussion section might then be desirable. In addition, many primary journals publish "Notes" or "Short Communications," in which the IMRAD organization is abridged.

Various types of organization are used in descriptive areas of science. To determine how to organize such papers, and which general headings to use, you will need to refer to the Instructions to Authors of your target journal. If you are in doubt as to the journal, or if the journal publishes widely different kinds of papers, you can obtain general information from appropriate source books. For example, the several major types of medical papers are described in detail by Huth (1990), and the many types of engineering papers and reports are outlined by Michaelson (1990).

In short, I take the position that the preparation of a scientific paper has less to do with literary skill than with *organization*. A scientific paper is not literature. The preparer of a scientific paper is not an author in the literary sense.

Some of my old-fashioned colleagues think that scientific papers should be literature, that the style and flair of an author should be clearly evident, and that variations in style encourage the interest of the reader. I disagree. I think scientists should indeed be interested in reading literature, and perhaps even in writing literature, but the communication of research results is a more prosaic procedure. As Booth (1981) put it, "Grandiloquence has no place in scientific writing."

Today, the average scientist, to keep up with a field, must examine the data reported in a very large number of papers. Therefore, scientists (and of course editors) must demand a system of reporting data that is uniform, concise, and readily understandable.

OTHER DEFINITIONS

If "scientific paper" is the term for an original research report, how should this be distinguished from research reports that are not original, or are not scientific, or somehow fail to qualify as scientific papers? Several specific terms are commonly used: "review paper," "conference report," and "meeting abstract."

A review paper may review almost anything, most typically the recent work in a defined subject area or the work of a particular individual or group. Thus, the review paper is designed to summarize,

analyze, evaluate, or synthesize information that *has already been published* (research reports in primary journals). Although much or all of the material in a review paper has previously been published, the spectre of dual publication does not normally arise because the review nature of the work is usually obvious (often in the title of the publication, such as *Microbiological Reviews, Annual Review of Biochemistry,* etc.). Do not assume, however, that reviews contain nothing new. From the best review papers come new syntheses, new ideas and theories, and even new paradigms.

A conference report is a paper published in a book or journal as part of the proceedings of a symposium, national or international congress, workshop, roundtable, or the like. Such conferences are normally not designed for the presentation of original data, and the resultant proceedings (in a book or journal) do not qualify as primary publications. Conference presentations are often review papers, presenting reviews of the recent work of particular scientists or recent work in particular laboratories. Some of the material reported at some conferences (especially the exciting ones) is in the form of preliminary reports, in which new, original data are reported, often accompanied by interesting speculation. But, usually, these preliminary reports do not qualify, nor are they intended to qualify, as scientific papers. Later, often much later, such work is validly published in a primary journal; by this time, the loose ends have been tied down, all essential experimental details are recorded (so that a competent worker could repeat the experiments), and previous speculation has matured into conclusions.

Therefore, the vast conference literature that appears in print normally is not *primary*. If original data are presented in such contributions, the data can and should be published (or republished) in an archival (primary) journal. Otherwise, the information may effectively be lost. If publication in a primary journal follows publication in a conference report, there may be copyright and permission problems (*see* Chapter 26), but the more fundamental problem of dual publication (duplicate publication of original data) normally does not and should not arise.

Meeting abstracts, like conference proceedings, are of several types. Conceptually, however, they are similar to conference reports in that they can and often do contain original information. They are not primary publications, nor should publication of an abstract be considered a bar to later publication of the full report.

In the past, there has been little confusion regarding the typical one-paragraph abstracts published as part of the program or distributed along with the program of a national meeting or international congress. It was usually understood that the papers presented at these meetings would later be submitted for publication in primary journals. More recently, however, there has been a strong trend towards extended abstracts (or "synoptics"). Because publishing all of the full papers presented at a large meeting, such as a major international congress, is very expensive, and because such publication is still not a substitute for the valid publication offered by the primary journal, the movement to extended abstracts makes a great deal of sense. The extended abstract can supply virtually as much information as a full paper; all that it lacks is the experimental detail. However, precisely because it lacks experimental detail, it cannot qualify as a scientific paper.

Those involved with publishing these materials should see the importance of careful definition of the different types of papers. More and more publishers, conference organizers, and individual scientists are beginning to agree on these basic definitions, and their general acceptance will greatly clarify both primary and secondary communication of scientific information.

Chapter 4
How to Prepare the Title

First impressions are strong impressions; a title ought therefore to be well studied, and to give, so far as its limits permit, a definite and concise indication of what is to come.

—T. Clifford Allbutt

◇ ◇ ◇

IMPORTANCE OF THE TITLE

In preparing a title for a paper, the author would do well to remember one salient fact: That title will be read by thousands of people. Perhaps few people, if any, will read the entire paper, but many people will read the title, either in the original journal or in one of the secondary (abstracting and indexing) publications. Therefore, all words in the title should be chosen with great care, and their association with one another must be carefully managed. Perhaps the most common error in defective titles, and certainly the most damaging in terms of comprehension, is faulty syntax (word order).

What is a good title? I define it as the fewest possible words that adequately describe the contents of the paper.

Remember that the indexing and abstracting services depend heavily on the accuracy of the title. An improperly titled paper may be virtually lost and never reach its intended audience.