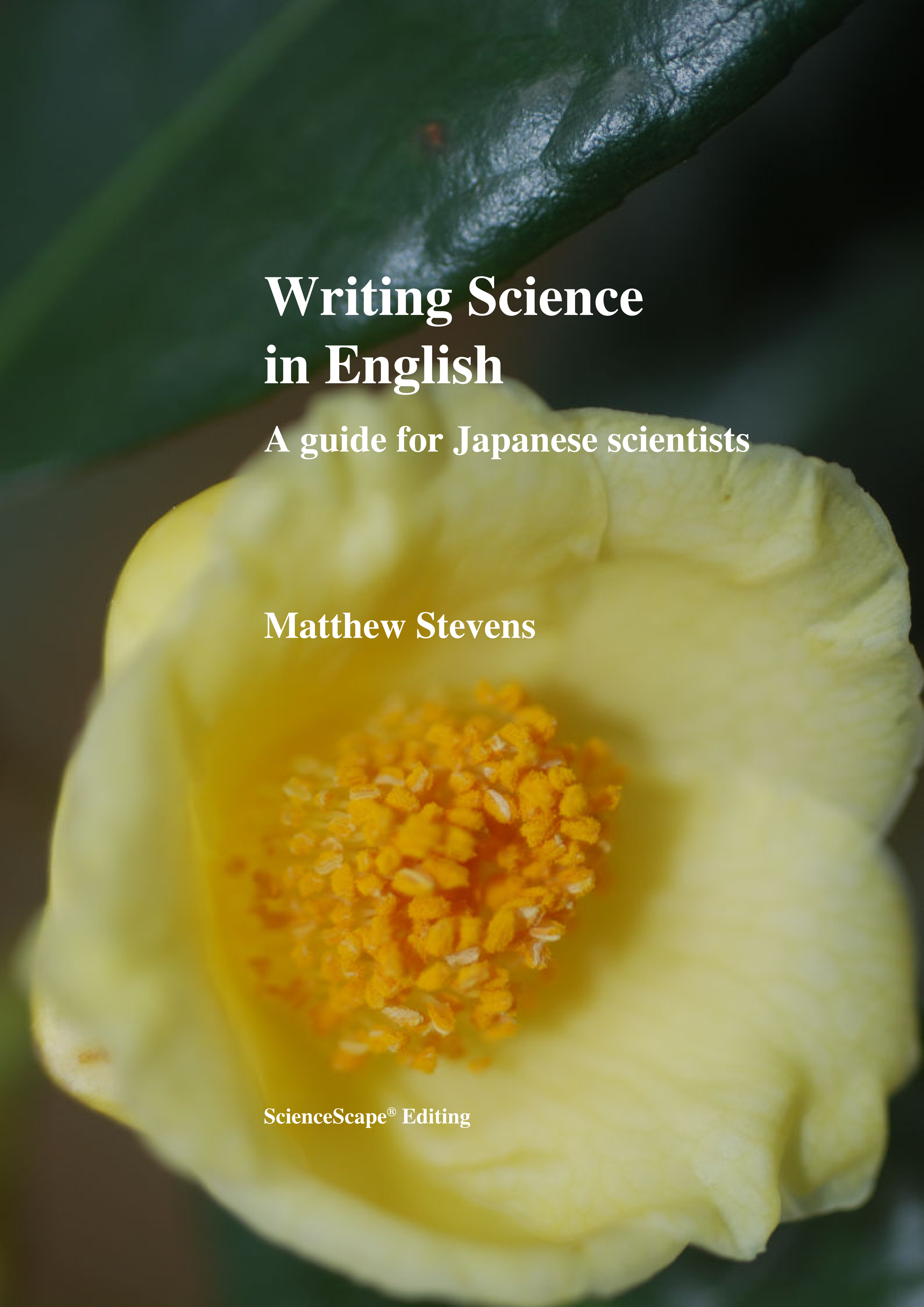


Writing Science in English

A guide for Japanese scientists

Matthew Stevens

ScienceScape® Editing

Foreword

This book is intended as a "how-to" guide for Japanese scientists who are required to write in English for publication. It is based on over 25 years of my experience working as a scientific editor with Japanese authors. It covers all of the common errors I see, and provides a basic structure for writing scientific papers in English, along with examples. It starts with a background to the English language and why it is so annoying. Then it presents the basic structure of a scientific paper. Then it shows examples of how to write and how not to write, and why.

I hope that you find it helpful. You can use it as a book of rules if you want, or you can choose the parts that you like, or you can disagree entirely with it. As the famous English author George Orwell (1903–1950) wrote:

"Break any of these rules sooner than say anything outright barbarous."

Matthew Stevens
ScienceScape Editing
Sydney, New South Wales, Australia
October 2021

If you find this book helpful, please send ¥1000 or $10 via PayPal to mls@mlstevens.net

Acknowledgements

This book had its genesis in a request from my friend and colleague Mr Kazuya Yoshimura, of World Translation Services, in Tsukuba, Japan. My earlier book, *Subtleties of Scientific Style*, which I wrote for other scientific editors like myself, proved popular among some of Kazuya's scientist authors. Kazuya asked me to write something more relevant to authors. The result is this book, although it took me several years of alternating enthusiasm and uncertainty.

I am grateful to Kazuya and his colleagues for the idea and to my friends and colleagues Dr Richard Weisburd (ELSS Inc., Tsukuba) and Mr Geoff Hart (Montréal, Québec, Canada) for their insightful revisions and corrections.

Cartoons on pages 12, 14, 16 and 26 © Makiko Kanazawa. Diagrams are drawn by me. Other illustrations are used under CC licence.

The SAMPL Guidelines (Appendix 4) are included by kind permission of their author, Tom Lang.

Cover: *Camellia nitidissima*. Photo by Matthew Stevens

Published by ScienceScape® Editing, Thornleigh, NSW, Australia
E-mail: mls@mlstevens.net

ISBN 978-0-6487079-3-6 (electronic edition only)

Contents

日本人研究者のためだけに書かれた英語論文ライティング・ガイドブック

英語で論文を書くことは日本語ネイティブの私にはとてもハードルが高く感じます。
翻訳会社の一員として、和文英訳・英文校閲のコーディネーター業務を行っていく中で、私と同じように感じる著者様がとても多くいらっしゃいました。
本著者であるMatthew Stevensさんは、弊社からも英文校閲をお願いしているエディターさんで、日本人の著者様の原稿を25年も見てきた実績があります。
Matthewさんがプロのエディター向けの本「Subtleties of Scientific Style」を数年前に出版した時、私はMatthewさんに今度は日本語ネイティブの研究者向けに英語論文ライティングのガイドブックも書いてもらえないかとお願いしました。そしてついに日本語ネイティブの研究者のために書かれた英語論文のライティング・ガイドブック「Writing Science in English」が出版されることとなりました。

本書は、Matthewさんがこれまでの英文校閲業務で見てきた日本人にありがちな間違いを網羅し、英語や英語論文の構造が日本語とは違うことを例とともに説明しています。つまり、日本語の感覚で英語論文を書いてはダメで、英語圏のスタイルに則って書く必要があるのですが、そのことに気づいていない方が少なくありません。
本書を読むことで、どのように書くべきか、どのように書いてはダメか、またなぜなのかが豊富な例や説明で容易に理解できるでしょう。

すでに英語論文の書き方をご存じの方には、本書には英語表現を向上させるためのセクションもあるため、より簡潔で明瞭な英語表現を確認することができます。

また、本書には英語の歴史といった説明や投稿された論文に日本人の名前を見た時に英語学術誌の査読者が感じることなど興味深い話も盛り込まれており、英語論文に関する予備知識も得られます。

日本に住む英語ネイティブの方が、日本人の間違った英語を指摘する書籍や英語論文の書き方を教えてくれる書籍はたくさんあると思いますが、日本人研究者の英語論文を数多く校閲した経験の中で共通する間違いを何度も見てきたエディターが、日本人研究者向けに英語論文の書き方を教えてくれる書籍は本書だけだと思います。出版されるまでに何年もかかったMatthewさんの思いが、日本人研究者の方々に必ず伝わり、理解してもらえると信じています。

英語の書籍を読むことに抵抗がある方もいらっしゃると思いますが、Matthewさんの英語は不思議と容易に理解できます。気がつけば数十ページも読み進めていることでしょう。ぜひそんなMatthewさんの言葉のマジックもご体験ください。

吉村和也
有限会社ワールド翻訳サービス

Writing Science in English

Two hundred years ago the language of trade and diplomacy was French. In another hundred years, it might be Japanese or Mandarin Chinese. For now, it is English.

To reach an international audience, all scientists must publish in English. English is the first language of 375 million people and the second language of up to 1000 million people. It is the official language in nearly 90 countries and territories. It is the official language of all international air travel and sea travel (en.wikipedia.org/wiki/English_language). For convenience, if nothing else, publishing in English will get your message to the largest number of people.

But English is a natural language, and in common with all natural languages, it is not logical. So I have written this book to help Japanese authors in particular (but other authors also) to write a scientific paper for English-language publications (with a focus on journals). Along with explanations, I have tried to present my advice like a cookbook: If you follow the examples given, you will minimise the problems you will face in publishing your work and maximise the chances that your work will be understood and accepted.

A history of English

English is not logical!

English is illogical and annoying. To understand why, it can help to understand its history.

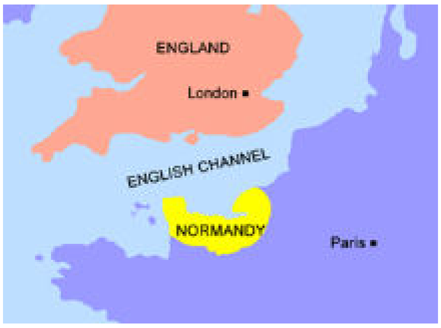
English began life as a trade language. When the ancient Romans left Britain in AD 410, they left no government in charge. With no formal army, the land was open to invasion. So the British rulers invited Anglo-Saxon mercenaries (free-lance soldiers) to come and help. The Anglo-Saxons stayed.

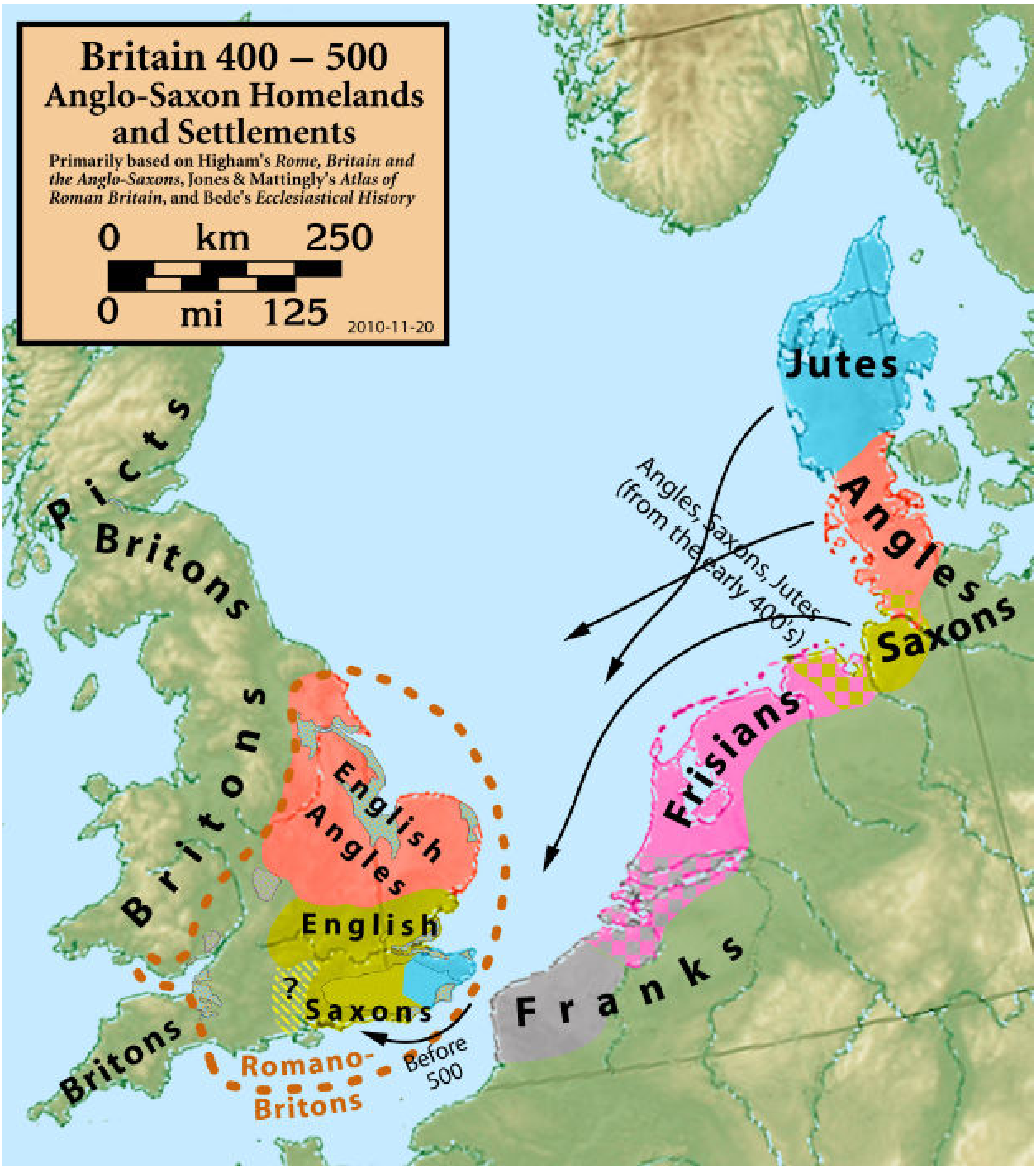
The name “England” and the word “English” come from “Angle”. The Angles came from a region called Angeln, in what is now northern Germany (Holstein), near Denmark. The Saxons came from what is now Lower Saxony (Niedersachsen), in northern Germany, just south of Angeln. A third group of people also arrived: the Jutes, from what is now Jutland (Jylland), in Denmark, just north of Angeln.

English is a Germanic language

Their languages, all from the Germanic family, were similar but not close enough to allow easy communication. So the different groups developed a trading language—a simplified language for trade. That language became English. Therefore English is a Germanic language, and shares many similarities with modern German and Dutch.

Then in 1066, William of Normandy (now part of France) invaded England. His army defeated the English army and William became king. The official language thus became French. Today, 950 years later, the language of government and the law still uses mostly French words.





The Angles became the English.

Anglo Saxon—daily words

French—law

Latin—church

Latin, Greek—science, medicine

Arabic, German—mathematics

Latin words entered English in 3 phases. First, the Romans left many Latin words in the local languages. Then Christianity arrived in Britain, and Latin words entered the language through the dominance of the Christian Church over people's lives. Finally, through the dominance of the Church in the universities in the Middle Ages (5th to 15th centuries), Latin words entered learning and then science and medicine.

The universities were also responsible for introducing Greek words into science and medicine. Today most medical words in English are derived from Latin or Greek. All species binomial names have Latin or Greek roots.

Because of the role of the early Islamic scholars, who translated Greek texts into Arabic, many words in mathematics (such as "algebra") are derived from Arabic.

English belongs to everyone

In the 16th century, Britain (England, Scotland and Wales) began to explore the world and to set up colonies in other countries. And so English imported words from many other countries: the Americas, India, Africa, Australia and more. At the same time, Britain exported English to many other countries, where meanings have slowly changed. Today English is the main language or a major language in nearly 90 countries, and every country has its own version of English, with different spellings, words and pronunciations.

In France, the French Government tries to preserve the French language against invasion by words from other languages (mostly English). The Académie Française bans the use of certain words and introduces French words in their place (although their recommendations have no legal force). There is no equivalent organisation in English. This is because no country owns English. Instead, everyone owns English.

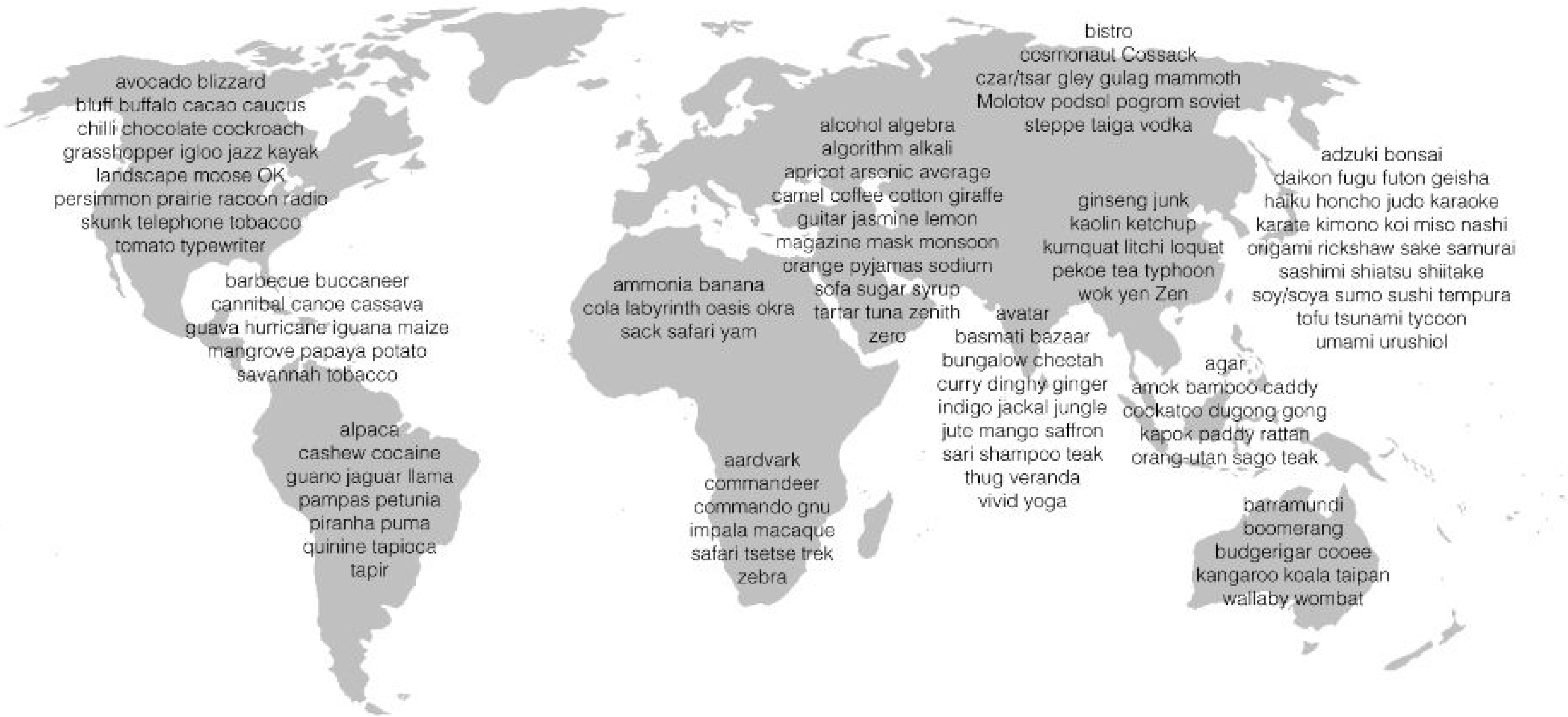
English is flexible

The result of this mixed history is that English vocabulary comes from hundreds of different languages (including Japanese: e.g. kimono, sushi, sashimi, fugu, harakiri, samurai, tamagotchi, origami). Its grammar is the simplified grammar of a trade language (but still tricky to learn). It has at least 500 000 words—more than any other language. Any word can be used in any grammatical position (often leading to confusion). There is more than 1 way to spell many words (e.g. color, colour; inquire, enquire; acknowledgment, acknowledgement; disc, disk), and several different words can have the same meaning (e.g. student, pupil; on, upon; chair, seat; pig, hog, swine), because they came from different places or times. The result is a mess that now dominates world trade and communication.

Basic English is easy; good English is hard

It is easy to learn basic English. It is very hard to learn good English. Even native English speakers have problems. Italian is much easier (at least for Europeans).

English has absorbed thousands of words from all around the world.



UK English or US English?

English is an international language with a history of about 1300 years, so it varies around the world. International scientific journals tend to publish in either US (American) English or UK (British) English. In addition, regional variants include Australian English (like UK English, but not exactly), Canadian English (somewhere between US and UK English), and variants in South Africa, India, Malaysia and Singapore, the Philippines, New Zealand, Ireland, Scotland, Jamaica and many more. You can safely forget all of these. You are unlikely ever to be asked to use them in a scientific paper.

So the choice in writing a paper in English will always come down to either UK English or US English. US English is dominant (and is therefore sometimes referred to incorrectly as "international" English), but UK English is significant. You must always choose one or the other form. Do not mix them (mixing implies uncertainty or lack of attention, even if that is not true).

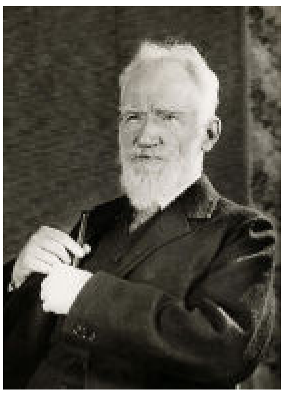
US English ✓
UK English ✓
US + UK English ✗

The journal's instructions to authors will often say which English is required, although you may need to guess this from the name of the dictionary they recommend; Webster's dictionary indicates US English, and Oxford dictionaries indicate UK English. Some journals accept either, but not a mixture of both.

