Writing the Scientific Paper

Introduction

When you write about scientific topics to specialists in a particular scientific field, we call that scientific writing. (When you write to non-specialists about scientific topics, we call that science writing.)

The scientific paper has developed over the past three centuries into a tool to communicate the results of scientific inquiry. The main audience for scientific papers is extremely specialized. The purpose of these papers is twofold: to present information so that it is easy to retrieve, and to present enough information that the reader can duplicate the scientific study. A standard format with six main parts helps readers to find expected information and analysis:

- Title--subject and what aspect of the subject was studied.
- Abstract--summary of paper: The main reason for the study, the primary results, the main conclusions
- Introduction--why the study was undertaken
- Methods and Materials--how the study was undertaken
- Results--what was found
- Discussion--why these results could be significant (what the reasons might be for the patterns found or not found)

There are many ways to approach the writing of a scientific paper, and no one way is right. Many people, however, find that drafting chunks in this order works best: Results, Discussion, Introduction, Materials & Methods, Abstract, and, finally, Title.

Title

The title should be very limited and specific. Really, it should be a pithy summary of the article's main focus.
Abstract

This is a summary of your article. Generally between 50-100 words, it should state the goals, results, and the main conclusions of your study. You should list the parameters of your study (when and where was it conducted, if applicable; your sample size; the specific species, proteins, genes, etc., studied). Think of the process of writing the abstract as taking one or two sentences from each of your sections (an introductory sentence, a sentence stating the specific question addressed, a sentence listing your main techniques or procedures, two or three sentences describing your results, and one sentence describing your main conclusion).

Example One

Hypertension, diabetes and hyperlipidemia are risk factors for life-threatening complications such as end-stage renal disease, coronary artery disease and stroke. Why some patients develop complications is unclear, but only susceptibility genes may be involved. To test this notion, we studied crosses involving the fawn-hooded rat, an animal model of hypertension that develops chronic renal failure. Here, we report the localization of two genes, \textit{Rf-1} and \textit{Rf-2}, responsible for about half of the genetic variation in key indices of renal impairment. In addition, we localize a gene, \textit{Bphh-1}, responsible for about 26\% of the genetic variation in blood pressure. \textit{Rf-1} strongly affects the risk of renal impairment, but has no significant effect on blood pressure. Our results show that susceptibility to a complication of hypertension is under at least partially independent genetic control from susceptibility to hypertension itself.


Example Two

We studied survival of 220 calves of radiocollared moose (\textit{Alces alces}) from parturition to the end of July in southcentral Alaska from 1994 to 1997. Prior studies established that predation by brown bears (\textit{Ursus arctos}) was the primary cause of mortality of moose calves in the region. Our objectives were to characterize vulnerability of moose calves to predation as influenced by age, date, snow depths, and previous reproductive success of the mother. We also tested the hypothesis that survival of twin moose calves was independent and identical to that of single calves. Survival of moose calves from parturition through July was 0.27 ± 0.03 SE, and their daily rate of mortality declined at a near constant rate with age in that period. Mean annual survival was 0.22 ± 0.03 SE. Previous winter's
snow depths or survival of the mother's previous calf was not related to neonatal survival. Selection for early parturition was evidenced in the 4 years of study by a 6.3% increase in the hazard of death with each daily increase in parturition date. Although there was no significant difference in survival of twin and single moose calves, most twins that died disappeared together during the first 15 days after birth and independently thereafter, suggesting that predators usually killed both when encountered up to that age.

Key words: Alaska, Alces alces, calf survival, moose, Nelchina, parturition synchrony, predation


Example Three

We monitored breeding phenology and population levels of Rana yavapaiensis by use of repeated egg mass censuses and visual encounter surveys at Agua Caliente Canyon near Tucson, Arizona, from 1994 to 1996. Adult counts fluctuated erratically within each year of the study but annual means remained similar. Juvenile counts peaked during the fall recruitment season and fell to near zero by early spring. Rana yavapaiensis deposited eggs in two distinct annual episodes, one in spring (March-May) and a much smaller one in fall (September-October). Larvae from the spring deposition period completed metamorphosis in early summer. Over the two years of study, 96.6% of egg masses successfully produced larvae. Egg masses were deposited during periods of predictable, moderate stream flow, but not during seasonal periods when flash flooding or drought were likely to affect eggs or larvae. Breeding phenology of Rana yavapaiensis is particularly well suited for life in desert streams with natural flow regimes which include frequent flash flooding and drought at predictable times. The exotic predators of R. yavapaiensis are less able to cope with fluctuating conditions. Unaltered stream flow regimes that allow natural fluctuations in stream discharge may provide refugia for this declining ranid frog from exotic predators by excluding those exotic species that are unable to cope with brief flash flooding and habitat drying.


