How to write scientific papers

Hugh V. Danks

TABLE OF CONTENTS

Danks, H.V. 2024. How to write scientific papers.
(pp. 1–49, 22 July 2024) doi.org/10.7939/r3-0vzm-nj50

Also available at
Entomological Society of Canada blog
and at *Google Drive* until 2026

Introduction

This article recommends how to prepare papers for publication in scientific journals. The same recommendations apply to the preparation of graduate theses. All essential components of how to plan and complete publication of both experimental results and literature reviews are treated.

These treatments describe how to draft and revise a manuscript, including organization, clear writing, and detailed checks. Procedures for acquiring and using references, and for producing appropriate tables and figures, are specified.

Ways to deal with the journal stages of submission, review, and checking proofs are also addressed. Considered too are general elements such as choosing a journal to receive the submission, and ethical and legal issues related to plagiarism, authorship, and copyright.

The recommendations here are based on my experience in writing, editing, reviewing, and reading scientific papers, books, and newsletters. Many of them are summarized in tabular form, and illustrated with examples, mainly from entomology. Selected references are cited.

Planning for publication

Planning for publications about field or experimental research

Scientific research that is worth publishing must have been properly conceived, planned, and executed, including appropriate sample sizes and rigorous analysis. Moreover, any project should include a plan to disseminate the results. Unpublished work is not useful to others, even if it is brilliant. The best projects consider all of the elements from start to finish, not only addressing relevant subjects and using suitable designs, but also aiming for useful products (Table 1).

Table 1. Elements of research planning that enhance publication.

Organization and design
Key questions or hypotheses
Appropriate techniques and equipment
Unbiased selection of samples
Adequate sample sizes
A plan for outputs
Journal papers
Other publications
Specimens or samples
Images
Documents
Databases
Execution
Foolproof procedures
Thoughtful and rigorous analysis
Long-term preservation of necessary material and records
Follow-up through appropriate publication and other outputs
Familiarity with relevant literature
Meticulous preparation of manuscripts

Beyond good organization, a project needs constant care throughout, even for mind-numbingly repetitive chores. Formal written procedures help to maintain consistency, and are also useful to characterize the methods for publication. They ensure that different assistants (and even the same researcher on different days, perhaps widely spaced) contribute comparable data.

Therefore, small details are important. Otherwise, data and the publications derived from them are likely to include minor but potentially important inconsistencies, both known and unknown. For example, taking substrate samples to a fixed depth ensures that results are not distorted by unexpected differences in the vertical distribution of the insects. Keeping only one sample open at a time whilst sorting extensive insect collections eliminates the chance that specimens will be mislabelled.

Alongside careful methods, scrupulous recording of data and notes assists the writing process. In fact, when the data have been painstakingly collected and documented, it is possible to write up other people's results. This route is exemplified by work in the high arctic that was detailed and structured enough to be analyzed and interpreted by someone who had never visited the study area (Danks and Oliver 1972). Indeed, researchers should think carefully about what data and observations to record (including ancillary observations), to ensure that the results are complete and easy to interpret. The records will then be more easily used by the researcher, even many years later if required, and others could use them too.[1]

Publications (of any sort) are not the only possible products of research. Images, videos, and collections are also potentially useful. Specific plans might have to be developed to ensure that assets such as voucher specimens are preserved. Therefore, simply leaping into interesting experiments is inefficient.

Researchers who are not familiar with potential outputs other than journal papers can consult specialists for advice on how to manage them. For example, taxonomists will be able to recommend appropriate methods to preserve and label collections and specimens.

Last but not least, even the smallest projects require a thorough search of the literature (see the next subsection, as well as *References*). Early in my career, an entomologist who worked in systematics, and was fully familiar with all of the taxonomic literature for his group, proudly published a paper about a new ecological feature found in some of the species. However—in the absence of a check of the wider literature—a key paper had been overlooked. It demonstrated that the feature was already widely known!

Planning for literature reviews

It is easy to underestimate the amount of effort and skill required to search the literature properly for references, and to synthesize published information for reviews and theses. The work requires planning and diligence equal to that for studying the actual insects.

Incomplete literature reviews, and "reviews" that compile a set of other people's abstracts without integrating them, are of limited use. Proper syntheses use the power of all available sources of information to find new generalities and insights that will help future work.

Searching for references, generating notes, and keeping track of all of them are assisted by modern digital capabilities. These tools include web and database searches, sophisticated word-processing programs, and reference-management software (see *References* below).

Relatively short reviews are facilitated if many papers are in hand in digital or hard-copy form. Even so, no-one is ready to start until the information in notes or in their head is sufficiently organized.

[1] The way that complete and organized records aid in preparing manuscripts is illustrated by the following anecdote. The numerical data for a paper based on arctic mosquito collections (Corbet and Danks 1973) provided no clues as to why adult activity abruptly declined much earlier than usual in one year, despite permissive temperatures. The only remaining possibility seemed to be that local plants had finished flowering unusually early that year, depriving adults of the nectar needed to fuel activity. Fortunately, detailed field notes had been kept, although they were stored elsewhere. They substantiated the explanation.

Preparing major reviews and books entails greater problems. So much information has to be assimilated that a systematic way to record it is essential. Otherwise, the contents of references consulted early in the project are likely to be forgotten.

I recommend establishing a detailed initial list of contents, which can be used to separate notes on different subjects into appropriate subcategories. Information from a single reference can then be placed into as many different topic files as necessary. Provided the notes are sufficiently detailed and specific, it will be possible to access information for and integrate each individual section and subsection, and to capitalize on them by further synthesis. If the information remains combined, on the other hand, all notes (or even the original papers) will have to be searched again to find items that belong in a particular subsection.