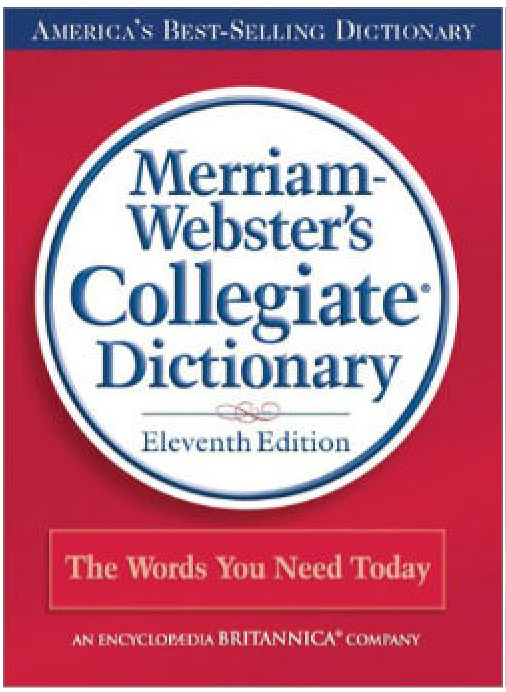
If the instructions to authors don't say, you can make your own choice (which one are you more comfortable writing in?), or you can look for clues in the journal or its instructions to authors. The following list of differences between UK and US English will help you decide which form of English the journal has used:

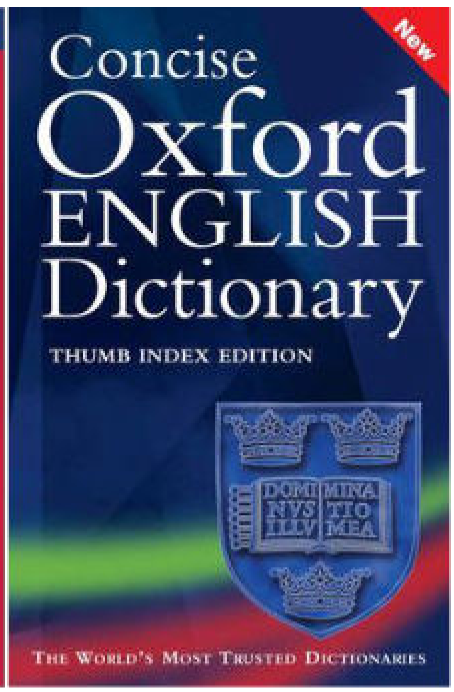
UK	US	Examples
-is-, -iz-	-iz-	organise/organize
-our	-or	colour/color
-yse	-yze	analyse/analyze
practis- (verb)	practic-	practised/practiced
-lled/-lling	-led/-ling	levelled/leveled
haem-	hem-	haematoma/hematoma
-aem-	-em-	anaemia/anemia
-tre	-ter	theatre/theater
-ogue	-og	homologue/homolog
grey	gray	
storey	story	understorey/understory
-wards	-ward	backwards/backward
besides	beside	
litre	liter	
metre	meter	millimetre/millimeter
oes-	es-	oestrogen/estrogen
		oesophagus/esophagus
paed-	ped-	paediatrics/pediatrics
palae-	pale-	palaeontology/paleontology
-or	-er	adaptor/adapter
non-	non_	non-invasive/noninvasive
-oeu-	-eu-	manoeuvre/maneuver
-se	-ce	defence/defense

The Irish author George Bernard Shaw once said: “England and America are two countries divided by a common language.”



Use one or the other.





UK English and US English differ in 3 main areas: spelling of some words, punctuation and idiom (cultural choice of words).

Using the right **spelling** is easy. Just set the language in your word processor. Or you can refer to a dictionary: use *Webster's Collegiate Dictionary* for US English and the *Concise Oxford Dictionary* for UK English.

Punctuation differs mainly in the use of the comma (",") in lists of words. UK English (usually but not always): "Values were 12, 16, 14 and 13 g." US English (usually but not always): "Values were 12, 16, 14, and 13 g." The difference is that US English almost always uses a comma before the last "and". Neither is always right or always wrong, despite what some people will tell you. Both styles have their advantages and their disadvantages.

Another difference occurs in the placement of punctuation in relation to quotation marks. UK English usually uses "logical punctuation": "The top Japanese rice is 'Koshihikari'." (In this example, the full point falls outside the single quotes, which indicate a cultivar.) US English usually uses "traditional punctuation": "The top Japanese rice is 'Koshihikari.'" (In this example, the full point falls inside the single quotes, even though it is not part of the cultivar name.) You can safely leave this difference to the journal editors.

Finally, differences in **idiom** do not often arise in scientific writing. Idioms are phrases that are unique to a dialect and that have evolved differently in different parts of the world. Many rely on knowledge of unusual aspects of a particular culture, such as when an American states that their study's results represent a "home run" (a baseball metaphor that means a great success). For example, one common variant is "in light of" (US English) versus "in the light of" (UK English). Other variants are not important; the journal editors will change them if necessary. (Examples include control/check, catchment/watershed, different from / different than, studies on / studies of.) Because idioms are not always clear, avoid using them.

In summary, choose either US English or UK English. That's all you need to remember.

Latin

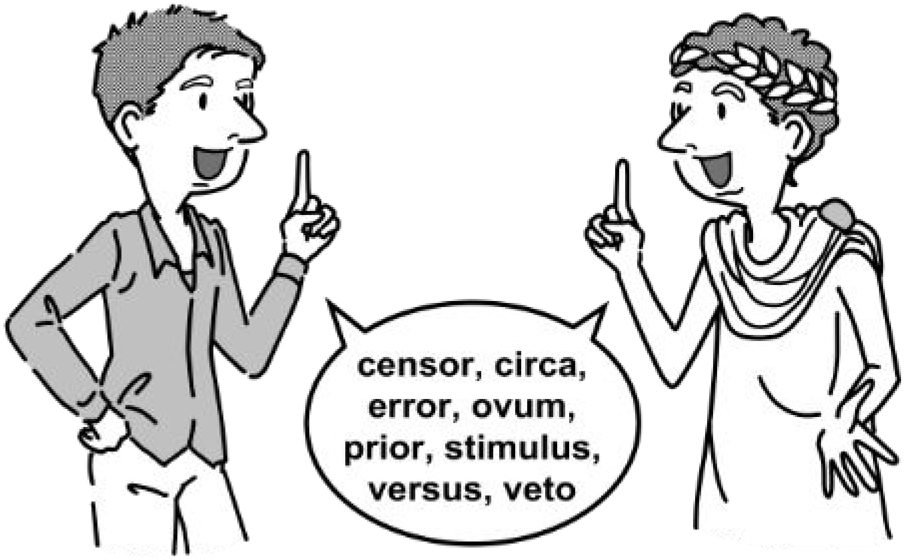
You can use these Latin abbreviations:
e.g.
i.e.
etc.
a.m.
p.m.
et al.

Latin—the language of ancient Rome, 2000 years ago—is part of the ancestry of modern English. (In the same way, Chinese is part of the ancestry of modern Japanese.) Some Latin words are still used in English exactly as they were spelled 2000 years ago; for example, censor, circa, error, ovum, prior, stimulus, versus, veto. But these words are now all English words, and are not treated as Latin.

Use these Latin terms if they are normal in your field of study

Many Latin abbreviations are also used in writing. Unfortunately, many native English speakers don't know what they mean, and often use them the wrong way. This tells us that we should not use abbreviations that the readers might not know. Only 5 Latin abbreviations remain in common use in English: "e.g." (= *exempli gratia* = "for [the sake of] example"), "i.e." (*id est* = "that is"), "etc." (*et cetera* = "and the rest"), "a.m." (*ante meridiem* = "before midday") and "p.m." (*post meridiem* = "after midday"). In addition, scientific writing uses "et

Some words have not changed for 2000 years.



al." (*et alii* = "and others"). You should use English words in place of all other Latin abbreviations (such as *loc. cit.*, *q.v.*, *idem.*, *ibid.*, *cf.*, *op. cit.*).

Some Latin terms also find common use in English:

in vivo	in life, in living material
in vitro	in glass
in utero	in the uterus
in ovo	in the egg
inter alia	among others
ex nihilo	out of nothing
ab initio	from the beginning
per diem	each day, daily
a priori	from before
a posteriori	from after
ad hoc	for one purpose only
ad libitum	to liberty, freely
post mortem	after death
de facto	in fact (not in law)
reductio ad absurdum	reduction (of an argument) to an absurd conclusion
de novo	new
per se	as such
prima facie	at first sight
pro rata	in proportion

If a Latin term is commonly used in your field of study, then use it. But if the readers are not likely to know what it means, then use English. If you are not sure of the right word, then refer to the international English literature to see what native speakers use.

When journal reviewers see Japanese names

Some Western reviewers think, "Japanese author = bad English"

When Japanese authors submit a paper to an English language journal, many journal reviewers see the Japanese names and immediately think, "Japanese authors = bad English", and they reject the paper. I have seen this, even when the English is perfect (because I or my colleagues have already rewritten the paper!). The authors feel upset, and do not understand what is wrong. Sometimes the authors think that I or my colleagues have not done a good job. That is not correct. In every example I have ever seen, *either* the journal editors look at the Japanese names and assume that the English is bad *or* the authors have added more text *after* we have rewritten the paper. These are different problems, and they have different solutions. (It is worth mentioning 2 other minor situations. Sometimes the cover letter is badly written, and the journal editors don't read anything else. And occasionally journal editors forget to delete a standard instruction to use a professional English rewriter.)

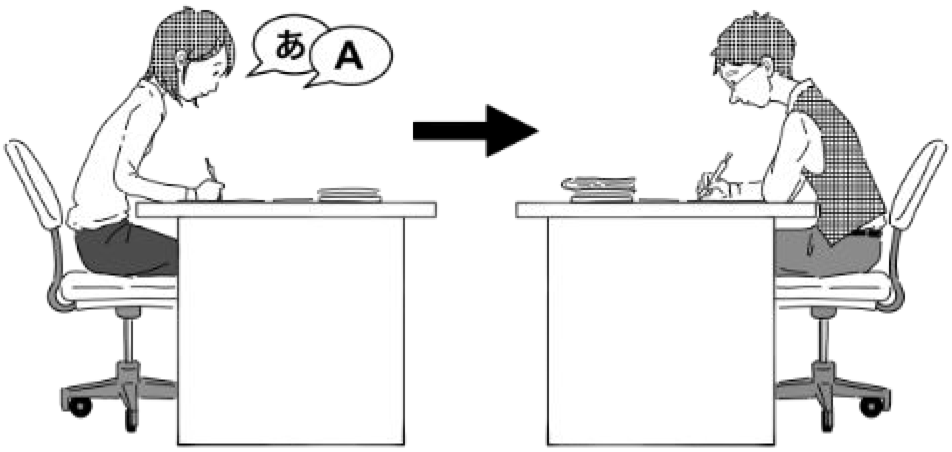
Some Japanese authors write perfect English. I have been fortunate to work with 2 or 3 of them. In fact, they write better English than most native English speakers! But they are the exception. Most Japanese authors need the help of professional rewriters in English. (Most native-English-speaking authors need the help of professional rewriters also.) Even if you hire a skilled translator to produce the English version of a manuscript for you, it is wise to pay a professional rewriter also to help you with the wording. Then, when you submit your paper to the journal, in your covering letter you can say, "The English has been extensively revised by two native-English-speaking scientific editors of Xyz Scientific Language Service."

The other problem can occur if an author adds text *after* the rewriter has rewritten the paper, or misses or rejects their editor's advice, leaving an error in the manuscript. I have seen a few examples of this in which the author is angry because the journal rejects the paper. However, the journal rejects the new, *un-rewritten* text. If you need to add more text, always ask your rewriter to check that text too. Of course, if you believe that a change made by your rewriter is wrong or imprecise, you should explain the problem and ask for clarification rather than simply ignoring the change.

Why use a translator *and* a rewriter?

Why should you pay money for a translator to translate Japanese into English, and then pay more money for a rewriter to rewrite the English into different English? There are good reasons:

- A translator understands both Japanese and English better than most people, and therefore can translate your science into good English.
- A rewriter (scientific editor) understands scientific English better than most people, and therefore can craft your paper into good scientific English.
- Although a scientist can be both a good translator and a good rewriter, nobody can objectively edit their own work.
- A translator and a rewriter have different sets of expertise.



The translator and the rewriter (editor) have different skills.

Good scientific English is not just good English. Using it includes:

- using *exactly* the right word for the intended meaning
- knowing what information to include and what is not needed
- understanding the science—scientific editors are often trained scientists who understand how science works, and who understand what other scientists need to know
- understanding the conventions of science—understanding the many different ways in which science can be presented, particularly the Instructions to Authors, which are all different
- plain writing—explaining complex ideas in a straightforward way
- experience with the review process (i.e. an understanding of the kinds of points that reviewers most often criticise and the ability to help you avoid those criticisms).

History of the scientific paper format

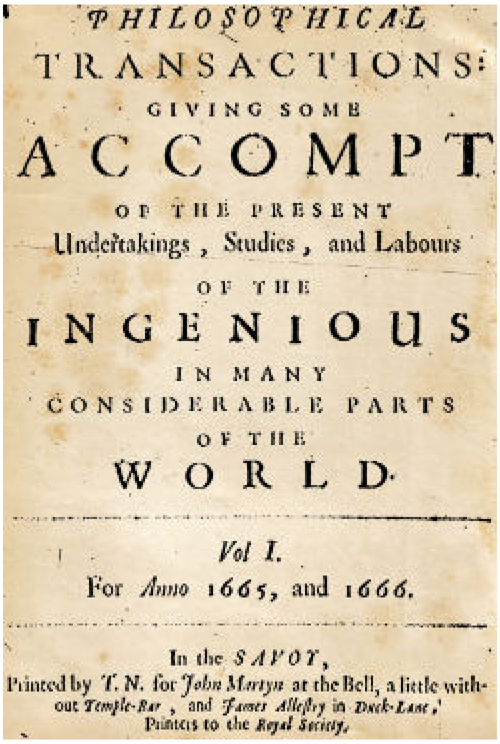
The first scientific journals were published in 1665: the *Philosophical Transactions of the Royal Society* in the UK and the *Journal des sçavans* in France. Both are still in publication today, more than 350 years later.

The first journals were simple accounts of what the authors saw and did. Today many papers still follow this pattern, such as medical case studies.

But as science and understanding grew, authors began to find that they had to provide proof to satisfy their critics, not just ideas. This led to an increase in the degree of mathematical and statistical rigour. In addition, detailed descriptions of methods were included. This allowed experiments to be repeated. Today, repeatability of experiments is an essential feature of science.

~1890:
Introduction
Methods
Results
Discussion

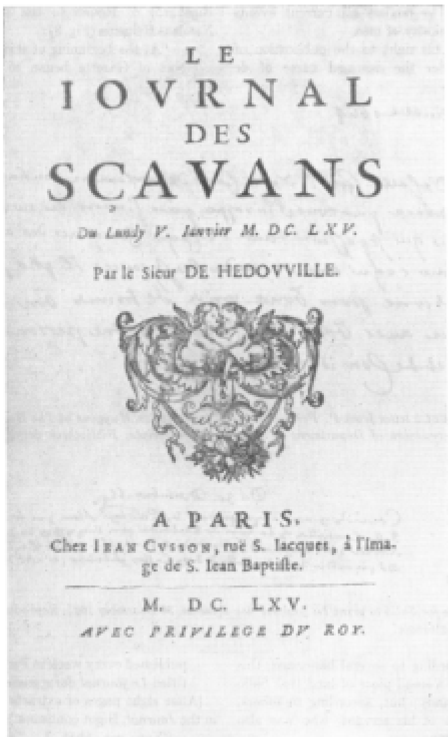
Beginning in the late 19th century (only about 130 years ago), the structure of the modern paper was introduced. The structure is Introduction, Methods, Results and Discussion, often abbreviated as IMRaD. It took until the middle of the 20th century for this structure to be generally accepted. Today, most scientific papers follow the IMRaD format.



PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane, Printers to the Royal Society.



LE
IOVRNAL
DES
SCAVANS
Du Lundy V. Ianvier M. DC. LXV.
Par le Sieur DE HEDOVVILLE.

A PARIS,
Chez IEAN CVSSON, ruë S. Iacques, à l'Image de S. Iean Baptiste.
M. DC. LXV.
AVEC PRIVILEGE DV ROY.

The first scientific journals in the world (MDCLXV = 1665).

The IMRaD format is convenient. It is not the only way you can write a paper, and it does not accurately describe the sequence of events in the study. We can think of it as an idealised sequence that we would have followed if we had known everything at the beginning. In reality, science is messy: it involves false starts (choosing a poor strategy), dead ends (not getting the results you expected), mistakes and surprises. Readers are more interested in a clear description of what you eventually achieved, not in a description of all the detours you took along the way.

Science is messy. IMRaD imposes structure

The IMRaD format is a standardised format. Any scientist can pick up a paper and know where to find particular information. It also sets out the study in a logical order: Why did the authors do the study? How did they do the study? What did they find? What does it mean? Although science is messy, the IMRaD format allows us to understand a study in a logical order, if not the actual order.

Abstract
Introduction
Methods
Results
Discussion

All papers these days include an Abstract. An Abstract is an essential part of a scientific paper. It allows hurried readers, with little free time, to understand the main points of the study and therefore to decide whether or not they want to read it in detail. So we can modify the name to AIMRaD.

Structure of a scientific paper

This section briefly describes the AIMRaD structure. Some studies don't fit easily into this format (for example, computer modelling studies; medical case studies), and some journals specify a different format (the same elements in a different order; or additional elements such as a graphical abstract). If so, follow the instructions for the journal.

Abstract

In about 250 words, introduce the problem, briefly describe what you did, state the main results and draw your main conclusion.

Introduction

Introduce the background to your work, review the literature and state the hypothesis or the main research question or objective.

Materials and Methods

Also called "Experimental Procedures". Here you describe what you did.

Results

Present the results, all the results, and nothing else (no interpretation, no literature citations).

Present the results only once: in words **or** in tables **or** in figures.

Present the results of the statistical analyses.

(Note that the Results can be combined with the Discussion if the total is small and is more easily understood that way.)

Discussion

Restate the reason for the study.

Describe what your results mean. Do they agree with previous results? Do they disagree?

Can you explain any disagreements?

What generalisations can be drawn and how widely are they applicable?

What should future work look at?

Was the study successful or not?

The next section describes each section in much more detail.

Writing a scientific paper in English

First, a note about word processors. Microsoft (MS) Word is the most popularly used word processor, largely because Microsoft promotes it aggressively. It is very complex and capable, but it could be a lot better. When I wrote this book, I planned to include screenshots from MS Word to illustrate key points, but with each new version, Microsoft changes features, and it is not practical for me to continue to update the screenshots as features are added, removed or changed. Therefore, whether you use MS Word, WordPerfect, LibreOffice, Apple Pages, Google Docs or another program, when I refer in this book to a particular tool I assume that you know how to use it without instruction from me.

Are you ready to start writing? Then let's begin.

Instructions to Authors

Always read the Instructions to Authors

If you have decided on a particular journal, read the journal's **Instructions to Authors**. You must follow these instructions or your paper might be rejected without first being reviewed.

Most journals have their own Instructions to Authors (some short, some extremely long). Some refer you to the *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals* (www.icmje.org/recommendations/). Unfortunately, every set of Instructions to Authors is at least slightly different from every other set.

The Instructions to Authors have 2 important advantages: (1) If you follow them exactly, the journal is less likely to reject your paper before the reviewers see it. (2) They provide a checklist of what you need to do to write the paper (so you don't have to remember).

As you read the instructions, note down all the tasks that you have to do. (Alternatively, cross out the tasks that you do not have to do.)

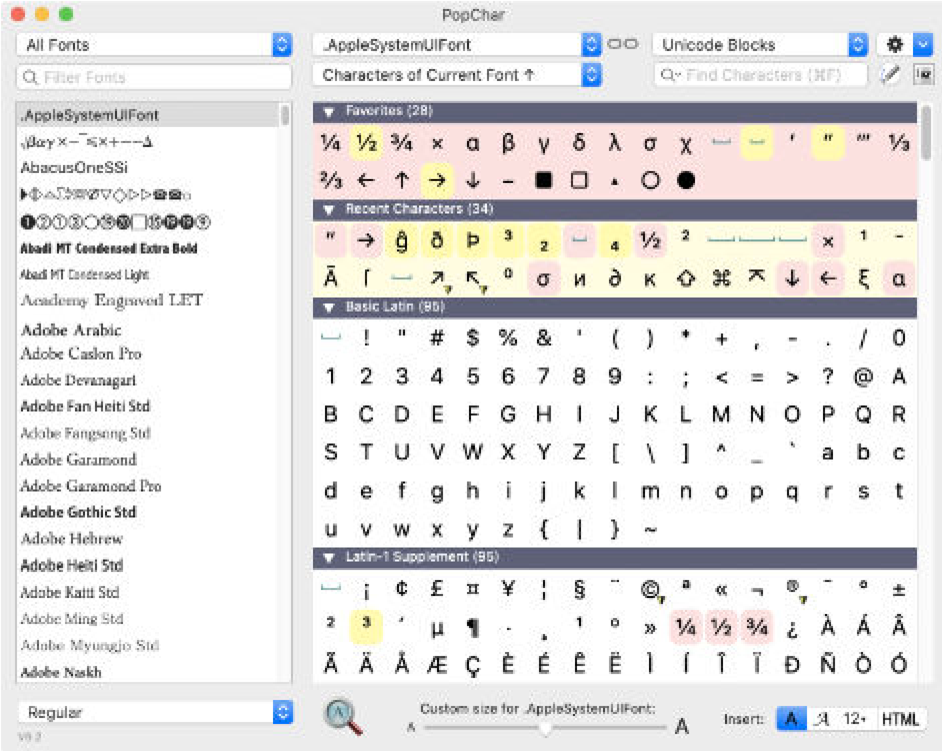
If the journal provides a **sample issue** or sample papers, download those and look at them. This is what your paper will look like when it is published. It is easy to copy the format as you write. An example is better than instructions, because you can copy the same style. In addition, the guidelines published on journal Web sites are often old or incomplete; a published article shows you how the journal actually follows its own guidelines.

If the journal supplies a **template** (usually for MS Word), download it and type in that. Some templates include examples of text to get you started.

Use Western typefaces (fonts):

Times New Roman
Helvetica
Arial
Courier

Please use a standard (Western) typeface (font), such as Times New Roman, Helvetica, Arial or Courier. Use a large enough size—10, 11 or 12 points. A size of 12 points is most common. If you find this too large for easy reading, you can change the magnification of the page on your computer. For special symbols, your word processor should allow you to insert any of hundreds of different characters. Please do not use Japanese typefaces (e.g. ＭＳ 明朝), as many Western computers do not have these installed, and the journal editors might not be able to see them. In addition, some of these symbols may be changed to the wrong symbol by the journal's automatic production software.



An example of a program that allows you to select uncommon characters. This is PopChar (from www.ergonis.com), which is available for both Mac and Windows.

Type of contribution

Is it a research paper, a review paper, a short note, a letter to the editor? The type of contribution will determine the format that you use and the length.

Title

The title of your paper is an important part of the paper. If readers notice it, they will read it and they might then read your paper. If it is boring or unclear, they will not read your paper. So make your title interesting. The best titles give a brief summary of the results.

Make the title interesting

✓ Effect of nitrogen on rice yields (It does not say what the effect is.)

✗ Nitrogen improves rice yields (This tells us what the result is.)

Write the title on a piece of paper and stick it to the wall above your desk. It will remind you of your focus. You can change the title at any time.

Authors and addresses

List all authors and state where they work. Make a clear distinction between personal names and family names; for example, "Akira KUROSAWA" or "Akira <u>Kurosawa</u>". This is particularly important because Western reviewers and journal editors often have trouble recognising Asian family names. If the journal incorrectly records your paper under your personal name, future readers will have trouble finding your papers in a literature search.

Corresponding author

Identify the author for correspondence (for example, with an asterisk: *), and give the *full* mailing address: building number, street number, school or department, university or company, city, prefecture, postal code, country; telephone number with country code (+81 for Japan), fax number with country code; e-mail address. For example:

> *Corresponding author: Akira Kurosawa, Laboratory of Biotechnology, Graduate School of Environmental Studies, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8765, Japan. Tel: +81 (0)3-9876-5432; Fax: +81 (0)3-9876-5431; E-mail: akirak@gses.u-tokyo.ac.jp.

Abstract

A brief summary

The Abstract gives a brief summary of the study. It states the **background**, the **aims**, the **methods** (an overview only), the **results** and an important **conclusion**. It must not include information that is not presented in the main text, and it must include the most important results. A typical maximum length is **250 words** (sometimes fewer, sometimes more).