Introduction

The introduction is where you sketch out the background of your study, including why you have investigated the question that you have and how it relates to earlier research that has been done in the field. It may help to think of an introduction as a telescoping focus, where you begin with the broader context and gradually narrow to the specific problem addressed by the report. A typical (and very useful) construction of an introduction proceeds as follows:

1. Open with two or three sentences placing your study subject in context.
   
   **Examples**

   "Echimyid rodents of the genus Proechimys (spiny rats) often are the most abundant and widespread lowland forest rodents throughout much of their range in the Neotropics (Eisenberg 1989). Recent studies suggested that these rodents play an important role in forest dynamics through their activities as seed predators and dispersers of seeds (Adler
"Our laboratory has been involved in the analysis of the HLA class II genes and their association with autoimmune disorders such as insulin-dependent diabetes mellitus. As part of this work, the laboratory handles a large number of blood samples. In an effort to minimize the expense and urgency of transportation of frozen or liquid blood samples, we have designed a protocol that will preserve the integrity of lymphocyte DNA and enable the transport and storage of samples at ambient temperatures." (Torrance, MacLeod & Hache, p. 64)

2. Follow with a description of the problem and its history, including previous research.
   • **Examples**

   "Despite the ubiquity and abundance of *P. semispinosus*, only two previous studies have assessed habitat use, with both showing a generalized habitat use. [brief summary of these studies]." (Lambert and Adler, p. 70)

   "Although very good results have been obtained using polymerase chain reaction (PCR) amplification of DNA extracted from dried blood spots on filter paper (1,4,5,8,9), this preservation method yields limited amounts of DNA and is susceptible to contamination." (Torrance, MacLeod & Hache, p. 64)

3. Describe how your work addresses a gap in existing knowledge or ability (here's where you'll state why you've undertaken this study).
   • **Examples**

   "No attempt has been made to quantitatively describe microhabitat characteristics with which this species may be associated. Thus, specific structural features of secondary forests that may promote abundance of spiny rats remains unknown. Such information is essential to understand the role of spiny rats in Neotropical forests, particularly with regard to forest regeneration via interactions with seeds." (Lambert and Adler, p. 71)

   "As an alternative, we have been investigating the use of lyophilization ("freeze-drying") of whole blood as a method to preserve sufficient amounts of genomic DNA to perform PCR and Southern Blot analysis." (Torrance, MacLeod & Hache, p. 64)

4. State what information your article will address.
   • **Examples**

   "We present an analysis of microhabitat use by *P. semispinosus* in tropical moist forests in central Panama." (Lambert and Adler, p. 71)

   "In this report, we summarize our analysis of genomic DNA extracted from lyophilized whole blood." (Torrance, MacLeod & Hache, p. 64)

**Methods and Materials**
In this section you describe how you performed your study. You need to provide enough information here for the reader to duplicate your experiment. However, be reasonable about who the reader is. Assume that he or she is someone familiar with the basic practices of your field.

It's helpful to both writer and reader to organize this section chronologically: that is, describe each procedure in the order it was performed. For example, DNA-extraction, purification, amplification, assay, detection. Or, study area, study population, sampling technique, variables studied, analysis method.

Include in this section:

- study design: procedures should be listed and described, or the reader should be referred to papers that have already described the used procedure
- particular techniques used and why, if relevant
- modifications of any techniques; be sure to describe the modification
- specialized equipment, including brand-names
- temporal, spatial, and historical description of study area and studied population
- assumptions underlying the study
- statistical methods, including software programs

**Example description of activity**

Chromosomal DNA was denatured for the first cycle by incubating the slides in 70% deionized formamide; 2x standard saline citrate (SSC) at 70°C for 2 min, followed by 70% ethanol at -20°C and then 90% and 100% ethanol at room temperature, followed by air drying. (Rouwendal et al., p. 79)

**Example description of assumptions**

We considered seeds left in the petri dish to be unharvested and those scattered singly on the surface of a tile to be scattered and also unharvested. We considered seeds in cheek pouches to be harvested but not cached, those stored in the nestbox to be larderhoarded, and those buried in caching sites within the arena to be scatterhoarded. (Krupa and Geluso, p. 99)

**Examples of use of specialized equipment**

1. Oligonucleotide primers were prepared using the Applied Biosystems Model 318A (Foster City, CA) DNA Synthesizer according to the manufacturers' instructions. (Rouwendal et al., p. 78)

2. We first visually reviewed the complete song sample of an individual using spectrograms produced on a Princeton Applied Research Real Time Spectrum Analyzer (model 4512). (Peters et al., p. 937)