The same strategy is useful during graduate work. Thesis preparation will be greatly simplified if notes and ideas accumulated over several years of study have been organized into separate files, including one for the discussion. However, in this instance a draft structure will take some time to develop, as information is gathered about the subject.

Such structured preparation makes it feasible to write one individual section at a time. That task is manageable, whereas trying to write a whole book or a whole thesis at once may seem overwhelming. Even so, some adjustments will certainly be beneficial as the project proceeds, and some sources will have to be consulted again during revision and checking.

In summary, it is as challenging to *organize* a long review as it is to *synthesize* the scientific findings. In particular, foolproof mechanisms to keep track of references and to document content are essential. A prerequisite for major reviews, therefore, is thorough knowledge of the subject area. Otherwise, the initial structure will be faulty, and potentially important topics will be missing or inadequately covered.

Preparing the manuscript

Introduction

The best papers bring new insights, have well written text that is easy to understand, are supported by tables and illustrations that are necessary and effective, and deliver focussed conclusions that stem from thoughtful analysis. They also integrate up-to-date literature, linking specific findings with known features of the species. Reaching this standard requires concentration at multiple stages, in addition to ensuring that the content is based in research that is well planned and executed.

Therefore, writing is hard work, because so many components have to be developed, improved, balanced, revised, and checked. In particular, the difficulty of keeping a manuscript organized, succinct, and internally consistent seems to increase rapidly as the length increases. As a result (although some projects run more easily than others), it is the longest manuscripts that most commonly prompt the sentiment: “I hate writing, but I like having written”!

The following recommendations stem from the procedures I developed during my career, which are summarized in Table 2. Each of these elements is explained below. Whatever the scheme, organization, diligence, and multiple drafts are essential.

Early in my career, manuscripts were handwritten, and any substantially revised draft had to be transcribed by writing out or typing a new hard copy. It was advisable to reduce the number of hard-copy drafts by composing each sentence of the first draft slowly and laboriously, processing possible options mentally before writing anything down. However, the advance of

Table 2. Elements of preparing a scientific manuscript.

Element	*Main components*
Purpose, target	Context clear within the subject area
	Key conclusions identified
	Preliminary choice of journal, based on subject matter, degree of detail, and other factors (Table 9)
Title, keywords	Title balances detail, precision, and length
	[Title revised later if necessary, and keywords chosen]
Writing	
Contents	Content well organized (typically Abstract, Introduction, Methods, Main sections and appropriate subsections, Discussion, Conclusions, and References)
Paragraphs	Each paragraph in the proper order, with an opening topic sentence, relevant detail, and a concluding (or bridging) sentence
Sentences	Each sentence in an easily understood order, usually opening with the main subject
	Grammar, spelling, and word use correct (Table 3 and text)
	English clear and concise with good style (Table 4 and text)
References	Literature covered completely through comprehensive searches using appropriate search terms
	References managed competently, including an easy-to-use database and painstakingly accurate citations
	All cited references examined closely and used carefully
Tables	All tables necessary, well-chosen, clear, concise, properly constructed, and accurately captioned (Table 5)
Figures	All figures necessary, supporting the intended purpose, and of the appropriate type (Table 6)
	Charts showing the required information in the simplest and clearest form (Figures 1–18)
	Figures with concise and informative captions
	Figures suitable for publication, and consistent with journal requirements (Table 7)
Revision, improvement	Multiple revisions to cement the logic, clarify the meaning, and improve the style
	Sentence-by-sentence checks for remaining errors of grammar, typography, consistency, and style
	Repetition and other distracting author habits eliminated
	Content of cited references verified if necessary
Standardization	Journal standards applied to minutiae such as abbreviations (Table 8), citation formats, and alternative spellings
Last checks	Careful re-read for logic, consistency, integration of tables and figures, and absence of errors
	Verification that the reference list contains all cited references but no uncited ones, and that entries are impeccably correct
	Two independent private evaluations of the “final” manuscript by colleagues
Submission	Adherence to all journal requirements

word-processing software on personal computers has made on-screen digital corrections easy. As a result, it is now more efficient to revise sentences through multiple written drafts, the system implied below.

The first decision for any particular paper is to decide on the message to be delivered, given the author's discoveries in the context of existing information. At an early stage, as results are analyzed and preliminary ideas developed, the author must decide on the interpretation, conclusion, or synthesis to be brought to the reader. Some publications are best divided into more than one paper, or combined into a single paper, to achieve this goal. Each paper must include new input from the author, and lead the reader somewhere, not just deliver a "laundry list" of related items..

In other words, work is ready to publish only when it reaches a staging point that has yielded useful discoveries. These discoveries, and the best way to transmit them, may suggest a suitable journal to receive the submission (see *Choice of journal*). The journal chosen affects many aspects of the manuscript, so it is wise to select a potential outlet early on.

Title

The title of a paper is important. Not only does it tell potential readers whether the content is likely to be of interest, but also it helps users to find references that are relevant.

The most helpful titles strike the right balance of detail, precision, and length. Existing knowledge of the field suggests the most useful approach. It may be helpful to consider the classical questions asked by enquirers: who, what, when, where, why (and how)—although "who" might be insect taxa!