Key words

These are used in computer database searches. If you want readers to find your paper, then supply a list of key words related to the study. The key words list is intended to supplement the words in the title, so you can use different words from those in the title (although some journals ask you to use the same words as in the title). Some journals require that key words be chosen exclusively from lists published on their Web site. Use specific words related to unique aspects of your study, but not so specific that readers will not use them in a search:

> ✗ Disease, growth, hunger, rice, borborygmus

> ✓ Rice blast, relative growth rate, food sufficiency, *Oryza sativa* L.

Introduction

Background

Problem

Aim

Hypothesis

Introduce the **background** to the work:

- What is the current situation?
- What is known now?
- What is not known?
- What is the problem?
- Why is it a problem?
- Why did you decide to investigate this problem?

Give enough information, but no more than necessary.

Present a brief **literature review**:

- What do other authors say?
- What did they say about the problem?

State the **aims** or **hypothesis** (or hypothes**es**):

- Why did you do the study?
- What is your new idea?
- What did you decide to test?
- What did you expect to find?

Some journals also ask you to summarise what you found in the last paragraph. Others tell you not to do this.

Materials and Methods or Experimental Procedures

Materials used

Equipment used

Where

How

- Describe what you did (as though there were no mistakes or repeats).
- Give enough information that any other scientist can repeat your work exactly.
- State chemical names, quantities, instruments, times, temperatures, concentrations ... .
- If you used a standard method, state the name of the method and cite a reference.
- If you created a new method, state all details: times, temperatures, quantities, manufacturers' names (and places), reagents, purities, illumination, wavelength, pressure—in short, everything that could influence the results.

Manufacturer's details

- Always state the **manufacturer's details**, including city, state or province (if applicable) and country—or the Web site. Any other scientist should be able to order the same materials or equipment.
- If you followed someone else's methods, then say "as previously described (Ito et al. 2010)" or "Following the method of Ito et al. (2010), we ...". Try to cite recent references.
- Cite sources of data that you did not collect (e.g. meteorological Web sites).

Ethics

- If your work involves experiments on humans or animals, then you *must* include an **ethics** statement: state that your experiments were approved by the ethics committee at your institution. If they were not approved, then you cannot publish the results. For an example of detailed guidelines on studies involving human and animal subjects, refer to Nature Publishing Group at https://www.nature.com/nature-portfolio/editorial-policies/ethics-and-biosecurity for advice on "Ethics and biosecurity". Other journals offer similar guidelines.

Statistics

- **Statistical analysis** is a crucial part of most scientific studies. Reviewers and readers must be able to judge for themselves whether you used the correct statistical tests. So you must include the names of the tests and software that you used. For more on statistics, see page 69.

Results

Results = what you found

- Unless the Results and Discussion section are combined, present the results, and nothing else.
- Present summaries of results (such as means), not raw data.
- Present *all* results here. Many authors present new results in the Discussion. This is confusing for readers. Some authors even describe new experiments in the Results. Again this is confusing.
- Present the results only once: in sentences **or** in tables **or** in figures, but not in multiple places. Repetition wastes space and the reader's time.
- Summarise the key results in your tables and figures rather than reporting all of the data.
- Cite the table or figure (or appendix), so that readers can confirm that the result is correct.
- Cross-check text and tables or figures. Errors are easy to make during revisions.
- If the data are not presented (for example, the results of the first experiment showed no effect at the maximum concentration, so you repeated the experiment with a higher concentration), say "(data not shown)".

Present statistics:

e.g. mean ± SD & *P*-value

- Always present the statistics with the results. Present both a measure of the middle (e.g. mean, median, mode) and a measure of the range (e.g. SD, 95%

CI, interquartile range). Indicate whether a result is significant or not significant, and show the probability (*P*) value.

Discussion

What do your results mean?

Hypothesis ✓ or ✗?

Future work

- Explain what the results *mean*, not what they *are*. Do not present new results here.
- What do your results mean? Do they agree with previous authors' results? Do they disagree? Why?
- Has your study advanced knowledge in your field? Where will your results be useful? What questions does your study raise? What should future work look at?

Conclusions

- Finally, in a few sentences, draw your conclusions. The study was or was not successful. Your hypothesis was right or wrong. Something is or is not possible. Were the results what you expected? What should happen next?

Acknowledgements

- Thank everyone who helped: colleagues, lab workers, field workers, other organisations, funding organisations.
- State all funding sources. Include the name of the funding body (e.g. MEXT), the name of the program (e.g. "Rice for the 21st century") and the grant number.

Conflicts of interest

- State any potential conflicts of interest. For example, you own the company that makes the new technology; you received money from a tobacco company for a health study.

References or Literature Cited

All references cited in the text must be listed in this section, and all references listed here must be cited in the text. Remember to include all information that will allow readers to find the same reference. If references are not yet published they must be either "in press" or "accepted". Otherwise, they are not references (that is, they are not yet verified as true). In that case, in the text, say "unpublished" or "personal observation".

Automatic referencing software will make your work easier—see "Automatic Referencing" on page 39.

Whether you use automatic referencing software or you prefer to compile references by hand, you must provide enough information to allow any other person to find the references. Different journals and publishers prefer different formats (which makes life confusing). But most of them require the following information in references:

A journal paper

1. Names(s) of author(s)—who wrote the paper?

2. Year—When was it published?

3. Title—What is the name of the paper?

4. Journal—What is the name of the journal?

5. Volume number—What are the numbers of the journal volume and issue?

6. Page numbers—On what page(s) will you find the article?

7. DOI—the Digital Object Identifier, a unique Internet address that always points to the article.

For example:

> [1] Saiki R, Scharf S, Faloona F, Mullis K, Horn G, Erlich H, Arnheim N. [2] 1985. [3] Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. [4] *Science* [5] **230** (4732): [6] 1350–1354. [7] doi: 10.1126/science.2999980

A book

1. Names(s) of author(s)—who wrote the book?

2. Year—When was it published?

3. Title—What is the name of the book?

4. Publisher—What is the name of the company or organisation that published the book?

5. City—Where is the company or organisation located?

For example:

> [1] Rabinow P. [2] 1996. [3] *Making PCR: A Story of Biotechnology.* [4] University of Chicago Press, [5] Chicago.

A chapter in a book

1. Names(s) of author(s)—who wrote the chapter?

2. Year—When was it published?

3. Chapter title—What is the name of the chapter?

4. Names(s) of author(s)—who compiled the book?

5. Book title—What is the name of the book?

6. Publisher—What is the name of the company or organisation that published the book?

7. City—Where is the company or organisation located?

8. Page numbers—On what page(s) will you find the chapter?

For example:

> [1] Pavlov AR, Pavlova NV, Kozyavkin SA, Slesarev AI. [2] 2006. [3] Thermostable DNA polymerases for a wide spectrum of applications: comparison of a robust hybrid TopoTaq to other enzymes. In: [4] Kieleczawa J. [5] *DNA Sequencing II: Optimizing Preparation and Cleanup.* [6] Jones and Bartlett, [7] Sudbury, MA, USA, [8] pp. 241–257.

Translation from Japanese

In many cases, you will need to cite a paper or book published in Japanese. You will need to provide additional information:

1. Translation of title into English.

2. Language of article or book.

For example:

> Miyata M, Ubukata M. 1994. [1] Genetic variation of allozymes in natural stands of Japanese black pine. *Journal of the Japanese Forest Society* **76**: 445–455 [2] (in Japanese with English summary).

If the published article has a title in English, then use that title (even if the English translation is poor). If the published article has no title in English, then you can supply your own translation (and ask the rewriter or editor to correct it). Remember, however, that if there is no published English title, then readers will not find the article in a search, even if your translation is accurate. In that case, consider providing the title in romaji for Japanese or pinyin for Chinese.

Revision

When you have written your paper, give it to your co-authors for careful review. Do they agree with what you've written under their names?

Next give it to colleagues or friends who are less familiar with your research. Does it make sense to them? If not, then the journal's editors and reviewers also will not understand it.

You will probably have to make some corrections. You might have to go through this cycle a few times: correct, review, correct, review ...

Finally you will be ready to think about submitting it to the journal. Now read the Instructions to Authors again to make sure that you have included everything.

Writing tips

Paragraph structure

State the most important information first. Then present supporting information

When you read a newspaper, you will see that the first paragraph contains the key information. If readers read nothing else in the article, they at least have the key fact. The newspaper has succeeded in telling the story. Similarly, if you read a press release, you will see that the first paragraph contains the key information. Again, if readers read nothing else, they at least know the main fact. The author has succeeded in telling you something.

In contrast, when you read a story, you want to be entertained. You want to enjoy being teased with facts that slowly build up to a climax. (This is not to say that a well written scientific study cannot also be a well written mystery story, in which important information is revealed.)

A scientific paper is like a newspaper, not a story book. So in writing a paper, we must present the fact first, and then support it with more information. This Western style of writing is the opposite of the Japanese style of writing, in which the information builds up to the climax. When you write a scientific paper, you must present the key information first, followed by the supporting information.

This is an example not of Western imperialism but of the time pressure that all scientists feel. There is not enough time in the day to read slowly. Scientists need to know the important information first, and then decide whether they want to know more.

Written information always has a hierarchy: word, sentence, paragraph, section or chapter, whole document. In scientific writing, a sentence tells 1 fact. A paragraph tells 1 idea (composed of 1 or more facts). A section or chapter tells 1 bigger idea (composed of 1 or more smaller ideas), and the whole manuscript assembles those ideas into a coherent story. In this way, we build up a coherent message from coherent ideas made of coherent facts.

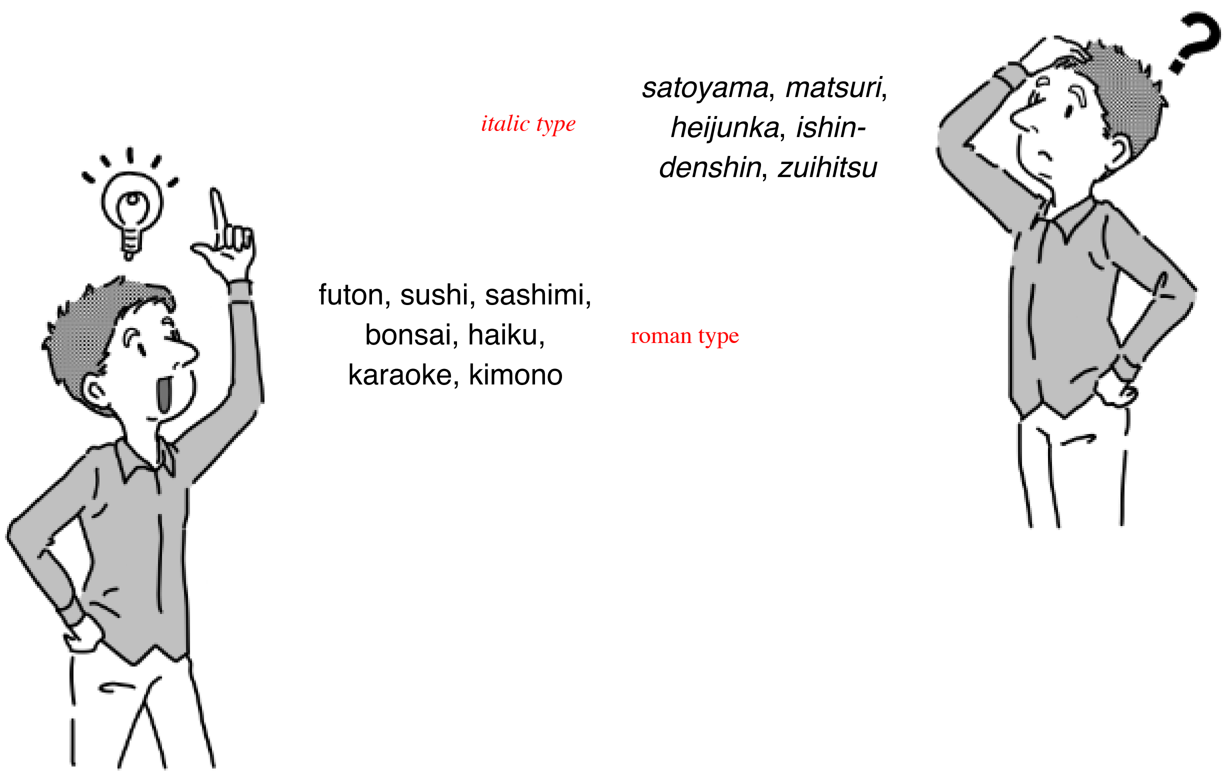
Using Japanese words in English text

Uncommon Japanese words in *italics*

Japanese authors often need to use Japanese words in English-language writing, because the concepts do not exist in English. Examples include *satoyama*, *matsuri*, *heijunka*, *ishin-denshin* and *zuihitsu*. When writing such Japanese words in English texts, write them in *italics*. This is a standard way to show that they are foreign words. There is no need to put them in "quotation marks". Capitalise words only at the start of a sentence or if they are the name of a unique person, place or commercial product (a trademark).

Common Japanese words in roman type

Many Japanese words have been adopted into English and are used as English words. Examples include futon, sushi, sashimi, bonsai, haiku, karaoke, karate, kimono, adzuki, ginkgo, sake, soya, tofu and tsunami. When writing Japanese words that have been adopted in English, write them as normal English words.



Don’t use word-to-word translation

Japanese and English had a last common ancestor more than 15 000 years ago [www.pnas.org/cgi/doi/10.1073/pnas.1218726110]. They have little in common. It is not helpful to think in Japanese and write in English. (This is also true of languages that are more closely related to English. For example, English and German had a last common ancestor only about 2000 years ago, and it is not helpful to think in German and write in English either.) Unless you are fluent in English and can think in English, then you should pay a good translator. (Although translation software such as Google Translate is getting better all the time, it doesn’t understand complex science.)

Quotation marks

English uses 2 forms of quotation marks: ‘single’ and “double”. (Other languages use a variety of other marks, including , ‘ „ “ ‹ › « » 「」.)

Both “ ” and ‘ ’ are correct. (Straight quotes can also be used: " and '.)

If you need to write a quote inside a quote, you can write it like this: “The phrase ‘swidden agriculture’ is often translated as ‘slash and burn’ agriculture.” Or like this: ‘The phrase “swidden agriculture” is often translated as “slash and burn” agriculture.’

“However” and “but”

Many native English speakers think that “however” is a more important way of saying “but”. They are wrong. There are both similarities and differences between “however” and “but”.

“But” is used as a conjunction—a word that joins 2 ideas, in this case to show **contrast**, not inclusion: “We enrolled 20 subjects, but 3 did not finish the trial.”

“However” has 2 very different uses. The first is similar to “but”, and is used to introduce an exception or a difference:

> “We enrolled 20 subjects. However, 3 did not finish the trial.”

Note that whereas “but” is used in the same sentence, “however” starts a new sentence, *and* it has a comma (,) after it: “... , but ...” versus “... . However, ...”.

Variations are allowed in which “however” starts a main clause of a sentence:

> “We enrolled 20 subjects; however, 3 did not finish the trial.” “We enrolled 20 subjects. Three, however, did not finish the trial.”

When used like “but”, “however” always has a full point (.) or a semicolon (;) before it and a comma (,) after it. (Alternatively, it has a comma before and after it when it is placed within the sentence instead of at the beginning.)

A very common error is to place a comma before and after “however” when “however” is used in the role of “but”:

> ✗ The minister asked for an adjournment, however, he was rejected. (The meaning is not clear.)

> ✓ The minister asked for an adjournment, however. He was rejected. (= Instead, the minister asked for an adjournment. He was rejected.)

✓The minister asked for an adjournment. However, he was rejected. (= The minister asked for an adjournment. Unfortunately, he was rejected.)

Notice that the meanings are slightly different.

The other use of "however" indicates not a contrast but a **condition**: "However you look at it, it's not going to be easy." This means "Whichever way you look at it, ..." or "In whatever way you look at it, ...". In this usage, there is no comma after "however". A very common error is to leave out the comma from the *contrast* use and suggest a *condition* use: "However many people go." As written, this looks like a condition, and we expect something to follow: "However many people go, they all come back happy." Insert a comma and the meaning changes: "However, many people go." This sentence is now complete.

In summary, **contrast**: "... , but ..." "... . However, ..."; **condition**: "However it might look ..."

Be clear, correct and complete (not perfect)

In my 25+ years of working with Japanese authors, I have found that some Japanese authors write better English than most native English authors. These Japanese authors make my work easy. On the other hand, some Japanese authors (a small number) write English so badly that I cannot understand anything. Those authors need to use a translator. Most authors, however, lie in the middle. Their English is not perfect, but it is good enough.

When you write in English (or even in Japanese), the language does not have to be perfect. The job of the rewriter is to make the language clear. Your job as the author is to be correct and complete. The translator (if you use one) and the rewriter can rewrite to make it clear. Think of the relationship as teamwork: you have the original ideas, the translator expresses them in English, and the rewriter makes them clear.

Be clear: Yes ✓
No: ⤵
Be correct: Yes ✓
No: ⤵
Be complete: Yes ✓

In summary, try to be **clear**. If you cannot be clear, then be **correct**. And if you cannot be correct, then be **complete**. Your colleagues can help you to change complete into correct. The translator and the rewriter can change correct into clear.

"Significant"

In science, "significant" has a special meaning. It refers specifically to statistical probability. A difference is significant ($P = 0.05$, for example) or it isn't.

In general usage, "significant" does not necessarily have this meaning.

If you want to say that something is significant (= meaningful) in the general sense, then you must use another word; for example, "marked" ("markedly") for an increase, "notable" ("notably") for important, "dramatic" ("dramatically") for very big.

Many Japanese authors write "drastically". This is usually not appropriate, because it suggests something bad or something radical. Words with similar meanings but lacking a negative connotation are "markedly" and "extremely".

Captions or titles

Captions describe figures, tables, maps, photographs and other objects.

A good caption tells the reader both *what the object is* and *what it means*.

Some authors repeat the methods in the caption. Unless the journal's guidelines require this repetition, it is a waste of space. If the methods are already described in the text, then don't repeat them in the caption. (Some journals ask you to put the methods in the caption and not in the text. To minimise the length, focus on the details that the reader needs in order to understand the figure, not all possible details.)

Many authors use the caption to describe the symbols in the figure. This is poor communication. The reader has to read the word, convert it into a symbol in their head, and then find the symbol in the figure.

Some authors place the symbols in the caption and describe them there. This is better, but the reader still has to look from the caption to the figure and back again.

The most effective authors put the symbols in a key in the figure or directly label their symbols and lines.

✗ Responses to temperature. Closed symbols indicate 25 °C; open symbols indicate 30 °C.

(✓) Responses to temperature. ■, 25 °C; □, 30 °C.

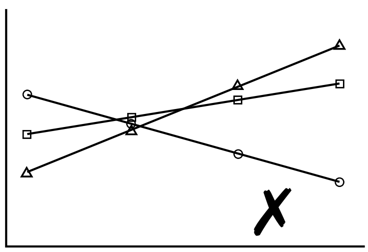


Figure 1. Caption goes here. Circles, 15 °C; squares, 20 °C; triangles, 25 °C.

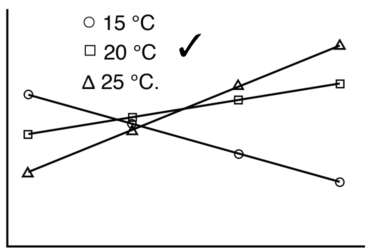


Figure 1. Caption goes here.

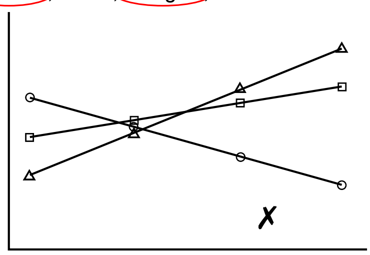


Figure 1. Caption goes here. ○, 15 °C; □, 20 °C; Δ, 25 °C.

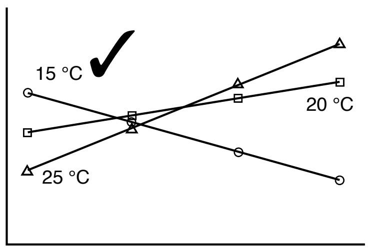


Figure 1. Caption goes here.

Verb tense

Like all European languages but unlike many Asian languages, English verbs use tenses. These are conventionally referred to as past, present and future, although they are more complex than that. In this section I stay with the idea of past, present and future tenses. (For more details on the complexity of English tenses, see http://www.geoff-hart.com/articles/2010/lecture-3.pdf.)

Scientific writing makes a distinction between the researcher's results and the results in the literature. When you report your results, normally you use *past* tense:

Your results: past tense (usually)

> The mean weight <u>was</u> 12.0 g.

This use of the past tense makes a distinction between what you found and what someone else *might* find. The result is true *only* in your study. The mean weight *was* 12.0 g in your study, but it might be different in a future study.

When you discuss results from the literature, you have 2 main options: past tense, particularly for research that was conducted a long time ago, and *present* tense (but try not the mix the 2):

Published results: present tense (usually)

> The mean weight <u>is</u> 12.0 g.

This use of the present tense takes the attitude that because those results are published (they have been through the peer review process), then they remain true unless later invalidated (in science, everything is provisional).

This difference between past and present tenses in scientific writing is a convention, not a law. So you do not have to obey it. In fact, there can be good reasons to present *your* results in the *present* tense and *published* results in the *past* tense (that is, the opposite). For example, you have just measured the atomic mass of element number 117. This characteristic does not change with time, so it can be considered a "universal truth": it is always true.

> Our results show that the atomic mass of element 117 <u>is</u> 294. Iida et al. (2009) <u>suggested</u> that the mass <u>was</u> 293.

Instead, I suggest that you consider the "truth" of any result. It is always true or is it likely to change with each measurement? If it is always true, then use present tense:

Always true: present tense

> Elephants <u>are</u> the largest land mammals.

If it could change, then use the past tense:

> According to Satoh (2005), the mean elephant body mass <u>was</u> 9750 kg.

Abstract—past

Introduction—present

Materials & Methods—past

Results—past

Discussion—past + present

Use present tense for presenting statistics and the results of complex calculations

This distinction applies in every part of your paper. In the Abstract, you use past tense, because it describes what you did. In the Introduction you use present tense to describe what is generally assumed to be true and to report what previous researchers have found. In the Methods you use past tense, because it describes what you did. In the Results, you use past tense to discuss what you found. In the Discussion, again you use past tense to discuss what you found and present tense to discuss what is considered always true: "Table 1 <u>shows</u> ...", because in a hundred years this statement will still be true (it will still show the same data). "The treatment 1 mean <u>is</u> higher than the treatment 2 mean", because every time you look at the 2 values, the treatment 1 mean is always higher; the relationship never changes.

Future tense

I mention future tense briefly here. You use future tense to indicate something that will definitely happen (tomorrow, next year, next century) or that you hope will happen (e.g. your plans for future research). If you are sure that your results will make a difference to the world, then you can say, for example, "Our new method will reduce the costs of genetic sequencing." If you are not sure, then use words that indicate a possibility: "Our new method would reduce the costs ..." or "Our new method might reduce the costs ... ."

In summary, use verb tense to reflect the truth of a statement: true at one time (past tense) or always true (present tense).