**Example of use of a certain technique**

Frogs were monitored using visual encounter transects (Crump and Scott, 1994). (Sartorius and Rosen, p. 269)
Example description of statistical analysis

We used Wilcox rank-sum tests for all comparisons of pre-experimental scores and for all comparisons of hue, saturation, and brightness scores between various groups of birds ... All $P$-values are two-tailed unless otherwise noted. (Brawner et al., p. 955)

Results

This section presents the facts--what was found in the course of this investigation. Detailed data--measurements, counts, percentages, patterns--usually appear in tables, figures, and graphs, and the text of the section draws attention to the key data and relationships among data. Three rules of thumb will help you with this section:

- present results clearly and logically
- avoid excess verbiage
- consider providing a one-sentence summary at the beginning of each paragraph if you think it will help your reader understand your data

Remember to use table and figures effectively. But don't expect these to stand alone.

Examples

Some examples of well-organized and easy-to-follow results:

- Size of the aquatic habitat at Agua Caliente Canyon varied dramatically throughout the year. The site contained three rockbound tinajas (bedrock pools) that did not dry during this study. During periods of high stream discharge seven more seasonal pools and intermittent stretches of riffle became available. Perennial and seasonal pool levels remained stable from late February through early May. Between mid-May and mid-July seasonal pools dried until they disappeared. Perennial pools shrank in surface area from a range of 30-60 m$^2$ to 3-5- M$^2$. (Sartorius and Rosen, Sept. 2000: 269)

Notice how the second sample points out what is important in the accompanying figure. It makes us aware of relationships that we may not have noticed quickly otherwise and that will be important to the discussion.

A similar test result is obtained with a primer derived from the human $\beta$-satellite... This primer (AGTGCAGAGATATGTCACAATG-CCCC: Oligo 435) labels 6 sites in the PRINS reaction: the chromosomes 1, one pair of acrocentrics and, more weakly, the chromosomes 9 (Fig. 2a). After 10 cycles of PCR-IS, the number of sites labeled has doubled (Fig. 2b); after 20 cycles, the number of sites labeled is the same but the signals are stronger (Fig. 2c)... (Rouwendal et al., July 93:80)

Related Information: Use Tables and Figures Effectively

Do not repeat all of the information in the text that appears in a table, but do summarize it. For example, if you present a table of temperature measurements taken at various times, describe the general pattern of temperature change and refer to the table.
"The temperature of the solution increased rapidly at first, going from 50º to 80º in the first three minutes (Table 1)."

You don't want to list every single measurement in the text ("After one minute, the temperature had risen to 55º. After two minutes, it had risen to 58º," etc.). There is no hard and fast rule about when to report all measurements in the text and when to put the measurements in a table and refer to them, but use your common sense. Remember that readers have all that data in the accompanying tables and figures, so your task in this section is to highlight key data, changes, or relationships.

**Discussion**

In this section you discuss your results. What aspect you choose to focus on depends on your results and on the main questions addressed by them. For example, if you were testing a new technique, you will want to discuss how useful this technique is: how well did it work, what are the benefits and drawbacks, etc. If you are presenting data that appear to refute or support earlier research, you will want to analyze both your own data and the earlier data--what conditions are different? how much difference is due to a change in the study design, and how much to a new property in the study subject? You may discuss the implication of your research--particularly if it has a direct bearing on a practical issue, such as conservation or public health.

This section centers on *speculation*. However, this does not free you to present wild and haphazard guesses. Focus your discussion around a particular question or hypothesis. Use subheadings to organize your thoughts, if necessary.

This section depends on a logical organization so readers can see the connection between your study question and your results. One typical approach is to make a list of all the ideas that you will discuss and to work out the logical relationships between them--what idea is most important? or, what point is most clearly made by your data? what ideas are subordinate to the main idea? what are the connections between ideas?

**Achieving the Scientific Voice**

Eight tips will help you match your style for most scientific publications.

1. Develop a precise vocabulary: read the literature to become fluent, or at least familiar with, the sort of language that is standard to describe what you're trying to describe.
2. Be as precise as possible: limit language.
   - Once you've labeled an activity, a condition, or a period of time, use that label consistently throughout the paper. Consistency is more important than creativity.
   - Define your terms and your assumptions.
3. Be honest about the limitations of your knowledge or your research: give the reader enough information to come to the same conclusions that you did (or to come to different conclusions)
   - Include all the information the reader needs to interpret your data.
• Remember, the key to all scientific discourse is that it be reproducible. Have you presented enough information clearly enough that the reader could reproduce your experiment, your research, or your investigation?

4. When describing an activity, break it down into elements that can be described and labeled, and then present them in the order they occurred.