A title that is too general (such as one from the 19th century that read *Some observations on moths in Ontario)* is of little use today. A title that is detailed but has relatively trivial content and is not integrated enough to add value (*The effect of pesticide A on species B in crop C in County D in State E in year F*) serves mainly to clutter search results. Most titles that try to be cute or amusing (*Why did the grasshopper cross the road?* [rather than *Swarming behaviour of early instars of the grasshopper* ...]) waste words at the expense of specific content. Titles that are extremely long tend to hinder search and transcription, even if a journal will accept them.

On the other hand, a title like *Latitudinal differences in diapause response in populations of ... in Canada*, for example, shows the essential elements. *Divergence among populations of ... demonstrated by morphology and nuclear DNA* highlights significant methodology. *First successful establishment of a laboratory colony of ... for research on ...* establishes the significance of the findings.

Different priorities apply for publication outside scientific journals. There, cuteness or extreme brevity, for example, may be an advantage.

The title is more critical than it was years ago, now that many more papers are published and their contents are tracked digitally. The early working title often needs no adjustment, but after the manuscript has been drafted the title should be improved, if necessary, by subjecting it to as much revision and editorial diligence as any part of the text. Most modern titles are considered too in the context of the keywords that, along with titles and abstracts, are used by internet search engines and reference databases to index content (see *Final title, and keywords*).

Writing

The keys to writing—in addition to scientific rigour—are purpose, structure, correct use of language, and attention to detail. Moreover, optimal content and expression come from determining exactly how to deliver the findings as clearly as possible, so that only relevant results are included.

Tables and illustrations are the quickest way to communicate large amounts of information that would be complicated or space-consuming to explain in text (see *Tables*; and *Figures*).

All elements of a manuscript should be organized in sequence, structuring them from the general to the specific. This sequence is preferred for the paper as a whole and for each of its parts, including sections, subsections, paragraphs, and sentences. Such a pattern helps the reader to follow the content. Planning the research itself will have entailed a similar progression.

More broadly, the manuscript profits from consistency of structure. For example, the ranking of headings and subheadings should be easy to follow, and the format of tables that display comparable data should be the same. For some purposes, setting off selected elements in a different font, or at a slightly smaller size (as for the illustrative examples included in this article), may facilitate reading and reference. Some journal editors are willing to entertain such possibilities.

Contents

Sections imposed by the journal provide initial structure to the manuscript. Within that framework, the sequence of information and ideas serves to lead the reader towards the conclusions, a goal that can be assisted by drafting a relatively detailed list of contents at the outset. Finding the most effective sequence continues in conjunction with the actual writing; and ideas or phrases can be jotted down at any time to incorporate later.

At a minimum, the components are as follows:

Abstract: An extremely tightly worded statement of the content. A few journals call for a closing summary instead, which may be more detailed. The abstract (or summary) gives specific information and findings, and does not simply outline what subject has been addressed or what studies have been done. I have sometimes written a tentative abstract—to be revised later—before the rest of the text. Stating the intended messages explicitly in this way, and confirming that they are logical, may help in developing the paper.

Introduction: A road map for where the paper is going, so that the reader sees the point. Typically, the introduction also includes well-chosen existing facts and conclusions to provide context for the new material. Once, I was dissatisfied with multiple drafts of an Introduction. However, the solution proved to be merely the addition of one pivotal set-up sentence to help readers understand the purpose of the paper and why the particular approach had been chosen.

Methods: An account presented in sufficient detail to allow others to repeat the work if they wish. This section calls for as much diligence as the rest of the manuscript, even though some authors familiar with the methods, and focussing on their findings, give it cursory treatment.

Main sections: Evidence and interpretations, organized into as many sections and subsections as necessary to ensure that the various components and findings of the work are clear, and presented in the optimal order to prevent the reader from getting lost. Carefully chosen and well-produced tables and figures support the presentation.

Discussion: More extensive interpretation of the findings, typically including substantial, and sometimes complete, reference to previous relevant work.

Conclusions: Pithy statements echoing the main findings, included separately or at the end of the discussion. Long complex works are strengthened by ending each main section with conclusions, in addition to a final synthesis.

References: An accurate list of cited publications (see *References*).

Writing paragraphs

Each paragraph should contain only one main idea, and open with an orientation to that idea for the reader, followed by a smooth sequence of pertinent materials and clear linkages among them. In practice, however, potential paragraphs can be split up or combined to avoid paragraphs that are particularly long or short.

A *topic sentence* begins each paragraph. It introduces the subject to be addressed and foreshadows the conclusion, but is not unduly elaborate.

For example, a typical topic sentence might read *A topic sentence begins each paragraph*!

The topic sentence is followed by relevant and well-organized facts and ideas that underpin the subject of the paragraph. These particulars show clearly how the paragraph fits in to the "story" being told in that part of the manuscript. Likewise, the sequence of paragraphs carries the larger story to the reader. When the relevance of the information is made readily visible in this way, even readers new to the subject will not become bogged down in details that they cannot associate with the main theme of the section.

Normally, a *concluding sentence* (or clause) then reinforces the topic of the paragraph.

For example, *The two main types of cold hardiness customarily recognized differ in whether or not water in the insect freezes* [Topic sentence]. *Freezing-resistant species supercool, the body fluids remaining unfrozen even well below 0 °C, whereas freezing-tolerant species survive the formation of extracellular ice…. However, some species that supercool are killed at temperatures above the supercooling point, …* [Main body of the paragraph]. *Nevertheless, the two customary categories serve to highlight the major significance of ice formation* [Concluding sentence].