Writing and drawing tools

Spelling checker

Even the best writers make spelling mistakes. I keep the spelling checker turned on permanently in MS Word. If my fingers slip, I can see the mistake.

When writing papers, you should always use the spelling checker.

Important: Ensure that the correct language is set for the entire document, and that "do not check spelling" is not set for important parts. A mixture of "UK English", "US English", "German" and "Do not check spelling" will hide errors.

However, the spelling checker's dictionaries do not know most scientific words. So your word processor will tell you that you have made a mistake when, in fact, you are correct. But how do you know whether the word is a correct scientific word or a wrong spelling?

Google will tell you. In particular, **Google Scholar** is focused on published scientific research. Type the word into the box and click on Search. Read the examples. Is the word used as you want to use it? Does it occur many times (perhaps 1000)? Or does Google suggest a different word? If so, is that word used the way you want?

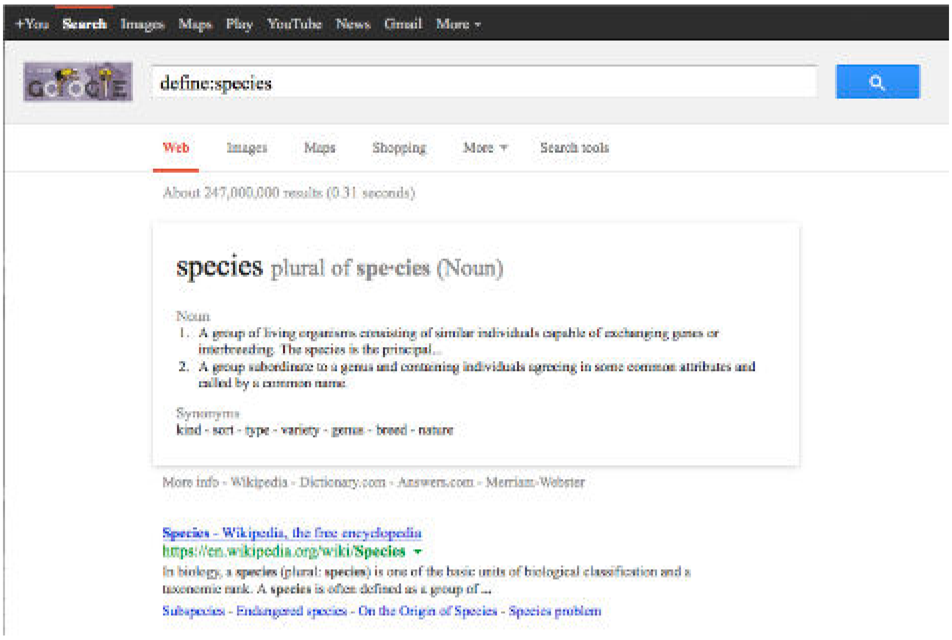
Pay more attention to scientific authors and less attention to non-scientific authors. (Remember, though, that even scientists make mistakes.)

In addition, you can ask Google to define a word for you. Type "define xxx", where "xxx" is your word (for example, type `define species`). If you get a definition that agrees with your use, then you are correct. But even if you do not get a definition, you might still be correct; some words are very rare but are still correct.

You know more than MS Word

Although spelling checkers are generally reliable, don't automatically accept their suggestions for the "correct" spelling. For example, "imine" is correct, but MS Word will suggest "amine"; "attenuata" is correct, but MS Word will offer to change this to "attenuate". Always consult a dictionary or Google before accepting the suggested word. You know more than your word processor does, especially in your field of expertise.

You can ask Google to define a word for you.



MS Word does not know …

Keep in mind that English spelling dictionaries do not include Japanese family names (in romaji), Japanese place names (in romaji), botanical names, zoological names, bacteria, fungi, genes, manufacturers, rare words and new words.

Download free spelling dictionaries from the Internet

To avoid telling the spelling checker to skip a word every time, you can add it to your user dictionary. Then the spelling checker will accept it every time in the future. You can also download free specialist dictionaries on-line. Or you can create your own list of words and then save them as a specialist dictionary file. See your word processor's Help files for adding dictionaries.

Grammar checker

You might find this useful. But it is wrong more than it is right. I never use it.

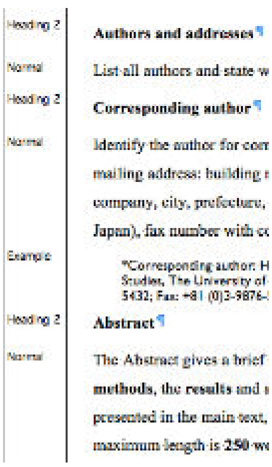
Examples from *Nature* ✓

Examples from Facebook ✗

Alternatively, you can use Google to search for examples of sentences. How do other authors say it? Pay attention to the sources: If the example occurs on Facebook, then forget it. But if it occurs in *Nature* or *Science*, then use it.

Templates and paragraph styles

MS Word and other word processors use document templates (*.dotx or *.dotm in MS Word). These are normal documents with formatting instructions saved in them. Within them you can define styles for paragraphs (e.g. Title, Subtitle, Heading 1, Heading 2, Normal, Caption, References, Table text—font, size, colour, alignment, spacing, language, numbering, tabs), add macros (for doing repetitive tasks within documents) and other things.



When you open a new (blank) document, your word processor will automatically assign the default template. You can replace the default template with a different one. Or you can download specialist templates from publishers.

I always use paragraph styles in everything I work on. This allows me to see whether headings are in the right place, to search for particular paragraphs (e.g. Caption), to reorganise a document in MS Word's Outline view, and to present the document in a way that visually indicates the structure. One big advantage of using paragraph styles is that you can change the appearance of the text by changing the style.

Use the journal's template

Some journals provide document templates for you to download from their Web sites. If you apply a journal's template to your document and then apply the correct style to every paragraph, the paper is more likely to follow the journal's style. Some journals provide templates that contain macros or special toolbars, which make preparing your document easy.

You can also download free templates from Geoff Hart's resources at http://www.geoff-hart.com/resources/journal-manuscript-template.doc.

Drawing tools

Drawing tools in word processor software cause many problems for a journal's production staff, and those problems are likely to affect you late in the review process, when your paper is being prepared for printing. Please *do not* use the drawing tools in your word processor software:

Do not use floating graphics

- Do not use Word Art or Shapes or floating rules (horizontal and vertical lines).
- Do not insert text boxes. In particular, do not put text boxes *on top* of figures—for example, graphs with labels on top of them.
- Do not insert any graphic (such as a figure) that sits *on top* of the text or that requires the text to *wrap around it* (it must sit *in line* with the text).

There are several good reasons not to use floating graphics:

- In MS Word, floating graphics are visible only in Page Layout, not in Draft (Normal) view. If a rewriter or reviewer works on your document in Draft view, they will not see the graphics and might not know that they are present.
- Text boxes are invisible to Find and Replace and to other functions in MS Word (such as word and character counts), so their contents are easily missed.
- Floating graphics tend to move about the page or to move from one page to the next (or even fall off the bottom of the page). In particular, objects such as arrows can become separated from their text, and it is no longer clear where they belong.
- They are not easily corrected.
- Image quality often suffers from being stored in an MS Word file (because the image is converted or compressed).

Graphics cannot be exported reliably from MS Word

- The graphics cannot be reliably exported from an MS Word document. Therefore, the journal must pay someone to redraw or reformat them. This costs you money and allows errors to be introduced into your figures.
- They waste your time. The journal decides where to place them.

Preparing figures correctly

Draw graphics in

Illustrator
Photoshop
GIMP
CorelDRAW
MS Excel

Instead of using your word processor's drawing tools, you should create figures in appropriate drawing software; for example, Illustrator, Photoshop, MS Excel (for graphs), CorelDRAW, GIMP or statistical software such as R. Supply them in a *separate file*: this makes it easier for the rewriter to see the text and the figures side-by-side, which helps them to understand your explanation.

Save graphics in PDF, TIFF, JPEG, PNG, EPS, AI, XLSX

Usually, the best file formats are vector formats such as PDF (.pdf), TIFF (.tif), JPEG (.jpg, with minimum compression), Excel (.xlsx), Adobe Illustrator (.ai), CorelDRAW (.cdr), PostScript (.ps) and encapsulated PostScript (.eps); and bitmap formats such as portable network graphic (.png) and Photoshop (.psd). Others might also be suitable—ask the journal.

If you expect the rewriter or journal to edit or change the figure, then you must supply it in its *original* format. For example, if you created an illustration in Adobe Illustrator, provide the .ai file, not a JPEG file. But more usually, the rewriter or journal will ask you to make corrections.

Some journals ask you to insert figures at the end of the word processor file. This is OK if the figure is a single graphics file (therefore, do not add labels or arrows on top of it). In this case, insert the figure in-line (not floating), so that it is visible all the time. You should be able to right-click on the figure and change its placement. The figure will then behave as a (very large) word in the line of text.

However, you must send the *final* images separately, in either their native file format or a standard file format.

Use Excel for graphs

Microsoft Excel is useful for preparing graphs. Please supply the Excel files separately; don't insert the graph into the text document.

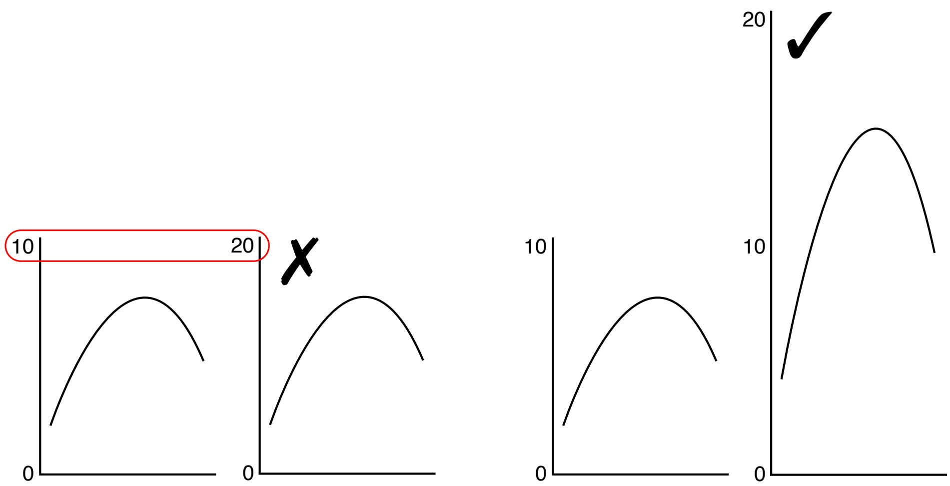
Don't use PowerPoint

Don't use **Microsoft PowerPoint**. PowerPoint makes preparing presentations easy, but the result is difficult to use in a journal. Many journal editors will be unhappy if you supply PowerPoint files.

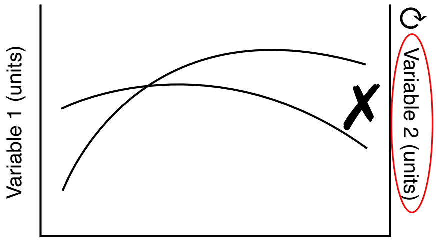
- In summary, supply figures separately (in native file format or a standard file format).
- If you insert them into the word processor file, insert them in-line.
- Don't use floating graphics or add details to a figure using the tools in your word processor.

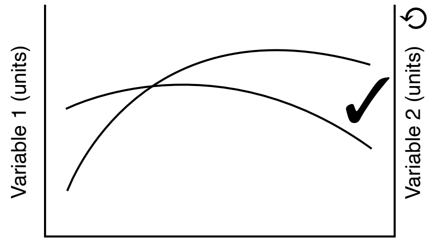
Basic elements of a figure

Use the same vertical scale – it aids comparison

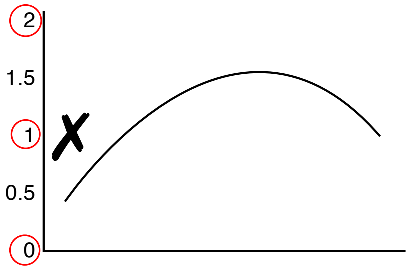


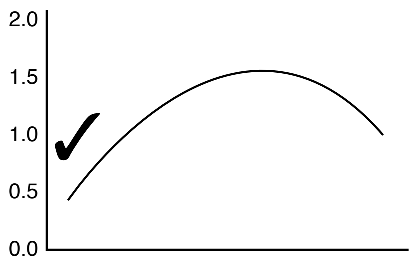
Don’t rotate *y*-axis label clockwise



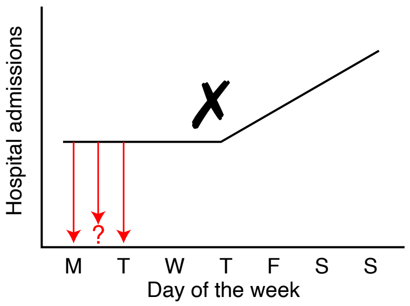


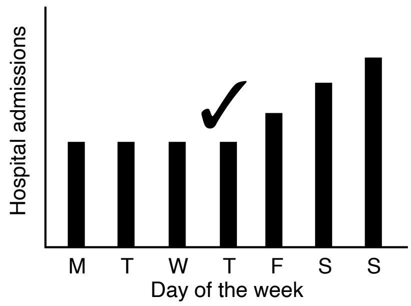
Use the same number of significant figures on all values on axes. For example, “0.5, 1.0, 1.5, 2.0”





Don’t join discrete vari-ables





Use a key! A key is much easier to understand than words that describe symbols

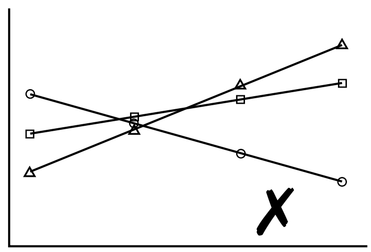


Figure 1. Caption goes here. Circles, 15 °C; squares, 20 °C; triangles, 25 °C.

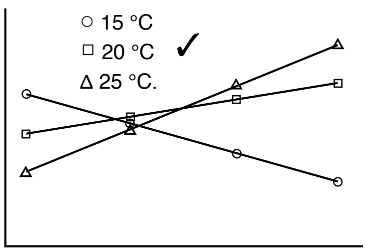


Figure 1. Caption goes here.

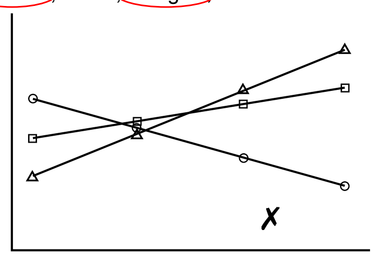


Figure 1. Caption goes here. ○, 15 °C; □, 20 °C; Δ, 25 °C.

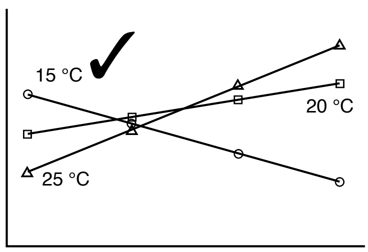
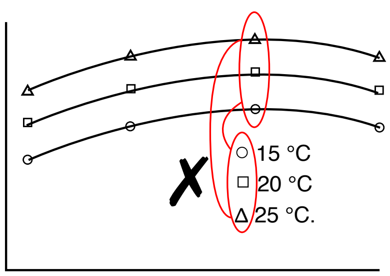
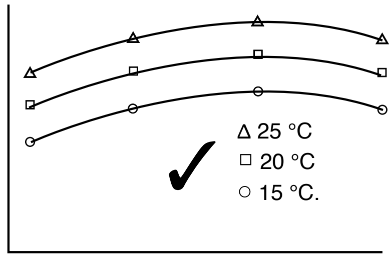


Figure 1. Caption goes here.

Place the elements of the key in the same physical order as used in the figure





Tables

Tables are essential in scientific writing.

Use your word processor's Table function or a spreadsheet

Always use your word processor's table function to create tables, or supply them in a separate spreadsheet file. This means that columns can be aligned, rows can have spacing adjusted, cells will expand to accommodate text, formatting can be applied, and numbers can be summed in a total. It also means (and this is crucial) that the journal can easily adjust your tables to fit within the available space.

Don't use images

Never insert *images* of tables (for example, a .png file), because images cannot be corrected. If you do not supply an editable table (in which corrections can be made), then the rewriter will not be able to correct any mistakes, and the journal editor will not be able to correct mistakes before the journal prints it.

Never add extra spaces

Some authors create tables by inserting spaces between numbers. This creates a mess for the journal. The biggest problem is that alignments change as text is inserted or deleted. The result is that numbers can end up in the wrong place.

Tabs are suitable for small tables

Some authors use tabs to create tables. This can work successfully for very small tables, but it still does not allow cells to get bigger as more text is inserted.

Don't add floating lines

Some authors draw rules (horizontal and vertical lines) on top of the text. As discussed above under "Drawing tools", these lines do not always move with the text, and they are not visible in Draft view.

Always use your word processor's Table function

The word processor's table function (or the spreadsheet software) will do everything you need. You can easily change the width of the whole table, the width of any column, the alignment in any column (left, centre, right, decimal); add or change rules; add background colours or shading; create repeating headers for very long tables; move columns or rows to new positions, insert new ones or delete them; and sort data by numerical or alphabetical order. In addition, the size of cells will automatically increase or decrease as you add or delete data. For more advice, see http://www.geoff-hart.com/articles/2013/word-tables.html.

In summary, always use the table function. Never use spaces to separate columns.

Left: An editable table. Right: An uneditable image of the same table.

Table 3 Summary of SSRs in pear cDNA sequences.

Type of motifs	No. of ESTs	
Dinucleotide	2,027	
AC/CA		76
GT/TG		77
AG/GA		702
CT/TC		982
AT/TA		189
CG/GC		1
Trinucleotide	1,012	
Tetranucleotide	55	
Pentanucleotide	11	
Hexanucleotide	8	
Total	3,113	

Table 3 Summary of SSRs in pear cDNA sequences.

Type of motifs	No. of ESTs	
Dinucleotide	2,027	
AC/CA		76
GT/TG		77
AG/GA		702
CT/TC		982
AT/TA		189
CG/GC		1
Trinucleotide	1,012	
Tetranucleotide	55	
Pentanucleotide	11	
Hexanucleotide	8	
Total	3,113	

A table full of spaces.

Table 2. Root elongation types of seedlings of two *Quercus* species (*Q. variabilis* and *Q. serrata* subsp. *mongolicoides*) at the site of bare granite area.

	Oblique	intermediate	Perpendicular	Total
Q. variabilis	1	1	15	17
Q. serrata subsp. *mongolicoides*	13	3	1	17

The same table after extra spaces are removed.

Table 2. Root elongation types of seedlings of two *Quercus* species (*Q. variabilis* and *Q. serrata* subsp. *mongolicoides*) at the site of bare granite area.

Oblique intermediate Perpendicular Total

Q. variabilis 1 1 15 17

Q. serrata subsp. *mongolicoides* 13 3 1 17

A table created with multiple tab characters and no tab stops.

Table 1. Annual resource production by host plant in two years.

	Year	
Type of shoots	2011 (*N* = 26)	2012 (*N* = 25)
Primary shoots		
Total number of shoots	1782	4295 (2163 in June)
Per-tree mean ± SD	68.5 ± 59.3	171.8 ± 112.1 (86.5 ± 66.1)
min–max	0–195	16–467 (4–328)
Secondary shoots		
Total number of shoots	787	358
Per-tree mean ± SD	30.3 ± 29.7	14.3 ± 25.1
min–max	0–93	0–114

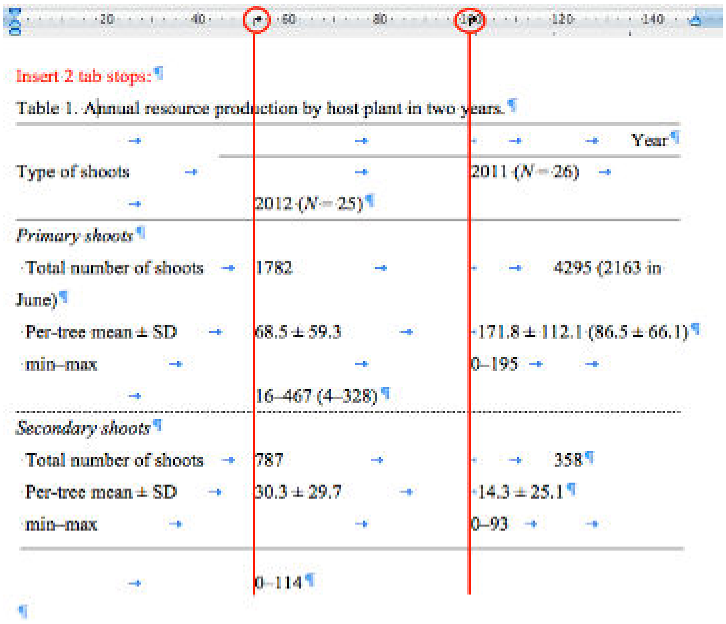
The same table after extra text is added.

Table 1. Annual resource production by host plant in two years.

	Year	
Type of shoots	2011 (*N* = 26)	2012 (*N* = 25)
Primary shoots		
Total number of shoots	1782 (insert some text)	4295
(2163 in June)		
Per-tree mean ± SD	68.5 ± 59.3	171.8 ± 112.1 (86.5 ± 66.1)
min–max	0–195	16–467 (4–328)
Secondary shoots		
Total number of shoots	787	358
Per-tree mean ± SD	30.3 ± 29.7	14.3 ± 25.1
min–max	0–93	0–114

Tab stops →

The same table as on the previous page when defined tab stops are set.



Insert 2 tab stops:

Table 1. Annual resource production by host plant in two years.

Year

Type of shoots 2011 (N = 26)
2012 (N = 25)

Primary shoots
Total number of shoots 1782 4295 (2163 in
June)
Per-tree mean ± SD 68.5 ± 59.3 171.8 ± 112.1 (86.5 ± 66.1)
min–max 0–195
16–467 (4–328)

Secondary shoots
Total number of shoots 787 358
Per-tree mean ± SD 30.3 ± 29.7 14.3 ± 25.1
min–max 0–93
0–114

The same table created with the word processor's Table function. This table can now be edited without accidental changes.

Table 1. Annual resource production by host plant in 2011 and 2012.

	Year	
Type of shoots	2011 (*N* = 26)	2012 (*N* = 25)
Primary shoots		
Total number of shoots	1782	4295 (2163 in June)
Per-tree mean ± SD	68.5 ± 59.3	171.8 ± 112.1 (86.5 ± 66.1)
min–max	0–195	16–467 (4–328)
Secondary shoots		
Total number of shoots	787	358
Per-tree mean ± SD	30.3 ± 29.7	14.3 ± 25.1
min–max	0–93	0–114

Document identifiers

When you submit your paper (to the rewriter, the reviewer or the journal), you need to be sure that none of it gets lost. To achieve this, you can add useful information to your document.

Generally at the top of the page, authors add their name and a brief title. (Note, however, that this is not allowed for "blind" reviews, where the reviewers are not told who the authors are.) Use the word processor's "header" function for this. Now every page has the number and your name on it. If the reviewer finds a loose page, they can put it back where it belongs.