5. When you use numbers, use them effectively. Don't present them so that they cause more work for the reader.

6. Include details before conclusions, but only include those details you have been able to observe by the methods you have described. Do not include your feelings, attitudes, impressions, or opinions.

7. Research your format and citations: do these match what have been used in current relevant journals?

8. Run a spellcheck and proofread carefully. Read your paper out loud, and/ or have a friend look over it for misspelled words, missing words, etc.

Related Information: Applying the Principles, Example 1

The following example needs more precise information. Look at the original and revised paragraphs to see how revising with these guidelines in mind can make the text clearer and more informative:

Before:

Each male sang a definite number of songs while singing. They start with a whistle and then go from there. Each new song is always different, but made up an overall repertoire that was completed before starting over again. In 16 cases (84%), no new songs were sung after the first 20, even though we counted about 44 songs for each bird.

After:

Each male used a discrete number of song types in his singing. Each song began with an introductory whistle, followed by a distinctive, complex series of fluty warbles (Fig. 1). Successive songs were always different, and five of the 19 males presented their entire song repertoire before repeating any of their song types (i.e., the first 10 recorded songs revealed the entire repertoire of 10 song types). Each song type recurred in long sequences of singing, so that we could be confident that we had recorded the entire repertoire of commonly used songs by each male. For 16 of the 19 males, no new song types were encountered after the first 20 songs, even though we analyzed and average of 44 songs/male (range 30-59).

Related Information: Applying the Principles, Example 2

In this set of examples, even a few changes in wording result in a more precise second version. Look at the original and revised paragraphs to see how revising with these guidelines in mind can make the text clearer and more informative:

Before:

The study area was on Mt. Cain and Maquilla Peak in British Columbia, Canada. The study area is about 12,000 ha of coastal montane forest. The area is both managed and unmanaged and ranges from 600-1650m. The most common trees present are mountain hemlock (Tsuga mertensiana),
western hemlock (*Tsuga heterophylla*), yellow cedar (*Chamaecyparis nootkatensis*), and amabilis fir (*Abies amabilis*).

**After:**

The study took place on Mt. Cain and Maquilla Peak (50°13'N, 126°18'W), Vancouver Island, British Columbia. The study area encompassed 11,800 ha of coastal montane forest. The landscape consisted of managed and unmanaged stands of coastal montane forest, 600-1650 m in elevation. The dominant tree species included mountain hemlock (*Tsuga mertensiana*), western hemlock (*Tsuga heterophylla*), yellow cedar (*Chamaecyparis nootkatensis*), and amabilis fir (*Abies amabilis*).

**Related Information: Two Tips for Sentence Clarity**

Although you will want to consider more detailed stylistic revisions as you become more comfortable with scientific writing, two tips can get you started:

First, the verb should follow the subject as soon as possible.

**Really Hard to Read:** "The smallest of the URF's (URFA6L), a 207-nucleotide (nt) reading frame overlapping out of phase the NH2-terminal portion of the adenosinetriphosphatase (ATPase) subunit 6 gene has been identified as the animal equivalent of the recently discovered yeast H+-ATPase subunit gene."

**Less Hard to Read:** "The smallest of the UR-F's is URFA6L, a 207-nucleotide (nt) reading frame overlapping out of phase the NH2-terminal portion of the adenosinetriphosphatase (ATPase) subunit 6 gene; it has been identified as the animal equivalent of the recently discovered yeast H+-ATPase subunit 8 gene."

Second, place familiar information first in a clause, a sentence, or a paragraph, and put the new and unfamiliar information later.

**More confusing:** The epidermis, the dermis, and the subcutaneous layer are the three layers of the skin. A layer of dead skin cells makes up the epidermis, which forms the body's shield against the world. Blood vessels, carrying nourishment, and nerve endings, which relay information about the outside world, are found in the dermis. Sweat glands and fat cells make up the third layer, the subcutaneous layer.

**Less confusing:** The skin consists of three layers: the epidermis, the dermis, and the subcutaneous layer. The epidermis is made up of dead skin cells, and forms a protective shield between the body and the world. The dermis contains the blood vessels and nerve endings that nourish the skin and make it receptive to outside stimuli. The subcutaneous layer contains the sweat glands and fat cells which perform other functions of the skin.

**Bibliography**
• *Scientific Writing for Graduate Students.* F. P. Woodford. Bethesda, MD: Council of Biology Editors, 1968. [A manual on the teaching of writing to graduate students--very clear and direct.]


• "A Quick Fix for Figure Legends and Table Headings." Donald Kroodsma. *The Auk,* 117 (4): 1081-1083, 2000.

**Citation Information**


**Copyright Information**

Copyright © 1994-2019 Colorado State University and/or this site's authors, developers, and contributors. Some material displayed on this site is used with permission.