The concluding sentence may be used to provide an appropriate bridge to the next paragraph, especially if a long paragraph has been broken up into smaller sections to make it less dense.

Writing sentences

Important elements of each sentence are order, grammar, style, clarity, conciseness, word use, and spelling. Hundreds of decisions are required throughout the manuscript to ensure that the language is impeccably correct, and to make the text as lucid and concise as possible whilst giving the proper nuances of meaning. Finding the best wording for the purpose, and in context, takes concentrated effort.

Order

The beginning of a sentence serves for orientation, in the same way as a topic sentence orientates the reader to a paragraph. Consequently, a sentence usually opens with the main subject, a general point, or suitable set-up words.

For example, when *writing* is intended to be the focus of a sentence, *Comprehension is improved by writing clearly* is not optimal because it defers the theme to the end, and the first word seen by the reader is *comprehension*. Changing the order would place the emphasis on writing: *Clear writing improves comprehension*.

That sentence demonstrates too how several requirements have to be kept in mind simultaneously.

For example, putting the subject *writing* ahead of *comprehension* by a simple change to *Writing clearly improves comprehension* is not suitable because the reader may slow down if the beginning of the sentence appears to read *Writing/ clearly improves…* rather than *Writing clearly/ improves…*

It may be useful to separate the main theme as a short set-up sentence for effect, rather than use it as the lead-in to a longer sentence. It then serves as the topic sentence of a paragraph.

For example, *Larvae cause extensive damage, and trees can be defoliated completely because each caterpillar consumes about one square metre of leaves as it develops* is less forceful than *Larvae cause extensive damage. Trees can be defoliated completely because each caterpillar consumes about one square metre of leaves as it develops.*

In parallel with an early focus on the main subject, sentences are easier to assimilate when the subject is relatively simple, and complex components are placed later in the sentence.

For example, *The milkweed leaf beetle, the milkweed longhorn beetle, the larva of the milkweed tussock moth, large and small milkweed bugs, the viceroy butterfly, and the monarch butterfly are orange-and-black species associated with milkweed in Canada* makes the reader work harder than *Many orange-and-black species are associated with milkweed in Canada; they include the milkweed leaf beetle, the milkweed longhorn beetle, the larva of the milkweed tussock moth, large and small milkweed bugs, the viceroy butterfly, and the monarch butterfly*.

In the same way, a list is easier to read when the most complex item is placed last.

For example, *Adaptations include timing of the life cycle, cold hardiness, and dormancy* is slightly less smooth than *Adaptations include dormancy, cold hardiness, and timing of the life cycle*.

Grammar

Proper use of English delivers the message without distracting the reader. The many available guides to the language include the well-indexed book by Swan (2016), which is also available in an online edition.

However, in my experience, many of the elementary errors emphasized in grammar books are relatively infrequent in journal submissions.

For example, non-existent words like *irregardless*, erroneous placement of apostrophes, use of adjectives as adverbs, and confusion of *its* with *it's* and *there* with *their* (or *they're*) are seldom seen.

More than a dozen of the commonest kinds of grammatical errors in submitted manuscripts are demonstrated in simple sentences in Table 3. Many of them are mentioned again below. Mistakes occur chiefly in complex sentences, where the errors are less obvious because components of the sentence are well separated.

Table 3. Grammatical errors that are relatively common in submitted papers*. The most frequent ones are listed first.

Type of error	*Explanation, example, and correction*
Dangling modifiers	Clauses misplaced and qualify the wrong thing: *After walking on to the leaf, the researcher found that the larvae took more than an hour to start feeding.* Correction: *The larvae took more than an hour to start feeding after they walked on to the leaf.*
Pronoun-antecedent disagreements	Pronouns not matching the nouns they replace in number or gender: *The colony of ants always guarded their nest.* Correction: *The colony of ants always guarded its nest.* [However, because there is no neutral singular pronoun in English, *they*, *them*, or *their* might be applied to a single person when the writer does not wish to identify gender.]
Mixed (or misused) tenses	Tenses not correct or consistent, as required by the content: *The butterflies flew away whenever a bird comes to the flower.* Correction: *The butterflies flew away whenever a bird came to the flower.*
Missing hyphens	Compound adjectives not hyphenated, and hence apply separately: *Pale bodied adults* Correction: *Pale-bodied adults*
Incorrect use of commas	Commas missing, although required before an introductory word or phrase, to separate elements in a series, to isolate independent clauses, and elsewhere: *At an air temperature of 9°C individuals were inactive.* Correction: *At an air temperature of 9°C, individuals were inactive.* [However, the comma can be avoided by rewording to *Individuals were inactive at an air temperature of 9°C*.] Commas unnecessary and obtrusive: *It was 10°C, when the mayflies emerged, from the river.* Correction: *It was 10°C when the mayflies emerged from the river.*
Lack of parallelism in lists	Listed items constructed differently: *The research aims to discover why the population of this pest is increasing, and finding control methods.* Correction: *The research aims to discover why the population of this pest is increasing, and to find control methods.*
Misplaced modifiers	Words misplaced and qualifying the wrong thing: *The pest almost survived for the entire month.* Correction: *The pest survived for almost the entire month.*
Incorrect use of *which* and *that*	*That* instead of *which* introducing a non-essential (non-restrictive) clause [can be removed without ruining the meaning of the rest of the sentence]: *The adult tick, that has eight legs, feeds on mammalian hosts.* Correction: *The adult tick, which has eight legs, feeds on mammalian hosts.* *Which* instead of *that* is used for a defining (restrictive) clause [with information essential to the meaning of the sentence]: *Any tick which finds a host tries to feed on it.* Correction: *Any tick that finds a host tries to feed on it.*
Comma splices	Two sentences separated only by a comma, when a period, conjunction, present participle, or semicolon is required instead: *Species A has defined pronotal punctuation, in species B pronotal punctuation is indistinct or absent.* Correction: *Species A has defined pronotal punctuation. In species B pronotal punctuation is indistinct or absent*

Continued...