Number every page

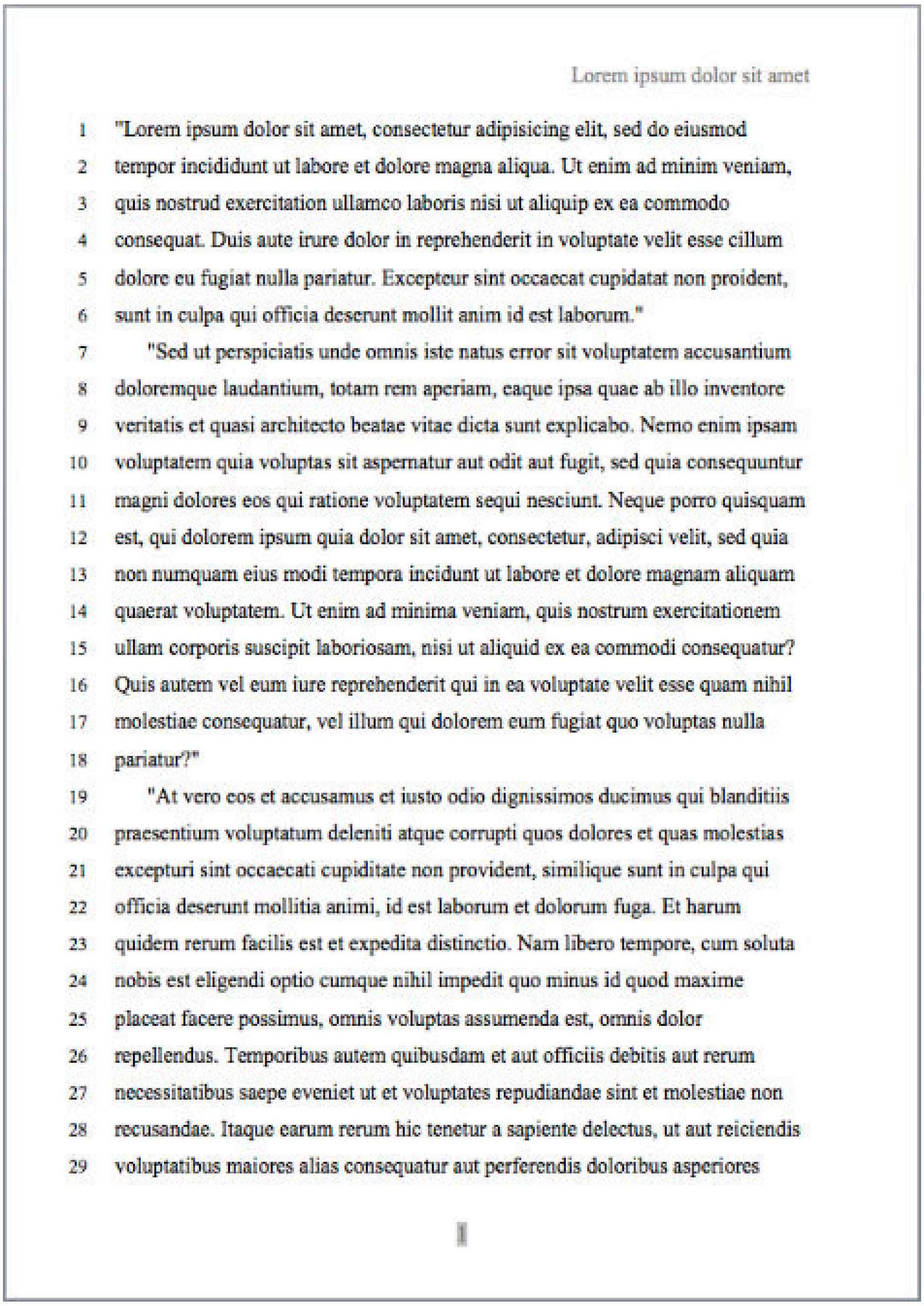
Page numbering is essential.

Add line numbers

Line numbering is useful, especially for reviewers. Use continuous line numbering (not starting at 1 on every page). With that information, you and the journal editors can be sure of discussing the same text, even on opposite sides of the world.

Beware, however, that if you submit your paper for "blind" review (the reviewer must not know who you are), you must delete all of your personal information!

Your name at the top. The lines are numbered. The page is numbered.


Lorem ipsum dolor sit amet

"Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum."

"Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?"

"At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores

1

Automatic referencing

All scientific texts use references. Different journals have different requirements for how references are to be cited in the text and formatted in the References list. If you submit your paper to one journal but it is not accepted, and you then submit it to another journal, you will probably have to reformat the references. Doing this by hand is tedious and can lead to mistakes. In fact, preparing the references in the first place is tedious and can lead to mistakes.

Many journals recommend the use of an automatic referencing system such as EndNote or Reference Manager. There are many referencing programs available, most of which work with word processors. Most can import full reference details from on-line publication databases such as PubMed and ArXiv and can format them automatically according to the required journal style. You can add references yourself too.

As well as automatically formatting your references, these programs look after an important part of referencing: all citations must be referenced at the end and all references must be cited in the text. If you delete a citation, the reference will be automatically removed; and if you add a citation, the reference will be automatically added.

If you import reference details from PubMed, for example, all details will be present and they should be correct. Once you have the correct details on your computer, they will always be correct. And if you take care when adding your own references, then you will never have to type them again. Finally, if you have to reformat a paper for a different journal, the references will be reformatted automatically and correctly.

As a rewriter, I almost always check the citations and the references, and I usually find errors. Errors can occur anywhere, most commonly in authors' names, the year, the title and page numbers. Often information is missing, most commonly the publisher's name or city. Using a referencing system such as EndNote can reduce and even eliminate errors. However, this works only if the details in the database are correct (any errors in the database will reappear each time you insert a citation), and you must use the software correctly; a common error involves typing some citations by hand instead of using the software, or inserting the wrong citation in the software.

For a comprehensive list of referencing software, visit Wikipedia at en.wikipedia.org/wiki/Comparison_of_reference_management_software.

Special characters

Use Western typefaces ("fonts")

Most Western journals should be able to cope with characters from Japanese typefaces (fonts) in documents. However, it is not safe to assume that they will have the correct Japanese fonts installed, and so Japanese characters might not appear correctly.

By "characters from Japanese fonts" I don't mean hiragana, katakana and kanji. I mean special symbols such as mathematical symbols (×, ≥, ÷), bullets (•, ○), stars (* ✩) and shapes (□, ◆).

If you want to use a character that is present in the Western character set (such as μ, ≥ or ±), please insert it using a Western font (such as Times New Roman). Your word processor will allow you to insert many less common symbols too.

Don't use Japanese typefaces ("fonts")

The Japanese character set is much larger than the Western character set, and it includes characters for symbols such as °C, mm^2 and Hz, that Western writers insert using standard characters. Please don't use these special Japanese characters—you create a risk that the journal editors and reviewers will not see them correctly and that that the journal's production software will change them to the journal's standard font and print the wrong character. In addition, Japanese characters are spaced differently from Western characters, and they distort character and line spacing.

In summary, where possible, use characters in Western fonts, not Japanese fonts.

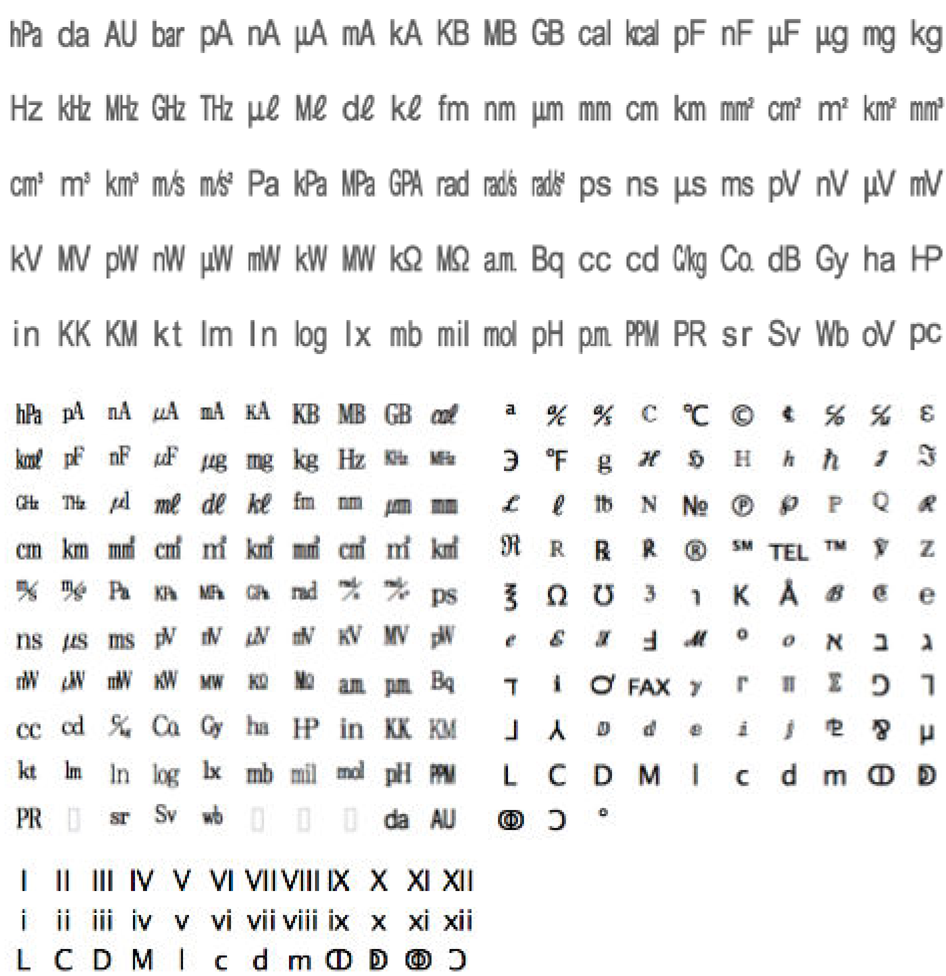
Paragraph numbering and bullets

Word processors offer a range of automatic paragraph numbering and bulleted list (•) styles. You can use these or insert your own bullets (° — § ※ etc.).

Use these symbols. ✓

°C © ® ™	m/s Pa kPa
μg mg kg	Bq dB Gy ha
Hz μL mL	ln log mol pH
nm μm mm cm km	I II III IV V VI VII
mm^2 cm^2 m^2	VIII IX X L C D M

Not these symbols. ✗



Common errors in English by Japanese authors

In 25+ years of working with Japanese authors, I have found several common errors that Japanese authors make in English. I am sure that they all reflect differences between Japanese and English. They are described in this section.

"Individual 3 plants" (adjective order)

In English, numbers come before adjectives:

number adjective noun

✗	✓
individual 3 plants	3 individual plants
pear 98 SSR motifs	98 pear SSR motifs
the enrolled 12 patients	the 12 enrolled patients

"More than 2"

">2" ≠ "≥2"

Many Japanese authors use this to mean "2 or more". But "more than 2" specifically excludes 2. You can see the difference in the mathematical notation: ">2" is not the same as "≥2". "More than 2" means "3, 4, 5 ...". "2 or more" means "2, 3, 4, 5, ...".

"As for xxx"

Wikipedia says "Thus Japanese, like Chinese, Korean, and many other Asian languages, is often called a topic-prominent language, which means it has a strong tendency to indicate the topic separately from the subject, and the 2 do not always coincide. The sentence *Zō-wa hana-ga nagai* (象は鼻が長い) literally means, 'As for elephants, their noses [are] long'. The topic is *zō* 'elephant', and the subject is *hana* 'nose'."

This description is true of English as well. The example sentence ('As for elephants, their noses [are] long') is correct in English as well as Japanese. However, scientific writing needs to be clear, and this sentence structure can be made shorter and clearer.

Instead, write "An elephant's nose is long." That is, keep the sentence as short as you can make it while still including all important information.

✗ <u>As for</u> late-sown rice, the harvest date was delayed by only 20 days.

✓ The harvest date of late-sown rice was delayed by only 20 days.

✗ <u>Regarding</u> tree regeneration, the proportion of tree species gradually increased ...

✓ The proportion of tree species gradually increased ...

L and R

The English sound /l/ does not occur in Japanese. (Conversely, the Japanese /r/ sound [ɾ] does not occur in English.) Because the English letter L/l does not have a corresponding sound in Japanese, Japanese authors often have difficulty with words that contain this letter. Because the Japanese pronunciation of the letter R/r lies somewhere between the English pronunciations of R and L, Japanese authors often replace an L with an R. This is the reason for the common confusion among Japanese people between English words such as lice/rice, light/right, low/row, allow/arrow.

This confusion continues into writing.

Always keep the spelling checker switched on

The spelling checker can help. Keep it always switched on. That way, if, for example, you type "crear", it will correct it to "clear". *However*, the spelling checker is not always right! It will not change a wrong (but correctly spelled) word to the right word. For example, it will not change "dairy" to "daily". Therefore, you need to be careful with L/R words. The following list of word pairs (and trios) is likely to include some terms that are relevant to your field of study. It includes only word pairs with a 1:1 substitution between l and r, and therefore does not include word pairs with different spellings, such as climb/crime.

Speller exclusion dictionary

If you find word pairs that you know cause problems for you (either in this list or in your work), it is a good idea to add them to the word processor's "speller exclusion" dictionary as words to mark as wrong, even if they are right. The most commonly cited example is pubic/public.

allay/array	lagging/ragging	locker/rocker
allow/arrow	laid/raid	look/rook
appeal/appear	lake/rake	loom/room
belly/berry	lamp/ramp	lope/rope
bland/brand	lank/rank	lot/rot
blew/brew	lash/rash	lout/rout
blight/bright	late/rate	low/row
blow/brow	lather/rather	loyal/royal
blown/brown	law/raw	lubber/rubber
blush/brush	lay/ray	luck/ruck
clash/crash	laze/raze	lug/rug
class/crass	lead/read	lump/rump
clime/crime	leader/reader	lung/rung
clown/crown	leap/reap	lure/rule
collect/correct	led/red	lush/rush
daily/dairy	leach/reach	lust/rust
fail/fair	leek/reek	molal/molar/moral
fallow/farrow	lib/rib	molality/molarity/morality
flair/frail	lice/rice	mole/more
flanking/franking	lick/rick	older/order
flee/free	lid/rid	play/pray
flight/fright	light/right	solely/sorely
fly/fry	link/rink	splint/sprint
glass/grass	lip/rip	tale/tare
gley/grey	load/road	temporary/temporarily/temporally
glow/grow	loam/roam	tile/tire
lace/race	lob/rob	umbel/umber
lack/rack	lobe/robe	wile/wire
lag/rag	lock/rock	

Repetition

Many authors say the same thing several times in scientific papers. This is not necessary, and it can draw attention away from the important message. If we are speaking, we often say the same thing several times to ensure that our audience hears the important point. Writing is different. In writing, the meaning is always present. The reader can re-read the same part many times if he or she wants. The writer must place the important information where it can be seen. This requires thinking, planning and maybe rewriting.

Scientists have to read a lot of information, and therefore take shortcuts to reduce the amount they must read. It is easy to miss the important parts. If a writer says the same thing 3 times, the hurried reader might think "Yes, you said that. Tell me something new," and jump to another part. If the text looks the same as before, the reader might miss some new information.

Say it once

In scientific writing, say it once—or not more than once per section. The fact will always be present. If it is important to remind the reader about the fact, say "As mentioned earlier in this section" or "As we saw in section 1", for example.

Spacing in units

When using numbers and units, *always* use a space between the number and the unit. This is not a convention of English. It is a convention of SI (Système International d'Unités, which is French for "International System of Units"). All science uses SI units or units derived from them. By international agreement, all countries (even the USA, which still uses traditional units) use SI units in most places.

Japanese typography (the presentation of printed text) is already spaced, so a space does not have to be inserted in many places. In English (and all European languages), a space must be inserted: between words, between sentences and, in SI usage, between numbers and units. Many Japanese authors write, for example, "5m", "12dS m^{-1}", "25°C", "30μg", "10L". This is not correct SI usage. By convention (not law or logic), SI units require a space. For example, 5 m, 12 dS m^{-1}, 25 °C, 30 μg, 10 L.

Note that "25 °C" has a space in it. This is because "°C" is a unit (derived, not SI). (But note that some journals require "25°C".)

In contrast, "%" and "‰" never have a space, because they are not units (they indicate a ratio, which has no units). So "100 %" is wrong. Write "100%".

✗ 5m, 12dS m^{-1}, 25°C, 30μg, 10L

✓ 5 m, 12 dS m^{-1}, 25 °C, 30 μg, 10 L

✗ 100 %

✓ 100%

"including"

"Including" introduces a partial list. "Comprising" introduces everything:

including

Oryza barthii, O. brachyantha, O. glaberrima, O. rufipogon, O. sativa, O. schlechteri

comprising

Passive voice, unidentified (dummy) subject and third person

In grammar, we distinguish between "passive voice" and "active voice". Passive voice: "Samples were collected." Active voice: "We collected samples."

Scientific writing has traditionally used the passive voice. In passive voice, the person or thing that performs the action (the *actor*) is not identified, or is identified later in the sentence; in active voice, the actor is identified by direct association with the action. The rationale has always been that the passive voice makes science objective, and that the active voice is egocentric.

Neither statement is correct. Science is subjective. The scientist is central to the work. The scientist is part of the science. Therefore the person who did the work must be acknowledged.

Using the passive voice can hide important information. If an author writes "A lot of thought has been given to the proposal ...", it is not clear who has given that thought. The authors? The government? Other authors? If the author changes this to "We have given a lot of thought to the proposal ..." or "The Standing Committee has given a lot of thought to the proposal ...", we now know who gave the thought, and we can assess for ourselves the importance of the fact.

A related usage involves a "dummy" subject: a subject that stands in for the real subject, while the real subject is lost. For example, if you read "It is thought that ...", you must ask "*Who* thinks this?" An important part of thinking critically is to understand who says what so you can assess the credibility of the source. "It is thought that ..." hides an important fact. "We think that ...", "The Standing Committee thinks that ..." and "Suzuki et al. think that ..." all give different and important information.

First person: I, me, we, us

Second person: you, your

Third person: he, she, they

English grammar also distinguishes between first person (I, me, we, us etc.), second person (you, your etc.) and third person (he, she, it, they etc.). You will often see instructions to authors that say "Write in the first person" or "Write in the third person". In some cases they actually mean "Write in the active voice" or "Write in the passive voice." To see the difference, compare "The sample was weighed out" (passive voice) and "The technician weighed out the sample" (active voice, third person). The second sentence is written in the third person, but this is not what the journal wants.

This example raises an important point: Although the active voice and first person can be important ("We interviewed the subjects"), sometimes the actor is not important. So "The sample was weighed" is correct, even if "we weighed the sample" might be clearer.

Pronouns

To avoid saying the same words several times, we can use pronouns: he, she, it, they etc. This can save space and improve readability.

> ✗ The *Hd1* gene was modified to incorporate the CaMV promoter upstream of <u>the *Hd1* gene</u>.

> ✓ The *Hd1* gene was modified to incorporate the CaMV promoter upstream of <u>it</u>.

But be aware of creating uncertainty with pronouns:

✗ The previous study used the original protocol but <u>it</u> was unsatisfactory.

The previous study was unsatisfactory or the original protocol was unsatisfactory?

Nouns as adjectives

In English, nouns can be used as adjectives. English (as do other Germanic languages) allows us to join nouns into long adjectival strings; for example:

measurements

topography measurements

sub-canopy topography measurements

lidar sub-canopy topography measurements

canopy lidar sub-canopy topography measurements

vegetation canopy lidar sub-canopy topography measurements

This can be very convenient, but it can also make reading difficult or even impossible:

Don't do this!

Space telescope wide-field planetary camera instrument definition team ground based charged-couple-device camera

This flexibility also leads many Japanese authors (and Western authors too) to try to force the words to do 2 different jobs. For example:

The <u>colonization</u> <u>ability</u> of the host intestinal tract

In this example, the author has tried to force "colonization" to play 2 different roles: "*colonization* ability" and "*colonization* of the host intestinal tract". We need to separate these 2 ideas, because the sentence, as written, says that the intestinal tract can colonize something:

✗ The colonization <u>ability of</u> the host <u>intestinal tract</u>

If we focus on the most important words in this sentence (underlined), we can see that it refers to the "ability of the intestinal tract"; that is, it says that the intestinal tract has an ability to colonize something. In fact, the intended meaning was the ability of *bacteria* to colonize the intestinal tract:

✓ The ability of [the bacteria] to colonize the host intestinal tract

✗ Development of risk assessment <u>methods of greening disease</u> ("development of *risk assessment* methods" + "*risk assessment* of greening disease")

This sentence literally refers to "methods of greening disease"; that is, it tells us that the disease uses risk assessment methods. The intended meaning was methods of *risk assessment*:

✓ Development of methods to assess the risk of greening disease

✗ A damage risk <u>map by wild boars</u> ("*damage* risk map" + "*damage* by wild boars")

This sentence tells us that wild boars use maps.

(✓) A wild boar damage risk map (Too many nouns in a row)

✓ A map of the risk of damage by wild boars

✗ eradication projects for invasive tree species

This sentence tells us that the tree species work on eradication projects.

✓ projects for the eradication of invasive tree species

"almost" → "most"

✗ Almost of the samples

✓ <u>Most</u> of the samples

✓ <u>Almost all</u> of the samples

"on the contrary" → "in contrast"

Usually, when Japanese authors write "On the contrary", they mean "In contrast". There is a difference:

Most plants in treatment 1 died. <u>In contrast</u>, most plants in treatment 2 survived.

Here, "in contrast" shows a simple contrast (fact 1 versus fact 2).

"On the contrary" is usually used to contradict another statement. For example, "I thought you were dead!" "On the contrary. As you can see, I am alive and well." It is rarely used in science. If you want to write "On the contrary", you probably need "In contrast."

"any" → "no"

Many Japanese authors write

✗ Any patients survived for more than 12 months

when they mean

✓ No patients survived for more than 12 months."

Similarly:

✗ Almost any patients ..."

✓ Almost no patients ..."

✗ Any splicing variants does not affect ...

✓ No splicing variants affect ...

"was shown" → "is shown"

Use present tense for presenting results

In referring to figures and tables, most Japanese authors say "The plot <u>was</u> shown in Figure 1" or "The results <u>were</u> shown in Table 1". Correct English usage is "is" or "are": "The plot <u>is</u> shown in Figure 1"; "The results <u>are</u> shown in Table 1". The correct usage requires present tense, on the basis that the fact remains true forever. The present tense shows what is true, not an event that happened in the past.

Common punctuation problems

Punctuation can have dramatic effects on meaning. Compare "Woman, without her man, is a savage" with "Woman: without her, man is a savage." The meanings are opposite. Why? In the first sentence, "without her man" is parenthetical; it is not crucial information and so can be deleted: "Woman is a savage." In the second sentence, "Woman:" is an introductory phrase; without that introduction, we have "Without her, man is a savage."

Hierarchy:

.

:

;

,

In English, the semicolon (";") indicates a brief pause, longer than a comma (",") but shorter than a full point ("."). It can also join 2 or more clauses of equal weight that together form a complete thought. In contrast, the colon (":") is preceded by context and introduces a list or a series of subordinate clauses that are meaningful in that context. The hierarchy in English is "." > ":" > ";" > ",".

Many authors (including English authors) use ";" where they should use ":".

✗ A; Aichi, I; Ishikawa, K; Kagoshima, M; Mie, S; Shizuoka, T; Toyama.

✓ A, Aichi; I, Ishikawa; K, Kagoshima; M, Mie; S, Shizuoka; T, Toyama.

Punctuation marks are summarised in Appendix 3.

Preposition confusion

Prepositions are words that define the relationship between nouns or other elements of a sentence; for example, "the plant was placed in the sample chamber". Prepositions include such words as "in", "on", "through", "by", "with" and "before". English uses many prepositions (about 150), and deciding which one to use can sometimes be tricky. Scientific writers in general (including native English speakers) seem to like "for" in places where other words work better. For example, "For rice, three samples were collected" would be better written as "Three samples of rice were collected."

Most native English speakers make mistakes with prepositions, so it is unfair to expect Japanese authors to learn them all. Instead, I suggest 2 techniques you can use:

1. Consult similar papers on your subject to learn how English authors have written the phrase. For example, search Google Scholar for articles containing "at June" and "in June" and see how English-speaking authors have used them. (The correct phrase is "in June".)

2. Refer to a list of nouns, verbs or adjectives followed by prepositions. You will find a good list at http://www.wordpower.ws/grammar/gramch26.html.

For a list of commonly confused prepositions and how to use them, turn to Appendix 2.

Elegant variation

The guru of English usage was an Englishman called H. W. Fowler. His *Modern English Usage*, first published in 1926, remains a standard reference book for English usage. Fowler created the name "elegant variation" for the practice of using different words to describe the same thing. Scientist authors do this a lot in scientific papers, perhaps because they were taught to do it that way. But this makes it harder, not easier, to understand the meaning. Every time you use a different word, you suggest to readers that the meaning is different also.

For example:

> ✗✗ Treatment 1 gave a mean of 12.5 g, the result in treatment 2 was 14.5 g, and 13.0 g was the average in the third treatment.

This sort of writing can cause confusion. Are the same things being compared, or are they different? When the author wrote “average”, is this the same as “mean” or does it mean “median”?

Parallel sentence structure

Let’s start by making everything parallel:

> (✓) Treatment 1 gave a mean of 12.5 g, treatment 2 gave a mean of 14.5 g, and treatment 3 gave a mean of 13.0 g.

Now that each concept (treatment and value) can be compared easily, we can rearrange and shorten this to:

> ✓ The means were 12.5 g in treatment 1, 14.5 g in treatment 2, and 13.0 g in treatment 3.

Another example:

> ✗ Among the 80 subjects, 26 patients were selected.

Why did the author refer to both “subjects” and “patients”? Do the subjects consist of a mixture of patients and non-patients? Introducing a different word for the same thing can imply a difference that does not exist. Rewrite this as:

> ✓ We selected 26 of 80 subjects.

Changing the subject by omission of words

> ✗ Ten colonies were picked and examined whether they were resistant.

This sentence structure is common among Japanese authors writing in English. The sentence is trying to say 2 things in the same grammatical structure that require different grammatical structures.

> ✓ Ten colonies were picked.

> ✓ Ten colonies were picked and examined.

> ✗ Ten colonies were ... examined whether they were resistant.

The problem is that the colonies were not examined. Instead, their resistance was examined.

> ✓ Ten colonies were picked, and their resistance was examined. (Two events, not one.)

Using active voice and first person can avoid this problem:

> ✓ We picked 10 colonies and examined their resistance.

> ✓ We picked 10 colonies and examined whether or not they were resistant.

“15<” → “>15”

Don’t write “15<”

Some Japanese authors use a non-standard approach to this mathematical notation.

✗ 15<, 150≤

✓ >15, ≥150

This mathematical notation implies the English statement “greater than 15”, not “15 less than”. For clarity, use “<15” if you mean that 15 is the largest value, but “15 units less than *x*” if you are comparing these values with *x*.

“~”

In Japanese, “15~20” means the range of numbers from 15 to 20.

In English, “15~20” has no meaning. Write instead “15–20” or “15 to 20”; the second option is particularly useful when you are presenting a range of negative numbers, because it avoids writing, for example, “–1––2”, which is confusing. This “–” character is not a hyphen (“-”). It is an “en dash” (so called because it is the width of the letter N).

In English, “~” is used to mean “about”, “around” or “approximately”. So “~100” (or “≈100”) means “about one hundred”.

✗ 15~20

✓ 15–20

Punctuation marks are summarised in Appendix 3.

Gender inclusion and singular “they”/“them”/“their”

“Gender” is a grammatical characteristic, and should not be confused with “sex”. English has 4 genders: masculine, feminine, neuter and common.

In English, gender is visible mainly in the pronouns that are used (he, his, him for masculine; she, her, hers for feminine; it, its, they, theirs for neuter and common). Unfortunately, English does not have a specific pronoun that means “either he or she”, or “either his or hers”. We cannot say “Ask the patient to open its mouth,” because “its” is used for non-human subjects. However, “they” and “theirs” have been used in this way for hundreds of years, so we can say “Ask the patient to open their mouth.” Or we can say “Ask the patient to open his or her mouth.”

“-”, “–”, “—”

Hyphen: -

En dash –

Em dash —

The hyphen (“-”) is used to join words into compounds (words that function as a single unit): “high-speed car”, “top-down programming”, “3-month-old baby”. The hyphen key is a standard key on the computer keyboard (to the right of the “0”).

The en dash (“–”) is used to join words into a range or to join 2 subjects of equal weight: “June–August” (a range), “Japan–Korea agreement” (equal weight). It is called an “en” dash because it is the width of the letter N. To type an en dash, on the Macintosh, hold down the **alt/option** key and press the hyphen key; in Windows, hold down the **alt** key, type 0150 on the numeric keypad, and release the **alt** key.

The em dash ("—") is used to make a break in the text and introduce a parenthetical phrase (i.e. words that add to the sentence but are not essential to its meaning): "The population of Japan—127 million—is concentrated in the lowlands". It is called an em dash because it is the width of the letter M. To type an em dash, on the Macintosh, hold down the **alt/option** key and the **shift** key and press the hyphen key; in Windows, hold down the **alt** key, type 0151 on the numeric keypad, and release the **alt** key.

Or in your word processor, look for the "Insert special symbol" option.

Note that it is called a parenthetical phrase because it serves the same function as parentheses: ( ).

Punctuation marks are summarised in Appendix 3.

"such as"

In English, "such as" is used to introduce a few examples, not the full list of possibilities. But some Japanese authors write "such as" and then give the full list:

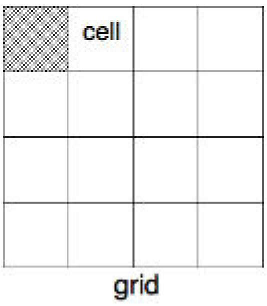
✗ There were 3 genes, such as *Hd1*, *Hd2* and *Hd4*.

Use "such as" only to show examples:

✓ Many countries, such as Japan, the USA and Australia, ...

Where you need to list everything, you can use a colon:

✓ There were 3 genes: *Hd1*, *Hd2* and *Hd4*.



"grid" → "cell"

A grid (also called a "mesh") is the pattern of horizontal and vertical lines, as on a map or in a table. A cell is a square (or other shape) defined by 4 lines or by (x, y) coordinates.

"alphabets" → "letters"

A, b, c etc. are *letters*. The set of {abcdefghijklmnopqrstuvwxyz} is an *alphabet*. This is the Roman alphabet.

Other *alphabets* include {αβγδεζηθιϰλμνξοπϱστυφχψω} (Greek, 24 *letters*) and {абвгдежзиклмнопрстуфхцчшщъыьэюяёй} (Russian, 33 *letters*).

Irrelevant information

Many authors like to include information that is not directly relevant. This extra information might be interesting but it is not critical. Delete it.

Keep to the topic of your study. Keep all the information directly relevant to your study, and delete everything else. There are 2 reasons for this approach. First, scientists do not have enough time to read everything. If you make them read irrelevant information, they might lose patience and read something else. If so, you have not conveyed your important message. Second, irrelevant information dilutes your message. Your message can seem less important if you mix it in with other facts.

Articles—“a”/“an”, “the”, “some”

	def.	indef.
singul.	the	a, an
plural	the	some

Like all European languages but unlike most Asian languages, English uses articles with nouns. The role of an article is to define whether you mean a specific thing (the *definite* article) or any member of a group of things (the *indefinite* article). English uses 2 types of article: definite (“the”) and indefinite (“a”). Each article comes in either singular or plural form. The plural of “the” is also “the”. The plural of “a” is “some” (there is no logic to this). In addition, “a” becomes “an” in front of most (but not all!) vowels. English has 4 common articles—“the”, “a”, “an” and “some”—along with demonstrative articles, such as “this”, “that” and “those”.

As well as using the definite and indefinite articles, English also leaves out the article in many places. This gives us 3 usages (definite, indefinite, none). How do we choose the right one?

“the” apple (that one there)

“an” apple (any apple is OK)

“apples” (as a concept)

We use the **definite** article (“the”) when we refer to a specific thing: “I like the apples in your kitchen” (specifically those apples, not the apples in your office).

We use the **indefinite** article (“a”, “an”, “some”) when it doesn’t matter what the thing is. It is not specific: “Please pass me an apple” (any apple, from either your kitchen or your office; it doesn’t matter).

We use **no** article when we are talking generally about something: “I like apples” (any apples, any type, any number).

Here are some examples:

The study [our study described here] was conducted in the phytotron [there is only 1 phytotron here] of the University of Tokyo.

The studies [our studies described here] were conducted in the phytotrons [we have already mentioned these, and we used all of them] of the University of Tokyo.

A study [1 study, not specified] was conducted in a phytotron [1 of several phytotrons] of the University of Tokyo.

Some studies [an unspecified number, but not all of them] were conducted in some phytotrons [there are several phytotrons, but we did not use all of them] of the University of Tokyo.

Studies [maybe some of them, maybe all of them] were conducted in phytotrons [there are several phytotrons, but we don’t say how many or which ones] of the University of Tokyo.

The insect [there is only 1, which we have already described] was collected in the field [there is only 1 field] of Tsukuba University.

The insects [we have already mentioned these specific individuals] were collected in the fields [there are several fields, which we have already mentioned] of Tsukuba University.

An insect [only 1 insect, but it doesn’t matter which one] was collected in a field [there are several fields, but we don’t say which one] of Tsukuba University.

Some insects [the number is not important, but not all of them] were collected in some fields [but not in other fields] of Tsukuba University.

Insects [the number is not important] were collected in fields [more than 1, but the actual number is not important] of Tsukuba University.

We analysed 9 strains of bacteria. [There were 50 strains, but we selected a few examples.]

We analysed the 9 strains of bacteria. [We listed these same 9 strains above.]

In 2 examples, ... [This is the first time we refer to these examples.]

In the 2 examples, ... [The same 2 examples were mentioned before.]

Put this in the box. [That box over there.]

Put this in a box. [Any box, it doesn't matter.]

Put this in the boxes. [Those boxes over there.]

Put this in boxes. [Use as many boxes as you need.]

When deciding whether to use "a" or "an", remember that the choice depends on the *sound* of the word, not on whether the first letter is a vowel or not. So "a book", "an apple", but "a uniform", "an honour".

Singular and plural

Like all European languages, but unlike most Asian languages, English distinguishes between singular (1 thing) and plural (2 or more things). Fortunately, it is not as complex as it could be: some Australian Aboriginal languages distinguish between singular (1), dual (2), trial (3) and plural (4 or more).

When a noun in English changes between singular and plural, so too do the article (a ↔ some), demonstrative article (this ↔ these, that ↔ those), pronoun (I ↔ we, it ↔ they) and verb ("he is" ↔ "they are", "plant grows" ↔ "plants grow"). I often see sentences such as "The genes was found on chromosome 4", and I have to ask the author, "Do you mean 'The genes were found' or 'The gene was found'?" This is probably the most common error of number (singular/plural) that Japanese authors make. The difference can be important to the meaning, so please take some extra time to check that the nouns and verbs agree with each other.

man/men
woman/women
foot/feet
goose/geese
tooth/teeth
mouse/mice
louse/lice
dormouse/dormice
ox/oxen
child/children

English is a Germanic language, and it still uses Germanic plurals for 12 words: man/men, woman/women, foot/feet, goose/geese, tooth/teeth, mouse/mice, louse/lice, dormouse/dormice, ox/oxen, child/children (actually a double plural), brother/brethren (usually "brothers") and cow/kine (now used only to surprise people; always "cows").

Beginning with the Norman (French) invasion of England in 1066, French rules began to be applied. Today, we add "-s" or "-es" to most words to make them plural: book/books, valley/valleys, calf/calves, table/tables.

Latin plurals

Greek plurals

In addition, English has many words of foreign origin, and many use their native plural forms. Examples include corpus/corpora, datum/data, genus/genera, focus/foci, index/indices, locus/loci, matrix/matrices (Latin); hypothesis/hypotheses, criterion/criteria, phenomenon/phenomena, stoma/stomata (Greek); plateau/plateaux, madame/mesdames (French); and many others.

Singular = plural

Further, some English words are the **same in both singular and plural**, including sheep, swine, fish, trout, cod (all of these are animals, and many other animal names follow the same rule).

No plurals

Worse, many English words have **no plural** spelling. Most of these words apply to things that cannot (or need not) be counted; for example, poultry, people, cattle, milk, wisdom.

No singular!

Finally, some English words have **no singular** spelling, including thanks (are), scissors (are), annals (are), news (is), means (is/are).

This is all very confusing.

If you are not sure, add "s"

Fortunately, your rewriter will always know the correct plural (or singular). My best advice to you is to remember the first 10 Germanic plurals listed above and add "-s" (or "-es") to everything else if you are not sure. (If you are sure, that's good.) The rewriter will correct any mistakes.

For a list of the most common irregular nouns used in scientific English, see Appendix 1.

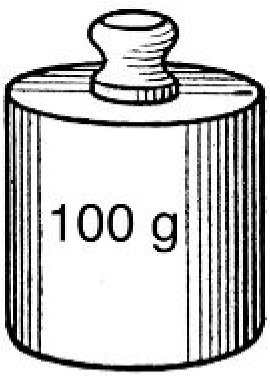
Count with integers, measure with real numbers

Integer: 0, 1, 2, 3, …

Real: 0.95, 1.0, π, 9.8

A **count noun** is something that is counted; for example, apples, people, cars, languages, ideas, cells, coins, atoms. Such things are counted with **integers** (0, 1, 2, 3, ...). A **mass noun** is something that is not counted but measured; for example, distances, times, volumes, masses, volts. Such things are measured in **real numbers** (0.5 m, 1.21 s, 12.7 L etc.).

Real numbers are typically used in the **plural** (because this is natural language) but describe a single scientific concept, so they can *also* be written in the **singular**:



one hundred grams *was* (that is, a single quantity massing 100 g)

after 30 s *elapses* (that is, a single period of 30 s)

twenty-four hours *isn't* enough in the day

9000 kJ *is* enough for anyone

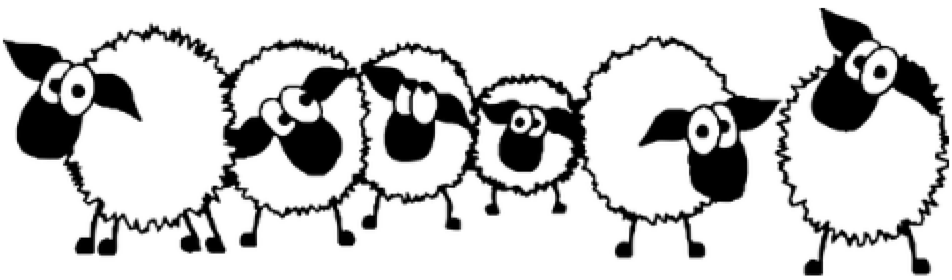
but integers are used in the **plural**:

20 sheep *were* treated

there *were* 2×10^9 cells

Six species *are* endemic

Note, however, that in *speech*, the actual units for real numbers are plural: "0.1 gram<u>s</u>", "1.0 metre<u>s</u>", "2.0 amp<u>s</u>".



Notional accord (using singular or plural depending on the intended meaning) describes the agreement we choose when faced with a choice of singular or plural. Do we make the verb agree with the members of the set or with the set itself? For example, "the F_1 generation were" or "the F_1 generation was"? Notional accord describes whether we are focusing on the members ("the F_1 generation were") or the set of members as a discrete unit ("the F_1 generation was"). Either can be correct according to context.

Similarly, when discussing patients in a study, should we say "the cohort is" or "the cohort are"? The answer depends on whether you are focusing on the group as a discrete unit (perhaps to distinguish it from another cohort; "the cohort is") or the individual members ("the cohort are"). Either can be correct depending on context.

For "number of ", we can use a simple rule: "a number are", "the number is". So, "a number of cattle are" (the animals themselves), "the number of cattle is" (how many).

Note that UK usage and US usage differ in relation to companies and teams. UK usage treats companies and sports teams as plural ("British Telecom are", "the England team are"), focusing on the members of the group, whereas US usage treats them as singular ("IBM is", "the American volleyball team is"). In both US and UK usage, you can avoid the problem by being explicit: "The employees of British Telecom are …"; "The band Pink Floyd is …".

"repeated *x* times"

> ✗ The residue was rinsed in ultrapure water and centrifuged at 10 000× *g*. This step was repeated 3 times.

So, how many times was the step performed *in total*? Most authors mean a total of 3 times. However, the strict meaning is a total of 4 times. Consider:

The step was performed once. Then it was repeated 1×. So the total is 2.

The step was performed once. Then it was repeated 2×. So the total is 3.

The step was performed once. Then it was repeated 3×. So the total is 4.

Native English speakers make this mistake too.

An easy solution to this confusion is to write:

> ✓ The residue was rinsed in ultrapure water and centrifuged at 10 000× *g*. This step was performed a total of 3 times. [That is, you state the total number.]

As another example: "IIEs [intermittent isometric extensions] were repeated 10 times." This implies that there were 11 cycles altogether (an original + 10 repeats). But the author intended 10 cycles. Solution: "IIEs were performed 10 times."

Similar confusion concerns "replicates" and "replications". A replicate is 1 of several copies. A replication is another copy. For example:

> We established a randomized block design with 3 replications and 10 plants per replicate. [$n = 4$]

> We established a randomized block design with 3 replicates and 10 plants per replicate. [$n = 3$]

The simplest solution is to use the standard notation for sample size (n): "We replicated the experiment ($n = 3$)."

Improving expression—being clear and concise

Scientific writing can be boring for those who are not interested in it. If authors have not taken care to choose the right words or to express themselves clearly, it can be misleading too. This section explains several ways to express yourself better through consistency and through conventions of clear communication.

Abbreviations

- Define abbreviations on the first use (often in the Abstract and again in the main text).
- Define them only if they are used several times (e.g. at least 3 times).
- Don't abbreviate short terms (e.g. don't abbreviate "postharvest" to "PH").
- Don't define very common abbreviations (e.g. DNA, AIDS, EDTA). Many journals provide a list of abbreviations that do not need to be defined.
- Don't define SI units (e.g. m, V, A, J, kg) or atomic symbols (N, P, K, Mg, Pt etc.).
- Always use standard abbreviations; don't make up a new one if a standard one exists.

Abbreviations should be immediately **meaningful** to the reader. They save space but must not reduce comprehension. For example, you could abbreviate "one-banded bee" and "two-banded bee" to "OBB" and "TBB", but it would be more meaningful to readers if you call them "1BB" and "2BB" (1 and 2 are obviously numbers, O and T are not). Or don't abbreviate them at all.

The remote control for my television set has many TLAs (three-letter abbreviations). I don't know what most of them stand for. They are not *meaningful*.

Active versus passive voice

Use **active** voice to show who did what

Use **passive** voice when "who" is not important

Use **passive** voice also to put familiar information in the topic position or new information in the stress position

Active: "Someone did something." (Actor → object)

Passive: "Something was done by someone." (Object ← actor)

An active construction:

- makes the sentence more forceful
- puts an actor in
- often uses fewer words
- is more interesting
- puts the actor at the beginning of the sentence, where (in English) it has more force.

Most passive sentences can be improved by being made active:

> Passive: The flowers were pollinated by the bees.
>
> Active: The bees *pollinated* the flowers.

Passive: The wetlands <u>were restored by</u> government workers.

Active: Government workers *restored* the wetlands.

However, sometimes passive is better:

The Prime Minister <u>was run over by</u> a car.

The subjects <u>were weighed</u> once a week.

Dates must be clear

What does the date "10.11.05" (or "10/11/05") mean to you? In Commonwealth countries and the US military it means the 10th day of November 2005. In the non-military USA it means the 11th day of October 2005. In China and Japan it means 5th day of November 2010. It may also mean any of these 3 things in any of these regions; even where standards exist, many people don't follow them.

Mixed usage is ambiguous. Scientific writing must be clear. *Scientific Style and Format* (Council of Science Editors 2006) recommends only 2 date formats:

- "2005 November 10" (largest to smallest)
- "10 November 2005" (smallest to largest).

Note that both follow a logical order. In these examples the meaning is obvious, as it is in "November 10, 2005" (but this order is illogical, and extra punctuation is needed).

In abbreviations, the only unambiguous format is "2005 Nov 10" (or "10 Nov 2005").

Exponents versus slashes

Most publishers require the use of exponents in mathematical usage. For example, $km{\cdot}h^{-1}$, $W{\cdot}m^{-2}{\cdot}s^{-1}$, $m^2{\cdot}s^{-1}{\cdot}K^{-1}$. This is because using multiple slashes ("/") can be ambiguous. For example, $W/m^2/s$ could be interpreted as either $W/m^2 \div s = W/m^2s$ or $W \div m^2/s = Ws/m^2$. The use of exponents clearly identifies the numerators (those units with a positive exponent, such as W here) and the denominators (those units with a negative exponent, such as m here).

There is no ambiguity with just 2 terms: km/h is as clear as $km{\cdot}h^{-1}$.

If you use slashes, then you should insert parentheses to remove any ambiguity. For example, does a/b/c mean a/bc or ac/b? Although there are mathematical rules for determining the calculation order, these rules are not followed consistently and may not be taught at all in some parts of the world. Substituting numbers will make this easier to follow. What is 15/3/2? If what is meant is $15/3 \div 2$, then the solution is $5 \div 2 = 2.5$. But if what is meant is $15 \div 3/2$, then the solution is $15 \div 1.5 = 10$. Inserting parentheses will make it clear: $(15/3)/2 = 2.5$; $15/(3/2) = 10$.

Note, however, that the use of mathematical exponents is not consistent with physical objects. It is not logical to write "3 mg kg^{-1} $sheep^{-1}$", because a sheep is not a mathematical concept, or "$sheep^2$" would be correct. In this case, write "3 mg kg^{-1} per sheep".

First-person science

We have all read sentences like "It is concluded that Pqr1 mediates the daylength response in petunia." Or "It is thought that this result was due to contamination." But "it is thought", "it is concluded" are slippery words: they allow the author to escape like a slippery fish. The reader can't answer the question "Who thinks it?" or "Who believes it?" because the author doesn't say. This is called writing in the "third person". (First person = I, we; second person = you; third person = he, she, it, they.)