Table 3 (*continued)*. Grammatical errors that are relatively common in submitted papers.

Type of error	*Explanation, example, and correction*
Incorrect use of *between* and *among*	*Between* instead of *among* used to compare more than two separate things: *The differences between the 25 replicates are surprising*. Correction: *The differences among the 25 replicates are surprising.* *Among* instead of *between* used to compare only two things: *There are considerable differences in foodplants among species A and species B.* Correction: *There are considerable differences in foodplants between species A and species B* [*Among* is also used to refer to things that are not clearly separated, as part of a group or mass.]
Incorrect use of *farther* and *further*	*Further* applied to physical distance: *The large larva travelled further than the small one.* Correction: *The large larva travelled farther than the small one.* [But some authorities use *further* for physical distances too] *Farther* applied to figurative and non-physical distances: *Nothing could be farther from the truth.* Correction: *Nothing could be further from the truth.*
Incorrect use of *less* and *fewer*	*Fewer* instead of *less* used for things that are not quantifiable: *Fewer DNA* Correction: *Less DNA* *Less* instead of *fewer* used for things that are quantifiable: *Less moths* Correction: *Fewer moths*
Incomplete comparisons	Comparison without reference to what is being compared: *These stag beetles are much larger.* Correction: *These stag beetles are much larger than the ones from area B.*
Illogical apposition	Words with opposite senses linked inappropriately: *The table includes some omissions; An accurate estimate.* Corrections: *Some data have been omitted from the table; A reliable estimate.*

*This summary is based on my professional experience, reviewing a few hundred papers for dozens of different journals, reviewing and editing more than 100 chapters or papers for books and symposium proceedings, editing various newsletters, and reading many published papers in which errors remained after review.

Clauses that are linked to the wrong subject (dangling modifiers and other constructions) confuse or misdirect the reader. They are particularly common.

For example, the linkage is faulty in *The species cannot survive in northern Ontario because the summer is too short to complete development.* This construction means that the summer, rather than the insect, cannot complete development.

Likewise, the subject must be stated rather than implied in follow-up sentences.

For example, the logic is faulty in *All text citations must accord with entries in the list of references. This task can be completed by a page-by-page review of the manuscript.* The noun *task* should not be the subject of the second sentence, because the intended action is matching. The link can be made, for example, by a change to *This requirement can be met by…*, or *Citations and references can be matched by...*

Pronouns that do not agree with their antecedents are common too.

For example, *Each insect was anaesthetized with carbon dioxide, marked according to the subarea of collection, and released in the centre of the whole area after their recovery.* [singular *each* ... plural *their*]

Singular subjects frequently have plural verbs (or vice versa), breaking an elementary rule of grammar.

For example, *The number of individuals attacked by each species of parasitoid, in relation to size and season, are shown in the top part of Figure 2.* [singular *number* … plural *are*]

Tenses that vary within a sentence or paragraph distract readers or introduce ambiguity. Switching from one tense to another should also be avoided in longer sections, even if the uses are infrequent.

For example, *Samples were sorted by flotation in sucrose solution according to the method of .... However, it proves possible before this treatment to sieve out fine particles through a 3-mm mesh without the loss of any larvae.* [*proved possible*—not *proves possible*—would both keep the tense consistent and confirm that the samples were in fact sieved.]

Omitting hyphens from compound adjectives can be amusing as well as misleading, because it causes the adjectives to apply separately.

For example, *vestigial winged Drosophila* [vestigial *Drosophila* with wings] must be extraordinarily difficult to detect, unlike the experimental population of *vestigial-winged Drosophila* [*Drosophila* with vestigial wings].

Other common errors listed in Table 3 include: too few or too many commas; sentences separated only by commas; lists containing items that are structured in more than one way; words that are misplaced or used inappropriately; and the wrong choices in pairs of words (*which*/*that*, *between*/*among*, *farther*/*further*, and *less*/*fewer*).

A few language rules are applied differently by different grammarians, however.

For example, the split infinitive [*to boldly go*] was once prohibited by leading authorities, a rule strictly enforced by the school system. The prohibition reflected a belief that the "infinitive" form of an English verb includes the preposition *to* and so is a single entity, but many others deny that claim.

Style

Limitations of style in scientific papers are at least as prevalent as grammatical errors. Some of them reflect the importance of sentence order, as noted above.

Consider the following sentences about faunal differences from one area to another:

Species lists that are complete are normally necessary to assess geographical differences reliably.

Reliable assessment of geographical differences normally requires species lists that are complete. [Main theme moved forward from the end of the sentence, and the awkward *are ...are,* and the floating adverb *reliably,* eliminated.]

Reliable assessment of geographical differences requires species lists that are complete. [Sentence simplified by deleting *normally*, deemed unnecessary here.]

Geographical differences can be established with certainty only when species lists are complete. [The main subject, *Geographical differences*, now opens the sentence, and the text is smoothed by replacing *assessment...requires* with the more direct construction *can be established*, and deleting *reliable* in favour of the more specific qualifier *with certainty*.]