Authors use the third person for many reasons:

They don't know who thinks it but don't want to admit it. This is poor scholarship. If a fact cannot be verified, it is no more useful than hearsay. If it's important enough to be used in an argument, then it's important enough to verify and support by a literature citation.

They can't be bothered looking it up. More poor scholarship.

They don't want to admit that *they* think it for fear that it won't sound important or reliable enough. This stems from the belief that "science is objective" (see below), or that their opinion doesn't count as much as someone else's.

They're unsure about the conclusion and want some wiggle room in case it turns out to be wrong. Again, poor scholarship. If the data are clear, then the conclusion is clear. If the data are unclear, then say so. (But note that uncertainty is acceptable if it is clearly defined.)

That's how they learned it as students. Old habits die hard, and are often passed on to the next generation.

"That's the way reputable journals do it." Yes and no. The best example I can give is the paper that formed the foundation of all modern genetics:

> **A Structure for Deoxyribose Nucleic Acid**
>
> We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest. A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons:
>
> (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other.
>
> (2) Some of the van der Waals distances appear to be too small.
>
> (Watson JD, Crick FHC. 1953. Nature 171: 737–738.)

If this wording is good enough for winners of the Nobel prize and for publication in *Nature*, it's good enough for other journals.

"Because science is objective." No. Science is **not** objective. Science is a product of the human mind. The human mind is subjective. Every step in the process of science takes place in someone's mind and is subject to all the limitations of the brain. We understand the world from the perspective of our learning, culture, beliefs, knowledge and expectations. The world might be "out there", but the only way we can interpret it is subjectively.

As Albert Einstein showed with his special theory of relativity, our points of view matter. And as Niels Bohr showed us through quantum physics, even the act of observing can determine the outcome. The observer matters.

There are several reasons for writing science in the first person. By acknowledging the people involved, the author:

- makes the article more accessible to the reader (turns a one-way lecture into a conversation)
- lets the reader know exactly who did the work
- takes credit for the work
- accepts responsibility for the work.

What about **Materials and Methods**?

Most authors write the Materials and Methods section in the third person. This is easy for readers to follow, but it can sometimes hide important information, such as whether a measurement was made in the laboratory or by a company.

Use of the third person can sometimes make sentences longer than necessary. For example:

> Eggs were counted, larval emergence was recorded, final instars were weighed, and dates of pupal emergence were recorded.

This can be rewritten in the first person and shorted to:

> We counted eggs, recorded larval emergence, weighed final instars, and recorded dates of pupal emergence.

Where credit is important, use first person

This reduces words and makes it clear who did those tasks—that it wasn't, in fact, another group, for example. To illustrate the point that it can be another group, here is a description from a plant breeding paper:

> A unique mapping population consisting of 54 chromosome segment substitution lines (CSSLs) in rye was recently developed (Smith et al., 2005). Briefly, 79 BC_1F_3 plants developed from self-pollinating BC_1F_1 (Silver/Golden//Silver) by the single-seed descent method were selected as the starting materials for development of CSSLs. Each BC_1F_3 was crossed with Silver, and then the resulting secondary F_1 (SF_1) was crossed with Silver to produce secondary BC_1F_1 (SBC_1F_1).

Good reasons to use the first person:

- It simplifies the text
- It explains why something was done (someone's conscious decision)
- It states a clear intention; it does not suggest that the method performed itself
- It makes it clear who did the work

It could be read either way—either the authors did this or Smith et al. did it—but in this case it was Smith et al. The methods section simply summarised the previously published steps, leaving the impression that the authors repeated the steps.

Third person can lead to **silly statements**:

> <u>To test the expression</u> of the chimeric promoter histochemically, <u>the *Sal*I–*Xba*I fragment</u> of the promoter was inserted into the *Hin*dIII–*Xba*I site of pBI221 …

In this example, the sentence construction implies that the *Sal*I–*Xba*I fragment wanted to test the expression of the promoter. Another sentence from the same paper put it much better:

> To create the chimeric promoters containing 4 or 8 copies of the enhancer-like element, *we digested* the constructs with *Xba*I …

In this case, it is clear that the researchers (not the constructs) wanted to create chimeric promoters.

Jargon

Jargon—the language of a particular profession or group—allows the members of the group to communicate concisely and to discuss concepts that don't exist outside their particular setting, without the need for lengthy descriptions. Jargon is **good** when it communicates clearly to a known audience. Jargon is **bad** when it alienates the audience.

Jargon can be used to exclude outsiders. When a computer salesman tells you about *bits* and *bytes* and *megaflops* and *gigahertz* and *cache*, he is using computer jargon, which is perfectly appropriate in the computer industry, to baffle you and persuade you that he is the expert and therefore you will buy the computer.

On the other hand, when a group of geneticists talk at a conference about *telomeres* and *proteomes* and *homoeoboxes* (or *homeoboxes*) and *PCR* and *sequencing*, they all understand the complex concepts behind these brief labels. In this setting, jargon is inclusive, not exclusive.

Specialised jargon is generally acceptable when you are writing for a specialised journal; it is less acceptable when you are writing for a journal such as *Nature* or *Science*, whose audience spans researchers from many disciplines. The key point when you are writing is to be sure that any term—jargon or otherwise—is appropriate to the audience. Ask yourself whether every single reader will understand each term. If yes (they know as much as you know), then use them. If no (they know less than you know), then can you find a more common term (e.g. "insect blood" for "haemolymph")? Or can you define the term first?

Manufacturer's details—include enough details to find the source

An important part of the information in the Materials and Methods section is the equipment, reagents and software used. Different models of equipment can give different results, different reagents can cause different results, and different software may have subtle bugs. Thus, it is important to include the **manufacturer's full location details** so the same equipment, reagent or software can be obtained, or so that differences in results between your study and previous studies can be explained by the different conditions. You must include the full details:

- model number or item name
- manufacturer's name
- manufacturer's city (and the state or province if applicable)
- the name of the country

or (more simply) a link to the Web site.

Order of events

Place **methods** in strict chronological (time) order. If anyone repeats a method, they can follow it exactly and get it right the first time. Sloppy writing often rearranges the order of events, making it unclear at first read what happened when. For example:

> ✗ Distilled water was added to the ethanol-insoluble fraction which was dried using a centrifugal dryer in vacuum, and the suspension was boiled for 4 h.

This sentence appears to describe **first** the addition of water to a powder, **then** the drying of the mixture, **then** the boiling of a (non-existent) suspension. If we rearrange it in the correct order, it becomes clear:

> ✓ The ethanol-insoluble fraction was dried in a centrifugal dryer in vacuum, distilled water was added, and the suspension was boiled for 4 h.

Much clearer, and shorter too. There's no chance that someone will repeat the steps in the wrong order.

> ✗ Samples were ground in liquid N_2 and then suspended in buffer, after weighing.

Again, this is the wrong order. Change it to:

> ✓ Samples were weighed, ground in liquid N_2, and then suspended in buffer [or]

> ✓ Samples were ground in liquid N_2, weighed, and then suspended in buffer

depending on the actual order of events. (See how it can be ambiguous?)

> ✗ All distal leafy growth units and their mother growth units were carefully detached from each harvested portion (distal leafy units correspond to current-year units in deciduous species; they can be previous-year or older units in evergreen species), after their mother–daughter relationship and the positions of daughter units on the mother units (i.e., terminal or lateral) had been recorded.

Here we have to work through 39 words (including the authors' definition of terms) before we find out that something else happened first. Then we need to rearrange the events in our head. It is a simple fix to rearrange the order:

> ✓ Distal leafy units correspond to current-year units in deciduous species; they can be previous-year or older units in evergreen species. After their mother–daughter relationship and the positions of daughter units on the mother units (i.e., terminal or lateral) had been recorded, all distal leafy growth units and their mother growth units were carefully detached from each harvested portion.

Note that the sentence has been rearranged to define an important term first.

Similarly, place other events in chronological order. For example, in the Results, describe the **preliminary** results that helped you decide reagent concentrations, then describe the results of the **main** experiment that was based on those concentrations. Don't describe the main results and then mention why particular concentrations were tested.

Parallel structures

Parallel structure means that the same concepts are presented in the **same place** in the **same order**, allowing readers to compare them quickly. For the sake of the reader, it is important to maintain a parallel grammatical structure in all writing. This is particularly important in science, where long sentences make keeping track of meaning a challenge.

A common example is the need to ensure that prepositions are in the right place. For example:

> ✗ 5FU/LV is still included in the current regimens, <u>either with</u> irinotecan or oxaliplatin

violates parallelism by placing the preposition ("with") in the wrong place, creating an expectation of something else to follow ("either with irinotecan or oxaliplatin <u>or without them</u>"). There are 2 solutions:

> ✓ … either with irinotecan or *with* oxaliplatin

> ✓ … *with* either irinotecan or oxaliplatin

In both solutions, the preposition now frames the intended object: "with [irinotecan] or with [oxaliplatin]" and "with [either irinotecan or oxaliplatin]". Both constructions are parallel.

The same problem crops up frequently in scientific writing where authors discuss similar results but vary their words and sentence structure, perhaps because they believe that it is wrong to repeat words (not true). Consider the following example:

> ✗ Under 665-nm light, phyB hypocotyls were more elongated than the wild-type and *phyA* while *phyA* was more elongated under 700-nm light than wild-type and *phyB*. Under 690-nm light, *phyB* and *phyA* hypocotyls were more elongated than the wild-type.

This example shows varied word order. Varying the word order can give the false impression that there is some difference in the meaning of the sentences, when there is no difference. This can be confusing, leading readers to expect that different situations are being discussed, when in fact the situations are the same in each instance. See how the information could have been much more clearly expressed:

> ✓ Under 665-nm light, phyB hypocotyls were more elongated than the wild-type and *phyA*. Under 690-nm light, *phyB* and *phyA* hypocotyls were more elongated than the wild-type. Under 700-nm light, *phyA* hypocotyls were more elongated than the wild-type and *phyB*.

The information is now presented in parallel. There is no suggestion that different concepts are being stated. Note that the sentence order was rearranged, as well as the word order in the "700-nm" sentence. Placing concepts in a predictable order helps readers follow the argument better, because the subject is not jumping backwards and forwards in the sentence structure.

Repetition of the same words in the same place or order is important. Consider:

> ✗ In the control, weights increased <u>to</u> 10.3 g <u>from</u> 8.7 g, and <u>from</u> 8.6 g <u>to</u> 18.2 g in the high treatment.

<u>from</u> *x* <u>to</u> *y*

This example, where parallelism is inverted, reverses the order of events. When presenting "from–to" results, *always* follow the order "*from* X *to* Y", never the other way around, because the word order mimics and reiterates the event order.

(The sentence structure thus becomes transparent, allowing the meaning to show through clearly.) By switching around the order of events, the sentence above gives the false impression of parallelism between {10.3 g and 8.6 g} and between {8.7 g and 18.2 g}. It also puts the treatment name in a different place, giving the impression on first reading that the second weight increase occurred in the control too. An appropriate rewrite is as follows:

✓ Weights increased *from* 8.7 g *to* 10.3 g in the control, and *from* 8.6 g *to* 18.2 g in the high treatment.

This restructuring highlights the facts that differ.

A related problem occurs when authors write "The control increased <u>by</u> 8.7 g <u>to</u> 10.3 g." Does this mean "The control increased *from* 1.6 g [the starting value] to 10.3 g [the final value]" or "The control increased *from x* g [the starting point] by 8.7–10.3 g [the magnitude of the increase]"? The meanings are very different.

Another important use of parallel structure occurs in the use of simple words such as prepositions, articles and verbs. We can often find a sentence like:

✗ The study focused on rice, corn, beans, and analysed biomass.

If we break this up into a list, in which the introductory phrase ("The study focused on") points to each item in turn, we can see the problem:

The study focused on
- rice
- corn
- beans
- analysed biomass.

So "The study focused on analysed biomass". The verbal markers are missing, causing the reader to stumble, because what was indicated (by the commas in this case) is not what was meant. Put the markers in, repeat the articles and use the correct verb (to reinforce the parallelism), and we have a comprehensible sentence with a parallel structure:

✓ The study <u>focused on</u> rice, corn, *and* beans *and* <u>analysed</u> biomass.

Plain English in science

I have had scientists say to me that they don't have to write in the style we call plain English, which emphasises simplicity, conciseness and describing key details explicitly instead of leaving them implicit. They consider themselves to be exempt.

This is not true.

Why should authors worry about plain English in scientific writing? Because they have readers. Unless they are writing solely for themselves, they must always consider the readers.

It is important to remember that the readers might include people who are not experts in your subject area. There are students entering a discipline, researchers in other disciplines, policy makers and interested amateurs. Even experts in a subject read to learn something new. All these people should be able to understand what they read.

In science written for lay readers, plain English is essential. But even scientists can benefit, where the aim has been to write clearly, not just to simplify. After all, jargon is essential when scientists communicate with other scientists (when one well chosen word communicates more clearly than a lengthy description).

One of the fundamental features of science is the sharing of knowledge. Poor writing is an impediment to this. A good illustration of this point is a paper by Oswald Avery, Colin MacLeod and Maclyn McCarty published in 1944 in the *Journal of Experimental Medicine*, which established that DNA was the substance that transmitted genetic information. Although it paved the way for James Watson and Francis Crick's milestone paper in 1953 in *Nature* (171: 737–738) that established the structure of DNA, it was not widely read or appreciated. Author Randy Moore has argued that the way it was written was the main reason for this (*Journal of College Science Teaching* 1994 November: 114–121). In comparison with Watson and Crick's paper, it is (as Moore wrote) hesitant, extremely dense, verbose, highly detailed, abstract, impersonal and dull. We've all heard of Watson and Crick. Who has heard of Avery, MacLeod and McCarty?

The *Journal of Natural Resources and Life Sciences Education* (1993; 22(2): 198–199) puts it well:

> The author and his or her closest colleagues will be the only people who read a truly murky piece of writing ….A truly outstanding piece of writing will be widely read, widely quoted and cited, and will bring great rewards to its writer … .
>
> The secret of producing an outstanding piece of writing is to always keep the reader in mind. Authors who keep readers in mind convey their information more lucidly than authors who write only for themselves. The scientist who has the attitude, 'Why should I worry about how this is presented; everybody *knows* what I mean,' is incorrect; everybody does *not* know.

Unfortunately, much scientific writing obscures the message by using fancy words, long sentences, flowery prose, convoluted phrases, poor grammar, imprecise passive sentence construction, copious abbreviations, unnecessary words and vague statements. As writers, we must avoid these impediments to clear understanding.

You will benefit too:

- Ultimately you will save time because journals will ask for fewer corrections.
- You will earn greater respect from your readers if you show that you have taken the trouble to express yourself clearly.
- You will have fewer rejections from funding bodies and consequently more funding.
- You will gain a better understanding of your own work. As Albert Einstein once said, "If you can't explain it to an eight-year-old, you don't understand it."
- Your work will be read more widely.

As Quintilian, a Roman rhetorician (AD 35–100), wrote, "One should not aim at being *possible* to understand but at being *impossible* to *mis*understand."

Position of adverbs

In the same way as an adjective modifies a noun, an adverb modifies a verb.

The position of an adverb can determine the meaning of a sentence.

In writing, particularly scientific writing, word order matters more than in speech. This is because in speech we have facial expressions, hand gestures and voice emphasis at our disposal, as well as the words, and we get immediate feedback from our listeners, allowing us to modify our message as we go. In writing, we have only words, and there is no feedback. So if we want our readers to understand our writing on first reading, then we must pay attention to word order.

Read the following similar examples out loud:

> I only like green tea.
>
> She only reads fiction.
>
> The Prime Minister only listens to his favourites.

In each case, where did you put the emphasis? On "green", "fiction" and "favourites"? And you de-emphasised "only"? No one listening to you would mistake your meaning. But in writing, alternative interpretations are possible. The second example, for instance, could imply "She doesn't *write* fiction, she only *reads* it". In speech, we would emphasise "reads" or "writes", as appropriate. But we don't have this capacity in writing, being limited to italics and punctuation, which are not as subtle as voice.

The same argument applies to other adverbs, too. For example:

> Imperfect clones were mostly contained in library D.

This could, conceivably, be interpreted to imply "But some clones were also *annotated* in library D". But see how the sentence is strengthened by repositioning the adverb:

> Imperfect clones were contained *mostly* in library D.

If we spoke the original sentence, we would emphasise "mostly", making the meaning clear. But in the rewritten version, the emphasis is placed naturally on "mostly". This allows the possibility that some clones were contained in library E.

An alternative solution comes from deleting every non-essential adverb. Far from strengthening prose, adverbs can weaken it. We can convert the adverb into an adjective, removing the ambiguity in the meaning of "mostly" that can result from regional idiom in the placement of adverbs:

> *Most* imperfect clones were contained in library D.

We could even quantify the meaning:

> Ninety per cent of imperfect clones were contained in library D.

Adverbs have a natural place in English, though this can vary with idiom.[1] The *Concise Oxford Dictionary* explains (10th ed, p 1686):

> The position of adverbs in phrases and clauses follows fairly clear rules, i.e. they normally come between the subject and its verb, or between an auxiliary and its main verb.

For example:

> We quickly froze the samples. (Between subject and verb.)

> The rats were painlessly euthanized. (Between auxiliary and main verb.)

I'd like to concentrate on this second situation. Writers will sometimes alter this order, putting the adverb up front:

> The rats painlessly were euthanized.

This is not grammatically wrong. However, it can make for clumsy sentences, and can sometimes alter the meaning. For example:

> The protocol was always followed carefully.

> The protocol always was followed carefully.

There's a subtle difference. The first sentence tells us that the authors always followed the protocol. The second shifts the emphasis onto "was", connoting a very long-term trial in the historical past. (Like in the sentence "Lord Melbourne always was a favourite of the Queen.") Try saying it out loud, putting the emphasis on "was" in the second sentence. Can you hear the difference? Another example:

> *Betula glauca* is also frequent in mountainous areas.

> *Betula glauca* also is frequent in mountainous areas.

The first sentence tells us that as well as being frequent in lowland areas, *B. glauca* is frequent in mountainous areas. The second sentence tells us that another species is frequent in mountainous areas, and *B. glauca* is frequent there as well. As this shows, this shift of emphasis can be subtle but important in the right place.

> The instrument sometimes was difficult to calibrate. (Confirming suspicions that it might have been.)

> The instrument was sometimes difficult to calibrate. (But mostly it was fine.)

But in most cases the "natural" word order is preferable. This natural word order also extends to "split infinitives" (an outmoded principle of style based on a false understanding of grammar). It is perfectly good English to write "to boldly go".

Be aware also that, although adverbs always modify a verb, they can sometimes have an association with ideas. Placement of an adverb can emphasise one idea over another. Compare the meanings implicit in the next 2 sentences (where "more or less" has an adverbial function):

1. English is annoying in having so many exceptions to all its rules. This is where you need to develop a good ear and not stick rigidly to the rules.

> Such variation was synchronized *more or less* among branches.

> Such variation was *more or less* synchronized among branches.

The first suggests that although variation was synchronised among branches, some variation was synchronised among other organs. ("More or less" associates with "among branches".) The second suggests that synchronisation was pretty much complete. ("More or less" associates with "synchronized".) However, these meanings are not explicit; they depend on the reader's interpretation. This shows a good reason to remove adverbs or to qualify their intended meaning.

Here's another example of where adverb placement can have subtle but important effect on meaning:

> The formation of axillary buds normally occurred.

> The formation of axillary buds occurred normally.

The first wording indicates that axillary buds were formed in most years. The second indicates that the axillary buds, when they were formed, looked normal, not malformed. Pay attention to adverb placement, because it can have an important effect on the meaning that the reader takes.

Put subject and verb at (or near) the front of long sentences

A common practice in scientific writing is to write a great long string of items before finally explaining their purpose or what happened to them. This is unfair to readers, who have to read through several lines, keeping all those words in their working memory before being told how to process them. The following sentence is a good example:

> Using these lines, the QTLs related to agronomic traits such as heading date (Brown *et al.*, 1998; Green *et al.*, 2005), ripening (Black *et al.*, 2002), and grain size (White *et al.*, 2002) or those related to physiological traits such as tolerances of excess Al (Gray *et al.*, 2002) and P deficiency (Gold *et al.*, 1998) have been identified.

Dangling preposition: introductory phrase that does not belong to the subject

That's 4 lines containing 29 words and 5 references before the reader is told what happened to the QTLs. Notice that this sentence is written in the passive form—there is no actor controlling the action—and contains a dangling preposition. Now see how it changes when we turn it into an active sentence, which is also written in the first person:

> Using these lines, *we have identified* the QTLs related to agronomic traits such as heading date (Brown *et al.*, 1998; Green *et al.*, 2005), ripening (Black *et al.*, 2002), and grain size (White *et al.*, 2002) or those related to physiological traits such as tolerances of excess Al (Gray *et al.*, 2002) and P deficiency (Gold *et al.*, 1998).

See how it now tells you immediately what happened? There's no need to suspend processing of all the words until you find out how to process them: you know right from the start that everything in the sentence is something the authors identified, rather than having to reach the end of the sentence to find out why the author is mentioning all these traits.

The original sentence attributes the action of "using these lines" to the QTLs. So the rewrite also removes that error and correctly attributes the action to the authors.

> The results of a socio-economic survey of farming systems practising rice culture with introduced fish, rice culture with indigenous fish and rice monoculture in the semi-deep waters of the Mekong Delta, Vietnam, are presented.

Again, you are required to suspend all these words and their possible fates until right at the end. Simply placing the verb up the front and making the sentence active makes it much easier to assimilate:

> *This paper presents* the results of a socio-economic survey of farming systems practising rice culture with introduced fish, rice culture with indigenous fish and rice monoculture in the semi-deep waters of the Mekong Delta, Vietnam.

Symbols

Many people don’t know how to type special symbols (such as mathematical symbols) on the computer, and so use substitutes. A common substitute is x (the letter) for × (the multiplication sign). You can improve the readability of a text by correcting such substitutions. This also improves the scientific accuracy of the text and the final appearance. As you can see in the following table, the differences can be significant, particularly if formatting is lost.

In Unicode, most common symbols are available in the base font. However, problems can arise during the automated typesetting that is the norm these days, when a non-standard character gets substituted with something entirely different. This makes sticking to standard symbols (and Western fonts) important.

Always check the publisher’s guidelines. Some publishers specifically forbid the use of the Symbol font; others insist on it. If this is the case, the safest approach is mark the symbols so that the journal staff will see them and convert them. Similarly, some publishers request the use of equation-editor software such as that included in MS Word, whereas others forbid it.

Symbol	Name	Don't use	Comments
×	Multiplication ("times")	x	In equations, the letter x can be confused with the variable *x*
·	Raised decimal point (in chemical formulae and mathematical expressions)	.	Full point (full stop or period)
°	Degree	^{o} ^{0} º ˚	Superscripted letter o or number 0, masculine ordinal[1] and ring diacritic[1]
–	En dash (or en rule)	- --	Hyphens
—	Em dash (or em rule)	- ---	Hyphens
′ ″	Minutes and seconds	’ ” ' " ´	Quotation marks, acute diacritic
≤ ≥	Less than or equal to; greater than or equal to	$\underline{<}$ $\underline{>}$	Underlining on < > (underlining can be lost)
µ	Greek letter mu	u	Roman letter u
±	Plus or minus	$\underline{+}$	Underlining on + (underlining can be lost)
β	Greek letter beta	ß	The German double-letter "ss" (eszett)
‘ ’ “ ”	Typographic (curly) quotes	' "	Typewriter (straight) quotes[2]

1. Masculine ordinal: e.g. 3º = 3rd. Ring diacritic: Å å.

2. By having opposite shapes, typographic quotes clearly indicate the start and end of terms. Straight quotes can make the start and end less clear.