All of these versions are grammatically correct and can be understood, but revised versions are easier to read and more precise.

Because style depends to some extent on author preferences, there is no single "best style". However, some style elements are universally regarded as suboptimal. Table 4 illustrates common deficiencies of this sort. Weaknesses of style in submitted manuscripts, as shown in Table 4, lead to unclear, rambling, or imprecise text. Long, dense, or convoluted sentences are difficult to read. Wordy expressions and vague or superfluous words clog the text. Complex constructions and lack of context slow down comprehension. Other examples and further explanations are included under *Clarity*, *Conciseness*, and *Word use* below.

A few elements of style depend on the journal.

For example, many journals omit the comma before the final item in a list (before *and*) [... *production of cryoprotectants, adjustments of water content and selection of sheltered habitats suggest...*]; but those following the standard of Oxford University include it. That comma ["the Oxford comma"] is often helpful, and sometimes essential, to an understanding of the sentence […*production of cryoprotectants, adjustments of water content, and selection of sheltered habitats suggest…*].

Clarity

Sentences are clear when they focus on the main points. Asides or unimportant details, and unnecessary qualifiers, add little value and can be moved to a subsequent sentence or deleted. Alternatively, ancillary material can be provided in a footnote when appropriate.

When all of the content has equal importance, a sentence can be clarified by placing part of the information into a separate sentence.

For example, a long sentence may be disjointed: *Species A, which is a predator of bugs, beetles, and other insects, develops more rapidly than species B, which consumes xylem fluid very low in nutrients, because its food is more nourishing*.

It is easier to understand when split into two sentences: *Species A develops more rapidly than species B because its food is more nourishing. Species A is a predator of bugs, beetles, and other insects, whereas the xylem fluid consumed by species B is very low in nutrients.*

Extremely long sentences replete with information make it likely that the reader will get lost as the sentence travels inexorably onwards by means of multiple commas, parentheses, dashes, colons, and semicolons. The writer already knows the connection between the parts and does not realize what might happen to the reader.

Information-rich text can be simplified by ensuring that material is in a logical order, and by breaking it up into successive sentences that lead into one another. In particular, clarity is improved by using a short topic sentence to set the scene, followed by subsequent short sentences that deal with only one item each.

For example, a sentence with many facts seems endless: *Insect adaptations to drought include [many lines of text that introduce and explain a long list of adaptations].*

It can be replaced by—Sentence 1: *Many different adaptations allow insects to survive drought.* Sentence 2: *[Adaptation 1 and its explanation.]* Sentence 3: *[Adaptation 2 and its explanation.]*; and so on.

This process of simplification can be repeated to break up the text further if necessary—Sentence 1: *Many kinds of ecological and physiological adaptations allow insects to survive drought.* Sentence 2: *The most common ecological adaptation is [Adaptation 1.]* Sentence 3: *[Explanation of adaptation 1.]* Sentence 4: *[Adaptation 2.]*; and so on.

Table 4. Deficiencies of style that are relatively common in submitted papers*.
The most frequent ones are listed first.

Type of deficiency	*Explanation, example, and improvement*
Paragraphs and sentences too long	Dense text is difficult to understand: *An example* would be too long for this table, but further details are provided in the text. [Uncomfortably long paragraphs and sentences should be split so that each one includes only a single main point. If there are many required details, they can be divided among several paragraphs or sentences.]
Sentences disjointed or convoluted	Reading is hindered when the text does not flow smoothly, because content is in the wrong order or the phrasing is awkward: *In this family—the Ichneumonidae—the parasitoids, attacking many kinds of insects, are widespread in the environments of fields and woodlands.* Improvement: *Parasitoids of the family Ichneumonidae attack many kinds of insects and are widely distributed in fields and woodlands.*
Few core words and many ancillary ones	Sentences in which few words contribute real content (relative to grammatical linkages and general words) are bothersome to read because they are insufficiently concise: *After this method had been used for an initial analysis, the data were transformed in the usual way and then plotted on the two axes shown in Figure 3 in order to indicate the effect of temperature and pressure on flight behaviour.* Improvement: *After initial analysis, the transformed data were plotted to show how temperature and pressure affect flight behaviour (Figure 3).*
Flabby expressions	Some commonly used expressions are superfluous or unnecessarily long: *It should be mentioned that…; An experiment was conducted to…; As is well known, …; Of course, …* Improvement: Delete all. *At this point in time; During the course of; By means of; In view of the fact that; In spite of the fact that; As a matter of fact.* Improvements: *Now*; *During* or *in*; *By*; *Because; Although*; *Indeed* (if needed at all)
Vague qualifiers	Some words are not precise enough: S*omewhat; rather*. Improvements: Delete or replace [Imprecise qualifiers, such as *some*, *many*, *partly*, and *normally* cannot always be avoided when data show complex outcomes, or if conclusions are not certain]
Unnecessary qualifiers	Qualifying a qualifier is usually superfluous: *Absolutely complete; Fully final; All finished.* Improvements: *Complete; Final; Finished*
Overuse of adverbs for action and events	Sentences with too many adverbs are intrusive: *The insect climbed the stem continuously and fed rapidly, but paused frequently.* Sample improvement: *The insect climbed the stem without interruption, and fed rapidly with frequent pauses.*
Use of *over* for *more than*	*Over* signifies a position above something, whereas *more than* shows that a number or quantity is exceeded: *There were over 100 moths.* Improvement: *There were more than 100 moths.* [This distinction was once an invariable rule of grammar, but—although it has been relaxed by some authorities—is still the preferred style.]

*This list is based on my professional experience (summarized in the footnote to Table 3).

These and other sentences can be linked by suitable starting or ending phrases.

For example, words at the end of a sentence might include: … *leading to two alternatives, ... can be classified as follows,* or *... for which several biochemical pathways have been identified.* Those at the beginning might include: *Moreover, In addition, Alternatively, In contrast*, or *Analyses aimed at [one of the group of items just introduced]*.

Sentences are only clear when readers understand their context. However, authors familiar with the subject, and with the usual specialist readership, might omit critical information or linkages. A reader who does not know enough to fill in context that has been assumed, or was provided only near the beginning of the paper, may misunderstand. Checking that the logic of the sentence is complete, and including helpful orientation, prevents misinterpretation.

For example, the sentence *These patterns of development mean that the species cannot survive in northern Ontario* may (depending on context) be easier to understand when expanded to *Therefore, the species cannot survive in northern Ontario because the summer is too short for the larvae to complete development before winter.*

Again, someone who reads *The caterpillars gradually darken as they grow, so that they are light brown in the third instar, and dark brown by the end of the final instar* might assume, if they were unfamiliar with this species, that the final instar is the fourth (as in some related species), whereas …*by the end of the final (fifth) instar* is unambiguous.

Conciseness

Thoughts should be expressed as simply as possible to deliver the main message. Multiple revisions might be required to prune each individual sentence. The famous quotation "I would have written a shorter letter, but I did not have the time" (variously attributed to Cicero, Blaise Pascal, and many later writers) encapsulates this reality.

Sentences can nearly always be condensed during initial drafting.

For example, *Samples that had been taken from each area were then sent to various specialists for identification* imparts only the information that *Samples from each area were identified by specialists.*

Consumption of leaves by larvae as they feed causes extensive damage can be shortened to *Larval feeding causes extensive damage.*

The first insects to emerge are males can, in most instances, be replaced by *Males emerge first.*

Superfluous content-free items often creep in as the writer tries to begin a sentence, and should be deleted. Other lengthy constructions have shorter equivalents.

For example, *It should be noted that…* and similar opening gambits add pointless words.

At the present time can be replaced by *now*, and *in view of the fact that* by *because*.

Other suboptimal expressions can be replaced by single words: *vast majority* (*most*), *all of* (*all*), *both of* (*both*), and *prior to* (*before*).

Table 4 shows further examples of these "flabby expressions".

Word use

Choosing the most effective words contributes to precision. Words that are regularly misused are listed in books on grammar, and should be employed with care.

Many of the general words that are used more or less interchangeably in ordinary speech vary in meaning in ways that matter for a paper.

For example, *Females were assessed* is less precise than *Adult females were examined*, which is less precise than *Adult females were dissected.*

The percentage of beetles carrying mites in sample 2 was estimated should read *The percentage of beetles…was calculated*, because the statistic came directly from the sample data.

On the other hand, *The percentage of the beetle population likely to be carrying mites was calculated* implies unwarranted precision, and should read *The percentage of the beetle population … was estimated*.

Words should be no more complex than necessary. However, reducing complexity is not the same as choosing a more precise word when appropriate.

Pretentious words are especially distracting. The tendency of bureaucrats to express themselves pretentiously has been parodied many times, but some scientists have adopted the habit in the mistaken belief that it appears scholarly.

For example, pretentious words include *conceptualize* rather than *plan*, *furnish* rather than *give* or *supply*, *orchestrate* rather than *arrange,* and *utilize* rather than *use*.

People are likely to stop reading when they encounter impenetrable prose: *Unremitting utilization of pretentious expressions categorically subjugates comprehensibility*. [Translation: *Pretentious writing is hard to understand.*]

Vague or otherwise unsuitable words should be replaced.

For example, there are better qualifiers than *about* (meaning *approximately*, not *close to*), *poor* (meaning *inadequate*, not *penniless*) and *meaningful* (provide an explanation).

Most meanings can be expressed correctly in many different ways, and the words are chosen by each author. Some choices are self-evident, but many are more difficult because there are so many shades of meaning in related English words. These subtle word choices depend on context, sentence structure, perspective, and other variables. Authors should use a thesaurus when required to find a word to express the appropriate nuance. Additional “wordsmithing” can be done during later revisions.

Spelling

Correct spelling is important. Anyone uncertain about a word should check with the applicable dictionary or style guide.

When there are alternative spellings or hyphenations, journal requirements may determine which forms are used. If authors can choose the form they prefer, it must be consistent throughout the manuscript.

For example, different journals use *characterize* or *characterise*, *behaviour* or *behavior*, *judgment* or *judgement*, *co-operate or cooperate*, and *pre-arranged or prearranged.*

References

Introduction

Information from cited references provides background information and confirms how the results or ideas in a submitted manuscript add to existing scientific understanding. Therefore, comprehensive knowledge of the literature is vital, a goal that demands great diligence.

It is unwise to stray into new fields without knowing about previous work. Two of my colleagues were crestfallen when their paper in a different field was forcibly rejected, but realized too late that such a fate might have been avoided by greater familiarity with relevant publications.

Even some papers written by specialists familiar with the literature omit relevant references, however. In particular, information published in journals from other countries or in other languages is nearly always under-represented, and recent publications are cited less than older ones that are more detailed or more applicable. These neglected papers would have provided further useful context.