Statistics

All science depends on statistics. Most scientific papers present statistics as evidence of mathematically justifiable results. Statistics is a very broad, complex area. I am not a statistician, and this book is not a lesson in how to use statistics. However, you must present enough statistical information that readers can be sure that your results are correct. In this section I have summarised some very good advice on reporting statistics. I recommend that you download the full information.

Uniform Requirements

The *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals* [http://www.icmje.org/recommendations/] says:

> *"Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as P values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used."*

In other words, you need to present the following information:

- The name of each test that you used. (If the test is not common, cite a reference to it.)
- A statement that the data meet the assumptions required for the use of each test (e.g. normally distributed).

- The number of samples (*n*).
- The results (in numbers, if possible).
- A measure of the centre: mean, median, mode or other.
- A measure of the measurement error or uncertainty: standard deviation (not standard error, which is a measure of precision), 95% confidence interval or range, for example.
- The probability that conclusion is correct—that is, the *P* value.
- The name of the computer software used.

Nature Statistical checklist

In more detail, *Nature* used to provide a brief (1 page) but thorough "Statistical checklist", which is summarised below:

Tests used

- What is being compared?
- Which tests have you used?
- Which statistical methods have you used?
- Why did you use this test (and not another)?
- How did you avoid sampling bias (e.g. randomisation)?
- Did you transform the data?
- Are the data normally distributed? If so, is the test appropriate?
- If the data are not normally distributed, which test did you use?

Details of the tests

- What is the value of *n*?
- What is the α (alpha) level (the threshold value for significance)?
- Is the test 1-tailed or 2-tailed?
- What is the *P*-value?

Summary of descriptive statistics

- What is the value of *n* for each data set?
- What is the measure of centre (mean, median, mode, other)?
- What is the measure of variability (SD, 95% confidence interval, range)?
- Does "±" mean standard deviation (SD) or standard error of the mean (SEM)?

Unexpected results

- Explain how you treated any unusual data, or any unusual statistical methods that you used.

Download the "Statistical checklist" from *Nature* for a fuller explanation.

SAMPL Guidelines

In much more detail, Tom Lang (of Tom Lang Communications and Training International, USA) and Douglas Altman (of Oxford University) have published the SAMPL Guidelines [https://www.equator-network.org/wp-content/uploads/2013/07/SAMPL-Guidelines-6-27-13.pdf]:

Name of test

n = ?

Mean ± SD (not SEM)

[normally distributed data]

α = ?

P = ?

Lang T, Altman D. 2013. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H, Polderman A, eds. Science Editors' Handbook, European Association of Science Editors: http://www.ease.org.uk/publications/science-editors-handbook.

The SAMPL (Statistical Analyses and Methods in the Published Literature) Guidelines present a comprehensive list of principles for reporting statistics. You can download them (5 pages) from the above link. They are reproduced in full in Appendix 4.

The covering letter

You will need to write a covering letter (or "cover letter") when you submit your paper to the journal. Keep it as short as possible.

Write to the person, not the position. For example, if the editor of *The Rice Journal* is Dr M. Satoh, address your letter to Dr Satoh, not to "the Editor".

✓		✗	
	Dr M. Satoh Editor, *Journal of Rice* [address] Dear Dr Satoh, ⋮		The Editor *Journal of Rice* [address] Dear Sir, / Dear Editor, / Dear Dr M. Satoh, ⋮

If you do not know the name of the editor, try to find it: on the Web site, in the journal, in the Instructions to Authors. If you cannot find it anywhere, then write "Dear Editor", not "Dear Sir or Madam".

Try to keep the letter brief. Here are some phrases you can use:

Please find enclosed ...	our paper entitled " ... "	for consideration for publication in *The Rice Journal*.
We would like to submit ...	our report on " ... "	as a proposed article in *The Rice Journal*.
We are pleased to submit ...	our latest study, titled " ... "	for your consideration.

- In 2 or 3 sentences, tell the editor why your work is important.
- Explain why the readers of the journal will be interested in it.
- Confirm that it is original research.
- State that all authors agree with the findings and agree to publication. (Remember to ask all authors first!)
- State that the paper has not been published in another journal, and is not under consideration by another journal. (Never submit a paper to 2 or more journals at the same time.)
- If it is relevant, list some recent publications by you.
- If the journal asks, suggest some reviewers. Provide their e-mail addresses to make it easy for the journal editor to contact them. If necessary, state unsuitable reviewers.
- If you used a professional rewriter or editor, state this too.
- State the name and address of the corresponding author.

Then finish with a friendly statement that you hope to hear from the editor soon.

Department of Genetics
Tokyo University of Genetics
[address]
[date]

Dr John Smith
Editor, *Journal of Rice*
[address]

Dear Dr Smith,

Please find enclosed our manuscript entitled “Genetic control of hull colour in rice”, by M. Satoh, S. Yano and A. Watanabe.

This report is the first, to our knowledge, that explains colour variation in rice hulls. We believe that it will be of relevance to readers of the *Journal of Rice* because hull colour has been linked to drought tolerance.

We confirm that the research is original, that this manuscript has not been published elsewhere, and that it is not under consideration by another journal. All authors have approved the manuscript and agree with its submission to the *Journal of Rice*.

Our recent paper published in *Nature Genetics* (...) explains our preliminary investigations.

We suggest Dr L. Arborio (University of Turin, Italy; arboriol@u-turin.it) and Dr T. Basmati (Lahore University, Pakistan; basmatit@lahore.ac.pk) as possible reviewers. Owing to research competition, we ask that you not approach Dr Aki Hikari (Lake Biwa University).

The English has been extensively revised by a native-English-speaking scientific editor of Xyz Scientific Language Service.

You can reply to me at the above address or by e-mail at [xxx]. I look forward to hearing from you at your convenience.

Yours sincerely,

[you]

Appendix 1: English plurals

(Note that some words are printed in grey: These words are valid but are almost never used.)

Anglo-Saxon (Germanic)

Singular	Plural
man	men
woman	women
foot	feet
goose	geese
tooth	teeth
mouse	mice
louse	lice
dormouse	dormice
ox	oxen
child	children

No plural form of the word

Singular	Plural
advice	–
eaves	–
equipment	–
information	–
research	–
summons	–

No singular form of the word

Singular	Plural
–	annals [are]
–	auspices [are]
–	cattle [are]
–	head (of cattle) [are]
–	intestines [are]
–	means [is/are]
–	measles [is]
–	mumps [is]
–	news [is]
(person)	people, persons [are]
–	poultry [are]
–	scissors [are]
–	statistics [is]
–	thanks [are]
–	tongs [are]
–	tweezers [are]
–	vermin [are]

Singular = plural

Singular	Plural
aircraft	aircraft
bison	bison
buffalo	buffalo
cod	cod
corps	corps
craft	craft
deer	deer
fish	fish
forceps	forceps
goods	goods
plankton	plankton (but note: plankter, plankters)
progeny	progeny
remains	remains
salmon	salmon
sheep	sheep
shrimp	shrimp
species	species
sperm	sperm
squid	squid
swine	swine
trout	trout

Latin

You can use the traditional plural spellings if you want. They are correct

Or you can use the simplified plurals if you prefer. "When in doubt, add "s"

Singular	Plural (traditional)	Plural (simplified)
-um → *-a*		
addendum	addenda	
agendum	agenda	agendas
bacterium	bacteria	
curriculum	curricula	curriculums
corrigendum	corrigenda	
datum	data	
dictum	dicta	dictums
effluvium	effluvia	
erratum	errata	
maximum	maxima	maximums
medium	media	
memorandum (memo)	memoranda	memos
millennium	millennia	
minimum	minima	minimums
ovum	ova	
speculum	specula	
spectrum	spectra	
stadium	stadia	stadiums
stratum	strata	
-us → *i*		
alumnus	alumni	
bacillus	bacilli	

In some cases, the traditional plural is never used in English

bronchus	bronchi	
cactus	cacti	cactuses
campus		campuses
census		censuses
focus	foci	focuses
fungus	fungi	
gladiolus	gladioli	
nucleus	nuclei	
radius	radii	
stimulus	stimuli	
stylus	styli	
terminus	termini	
uterus	uteri	uteruses
virus		viruses
-a → -ae		
alga	algae	
amoeba	amoebae	amoebas
antenna	antennae	antennas
formula	formulae	formulas
lamina	laminae	
larva	larvae	
medulla	medullae	medullas
mucosa	mucosae	
nebula	nebulae	nebulas
papilla	papillae	
sequela	sequelae	
theca	thecae	

Words written in grey are valid words that are almost never used with that specific meaning. They are correct, but they can confuse people

-us → -ora/-era		
corpus	corpora	
genus	genera	
viscus	viscera	
-ix/-ex → -ices		
appendix	appendices	appendixes
cortex	cortices	cortexes
helix	helices	helixes
index	indices	indexes
matrix	matrices	matrixes
radix	radices	
vortex	vortices	vortexes
-en → -ina		
foramen	foramina	
lumen	lumina	
Irregular		
axis	axes	
iris	irides	irises
stamen	stamina	stamens

Same spelling		
apparatus	apparatus (same)	
meatus	meatus (same)	
series	series (same)	
species	species (same)	
status	status (same)	
syllabus	syllabus (same)	syllabuses

Greek

Singular	Plural (traditional)	Plural (simplified)
-is → -es		
analysis	analyses	
basis	bases	
crisis	crises	
ellipsis	ellipses	
hypothesis	hypotheses	
metamorphosis	metamorphoses	
oasis	oases	
parenthesis	parentheses	
synopsis	synopses	
testis	testes	
thesis	theses	
-on/-um → -a		
gymnasium	gymnasia	gymnasiums
automaton	automata	automatons
criterion	criteria	
phenomenon	phenomena	
polyhedron	polyhedra	
-a → -ata		
dogma	dogmata	dogmas
lemma	lemmata	lemmas
stigma	stigmata	stigmas
stoma	stomata	stomas
Irregular		
calyx	calyces	
chrysalis	chrysalides	chrysalids
octopus	octopodes	octopuses
platypus	platypodes	platypuses
phalanx	phalanges	

French

Singular	Plural (traditional)	Plural (simplified)
bureau	bureaux	bureaus
milieu	milieux	milieus
plateau	plateaux	plateaus
tableau	tableaux	tableaus

Abbreviations

"Mr" (mister) and "Mrs" (mistress or missus) have no plural in English, so we use the French plurals instead

Singular	Plural (traditional)	Plural (simplified)
Mr	Messrs (French)	
Mrs	Mmes (French)	
Miss	Misses	
Ms	–	Mses
p.	pp.	
sp.	spp.	
cv.	cvv.	
pv.	pvv.	

Appendix 2: Prepositions

If you are not sure which preposition to use, follow the examples here. (Other words are possible too.)

in a company, department	**in** IBM, **in** the Department of Statistics
in a country	**in** Japan, **in** Nauru (as a country)
on a day or date	**on** day 7, **on** Monday, **on** 12 June
in a field, park	**in** a rice paddy, **in** the Ogasawara National Park
on an island	**on** Okinawa, **on** Nauru (as an island)
in a lake, sea, ocean	(fish) **in** Lake Biwa, (whales) **in** the Pacific Ocean
on a lake, sea, ocean	**on** Lake Biwa (in a boat), **on** the Pacific Ocean (in a ship)
on land	**on** the farm, **on** the campus of the University
on a map	**on** the 1:20 000 Hokkaido map
in a month, season, year	**in** June, **in** summer, **in** 2014
on <u>a</u> mountain, <u>a</u> hill	**on** Mt Fuji, **on** the tallest hill
in mountain<u>s</u>, hill<u>s</u>	**in** the Japan Alps, **in** the Tuscan Hills
in a prefecture, state, province	**in** Ibaraki Prefecture, **in** California
at a site	**at** the flux tower site
at a time	**at** 09:30, **at** dawn, **at** midday
in a town, city	**in** Tokyo, **in** New York
at a university, hospital	**at** Tsukuba University, **at** Tokyo General Hospital

Appendix 3: Punctuation marks in English

Symbol	Name	Comment
.	full stop or full point or period	Ends a sentence
,	comma	Weaker
;	semicolon	⇅
:	colon	Stronger
-	hyphen	Hyphens with different functions are available
–	en dash	ranges (10–20); to replace “and” (US–Japan)
—	em dash	Use—like this—to separate an idea (parenthetical statements)
( )	parentheses	Hierarchy: ([{ }])
[ ]	square brackets	
{ }	braces	
< >	angle brackets	
!	exclamation mark	
?	question mark	
‘ ’	single quotation marks	
“ ”	double quotation marks	
' "	typewriter quotation marks	
′ ″	single and double primes	
’	apostrophe	Same as single closing quotation mark
…	ellipsis or suspension points	Indicates an omission
/ ⁄	slash or solidus or virgule or oblique stroke	m/s ⅔
*	asterisk	}
†	dagger	}
‡	double dagger	} Often used for footnotes
§	section mark	} (in the order given here)
¶	paragraph mark	}
‖	parallels	}
•	bullet	
~	swung dash	Not used to mean a range, unlike in Japanese

Not normally used in English: ¡ ¿ ‹ › « » ‚ ‘ „ “ 「」 `

Appendix 4: SAMPL Guidelines

Reporting Basic Statistical Analyses and Methods in the Published Literature: The SAMPL Guidelines for Biomedical Journals

Guiding Principles for Reporting Statistical Methods and Results

Our first guiding principle for statistical reporting comes from The International Committee of Medical Journal Editors, whose Uniform Requirements for Manuscripts Submitted to Biomedical Journals include the following excellent statement about reporting statistical analyses:

> "**Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results.** [Emphasis added.] When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as *P* values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used".

Our second guiding principle for statistical reporting is to **provide enough detail that the results can be incorporated into other analyses.** In general, this principle requires reporting the descriptive statistics from which other statistics are derived, such as the numerators and denominators of percentages, especially in risk, odds, and hazards ratios. Likewise, P values are not sufficient for re-analysis. Needed instead are descriptive statistics for the variables being compared, including sample size of the groups involved, the estimate (or "effect size") associated with the P value, and a measure of precision for the estimate, usually a 95% confidence interval.

General Principles for Reporting Statistical Methods

Preliminary analyses

- Identify any statistical procedures used to modify raw data before analysis. Examples include mathematically transforming continuous measurements to make distributions closer to the normal distribution, creating ratios or other derived variables, and collapsing continuous data into categorical data or combining categories.

Primary analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- When possible, identify the smallest difference considered to be clinically important.
- Describe fully the main methods for analyzing the primary objectives of the study.
- Make clear which method was used for each analysis, rather than just listing in one place all the statistical methods used.
- Verify that that data conformed to the assumptions of the test used to analyze them. In particular, specify that 1) skewed data were analyzed with non-parametric tests, 2) paired data were analyzed with paired tests, and 3) the underlying relationship analyzed with linear regression models was linear.
- Indicate whether and how any allowance or adjustments were made for multiple comparisons (performing multiple hypothesis tests on the same data).
- If relevant, report how any outlying data were treated in the analysis.

Lang T, Altman D. Statistical Analyses and Methods in the Published Literature: the SAMPL Guidelines.

- Say whether tests were one- or two-tailed and justify the use of one-tailed tests.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- Name the statistical package or program used in the analysis.

Supplementary analyses

- Describe methods used for any ancillary analyses, such as sensitivity analyses, imputation of missing values, or testing of assumptions underlying methods of analysis.
- Identify post-hoc analyses, including unplanned subgroup analyses, as exploratory.

General Principles for Reporting Statistical Results

Reporting numbers and descriptive statistics

- Report numbers—especially measurements—with an appropriate degree of precision. For ease of comprehension and simplicity, round to a reasonable extent. For example, mean age can often be rounded to the nearest year without compromising either the clinical or the statistical analysis. If the smallest meaningful difference on a scale is 5 points, scores can be reported as whole numbers; decimals are not necessary.
- Report total sample and group sizes for each analysis.
- Report numerators and denominators for all percentages.
- Summarize data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean ± SD.
- Summarize data that are not normally distributed with medians and interpercentile ranges, ranges, or both. Report the upper and lower boundaries of interpercentile ranges and the minimum and maximum values of ranges, not just the size of the range.
- Do NOT use the standard error of the mean (SE) to indicate the variability of a data set. Use standard deviations, inter-percentile ranges, or ranges instead. (The SE is an inferential statistic—it is about a 68% confidence interval—not a descriptive statistic.)
- Display data in tables or figures. Tables present exact values, and figures provide an overall assessment of the data.

Reporting risk, rates, and ratios

- Identify the type of rate (e.g., incidence rates; survival rates), ratio (e.g., odds ratios; hazards ratios), or risk (e.g., absolute risks; relative risk differences), being reported.
- Identify the quantities represented in the numerator and denominator (e.g., the number of men with prostate cancer divided by the number of men in whom prostate cancer can occur).
- Identify the time period over with each rate applies.
- Identify any unit of population (that is, the unit multiplier: e.g., × 100; × 10,000) associated with the rate.
- Consider reporting a measure of precision (a confidence interval) for estimated risks, rates, and ratios.

Lang T, Altman D. Statistical Analyses and Methods in the Published Literature: the SAMPL Guidelines.

Reporting hypothesis tests

- State the hypothesis being tested.
- Identify the variables in the analysis and summarize the data for each variable with the appropriate descriptive statistics.
- If possible, identify the minimum difference considered to be clinically important.
- For equivalence and non-inferiority studies, report the largest difference between groups that will still be accepted as indicating biological equivalence (the equivalence margin).
- Identify the name of the test used in the analysis. Report whether the test was one- or two-tailed (justify the use of one-tailed tests) and for paired or independent samples.
- Confirm that the assumptions of the test were met by the data.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- At least for primary outcomes, such as differences or agreement between groups, diagnostic sensitivity, and slopes of regression lines, report a measure of precision, such as the 95% confidence interval.
- Do NOT use the standard error of the mean (SE) to indicate the precision of an estimate. The SE is essentially a 68% confidence coefficient: use the 95% confidence coefficient instead.
- Although not preferred to confidence intervals, if desired, *P* values should be reported as equalities when possible and to one or two decimal places (e.g., $P = 0.03$ or 0.22 not as inequalities: e.g., $P < 0.05$). Do NOT report "NS"; give the actual P value. The smallest *P* value that need be reported is $P < 0.001$, save in studies of genetic associations.
- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting association analyses

- Describe the association of interest.
- Identify the variables used and summarize each with descriptive statistics.
- Identify the test of association used.
- Indicate whether the test was one- or two-tailed. Justify the use of one-tailed tests.
- For *tests* of association (e.g., a *chi*-square test), report the *P* value of the test (because association is defined as a statistically significant result).
- For *measures* of association (i.e., the *phi* coefficient), report the value of the coefficient and a confidence interval. Do not describe the association as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, consider including the full contingency table for the analysis.
- Name the statistical package or program used in the analysis.

Reporting correlation analyses

- Describe the purpose of the analysis.
- Summarize each variable with the appropriate descriptive statistics.
- Identify the correlation coefficient used in the analysis (e.g., Pearson, Spearman).
- Confirm that the assumptions of the analysis were met.

Lang T, Altman D. Statistical Analyses and Methods in the Published Literature: the SAMPL Guidelines.

- Report the alpha level (e.g., 0.05) that indicates whether the correlation coefficient is statistically significant.
- Report the value of the correlation coefficient. Do not describe correlation as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, report the (95%) confidence interval for the correlation coefficient, whether or not it is statistically significant.
- For primary comparisons, consider reporting the results as a scatter plot. The sample size, correlation coefficient (with its confidence interval), and *P* value can be included in the data field.
- Name the statistical package or program used in the analysis.

Reporting regression analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, in linear regression indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying values were treated in the analysis.
- Report how any missing data were treated in the analyses.
- For either simple or multiple (multivariable) regression analyses, report the regression equation.
- For multiple regression analyses: 1) report the alpha level used in the univariate analysis; 2) report whether the variables were assessed for a) colinearity and b) interaction; and 3) describe the variable selection process by which the final model was developed (e.g., forward-stepwise; best subset).
- Report the regression coefficients (beta weights) of each explanatory variable and the associated confidence intervals and *P* values, preferably in a table.
- Provide a measure of the model's "goodness-of-fit" to the data (the coefficient of determination, r^2, for simple regression and the coefficient of multiple determination, R^2, for multiple regression).
- Specify whether and how the model was validated.
- For primary comparisons analyzed with simple linear regression analysis, consider reporting the results graphically, in a scatter plot showing the regression line and its confidence bounds. Do not extend the regression line (or the interpretation of the analysis) beyond the minimum and maximum values of the data.
- Name the statistical package or program used in the analysis.

Reporting analyses of variance (ANOVA) or of covariance (ANCOVA)

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying data were treated in the analysis.
- Report how any missing data were treated in the analyses.
- Specify whether the explanatory variables were tested for interaction, and if so how these interactions were treated.
- If appropriate, in a table, report the *P* value for each explanatory variable, the test statistics and, where applicable, the degrees of freedom for the analysis.

Lang T, Altman D. Statistical Analyses and Methods in the Published Literature: the SAMPL Guidelines.

- Provide an assessment of the goodness-of-fit of the model to the data, such as R^2.
- Specify whether and how the model was validated.
- Name the statistical package or program used in the analysis.

Reporting survival (time-to-event) analyses

- Describe the purpose of the analysis.
- Identify the dates or events that mark the beginning and the end of the time period analyzed.
- Specify the circumstances under which data were censored.
- Specify the statistical methods used to estimate the survival rate.
- Confirm that the assumptions of survival analysis were met.
- For each group, give the estimated survival probability at appropriate follow-up times, with confidence intervals, and the number of participants at risk for death at each time. It is often more helpful to plot the cumulative probability of not surviving, especially when events are not common.
- Reporting median survival times, with confidence intervals, is often useful to allow the results to be compared with those of other studies.
- Consider presenting the full results in a graph (e.g., a Kaplan-Meier plot) or table.
- Specify the statistical methods used to compare two or more survival curves.
- When comparing two or more survival curves with hypothesis tests, report the *P* value of the comparison
- Report the regression model used to assess the associations between the explanatory variables and survival or time-to-event.
- Report a measure of risk (e.g., a hazard ratio) for each explanatory variable, with a confidence interval.

Reporting Bayesian analyses

- Specify the pre-trial probabilities (“priors”).
- Explain how the priors were selected.
- Describe the statistical model used.
- Describe the techniques used in the analysis.
- Identify the statistical software program used in the analysis.
- Summarize the posterior distribution with a measure of central tendency and a credibility interval
- Assess the sensitivity of the analysis to different priors.

Lang T, Altman D. Statistical Analyses and Methods in the Published Literature: the SAMPL Guidelines.