Reviewing the literature thoroughly not only serves authors well for particular manuscripts, but also strengthens all of their future work. My own experience illustrates these benefits. The search for employment early in my career led me to investigate many subjects: the ecology of various insect groups and their natural enemies; seasonal and other adaptations; taxonomy of immature stages; linkages with agricultural systems; and the conditions and faunas in different regions and habitats. The resulting literature reviews widened my entomological knowledge and exposed me to contributions from experts in many disciplines. They provided information and insights that were invaluable for future manuscripts and syntheses. Similar benefits accrue during graduate studies from exposure to a range of subjects, rather than narrow focus on a single speciality or technique.

Reference searches

Adequate reviews of the literature depend on effective searches. Search methods and strategies have been evaluated especially as a basis for statistical meta-analysis (e.g., White 2009, and references cited there), although not always with perspectives that pertain to entomology.

Several approaches are recognized. Checking citations in other people's papers (sometimes called "footnote chasing") is especially useful because those papers have already been screened for relevance. Browsing library shelves in selected subject areas is a less efficient but sometimes profitable strategy. Helpful information may come to light from communication with colleagues, and from reviewing citation indexes to identify highly cited papers. I used a further strategy during some particularly broad reviews (e.g., Danks 1981, 2006): checking content in every issue of journals that were especially likely to contain relevant papers.

Increasingly, however, comprehensive searches are done using internet search engines and reference databases, which have replaced the printed bibliographic indexes and weekly computer-assisted title searches available earlier. For example, BIOSIS, which began in 1926, is now part of the Web of Science (Clarivate 2024a).

Millions of titles and keywords can now be searched. Google Scholar (Google Scholar 2024) covers many books as well as journals, and is the most effective search engine. Large reference databases that focus on papers published in scientific journals include Web of

Science (just noted), EBSCO (EBSCO 2024), and Scopus (Elsevier 2024). Available academic research databases are summarized by Paperpile (2024).

The success of searches depends on choosing the correct search terms, conducting multiple searches with different terms, and searching as many potential sources as possible.

Ways to choose search terms and optimize their results have been reviewed many times. For an overview and further references see Reed and Baxter (2009), for example. The effectiveness of choices depends partly on in-depth knowledge of the general subject area, reinforcing the conclusion that preparing a major review requires considerable familiarity with existing information.

Before a search begins, the scope of the topic should be established as narrowly as possible. Otherwise, searching finds large numbers of irrelevant titles.

Search terms are developed by considering words about the topic, their synonyms (to add terms), exclusions (to limit them: for example, searching without excluding unwanted taxa will dilute the results), and logical combinations (to increase the power of the search). Boolean operators (notably AND, OR, and NOT, although AROUND is useful too) help to limit or expand the search. Examining word clouds (see *Final title, and keywords* below) may help the process. Most database programs enable searches by title, author, date range, exact phrase, and other categories.

Preliminary searches show how the choice of terms might be improved. In particular, searches that are too narrow miss relevant titles, and searches that are too broad overwhelm the results with false positives. As the search proceeds, its effectiveness can be evaluated by examining the results for subtopics that are already particularly well known to the searcher. Learning what sorts of information have been missed suggests ways to improve the search.

A thorough search relies on multiple sources. Accessible content differs among the various custom databases, and from the web, and different search engines use different algorithms to organize the search. Several demonstrations of the need for multiple sources come from the medical and social sciences literature. For example, Wanyama et al. (2022) showed that only 7% of titles in one search were retrieved by all of the three major databases (Web of Life, Scopus, and EBSCO).

Some publications (especially from the period before keywords) are not fully indexed for web and reference-database searches. Most of them can be found by checking citations in other papers, but (depending on the topic) searching indexes in hard-copy compilations may also be needed.

A final decision is when to stop searching. Ideally, the search ends when no new titles are discovered. In practice, however, a typical search might be ended when additional papers no longer show anything new, and when the few minor papers that continue to be found merely confirm existing information.

Managing references

Managing large numbers of citations requires organization. In the past, storing and tracking the information depended on laborious annotation of individual cards that were filed in index-card holders. Notes might be written, abstracts copied, or photocopies pasted on some of the index cards. Now, reference-management software organizes records automatically and generates citations for individual researchers. Abstracts and other information can be included with each reference.

There are many reference-management programs. Although most of the free ones (including free versions of more capable programs) are adequate for one or a few projects, they are not powerful enough, or allow too small of a personal database, for career researchers.

Some common programs are characterized in Technical University of Munich University Library (2022). Widely used software includes EndNote (2024), which is part of Clarivate, Mendeley (Mendeley 2024), which is owned by Elsevier, and Zotero (Zotero 2024). One relatively common one (RefWorks) is available only by institutional subscription.

The better programs work on Windows or Mac platforms, interface with Microsoft Word and other common software, and output a range of citation styles. Most of them are web-based. The value of these databases is enhanced if every reference is checked when first consulted, to ensure that it has been transcribed exactly. Author, date, title, journal name, volume, issue, and page numbers should all be verified, because many citations, from various sources, include minor errors.

Each journal imposes its own citation style. However, following a given style is simplified if the chosen reference-management software can generate the list of cited references in the appropriate format.

Using references

The titles and abstracts of references suggest which papers are worth consulting for a particular project. Publications that prove not to be useful (or not useful for the current topic) are best included in the database but annotated accordingly, to prevent later duplication of effort.

When a digital or hard copy of a paper has been obtained, it should be read carefully and key items noted. The section above on *Planning for literature reviews* recommends a way to ensure that information from references remains accessible for any given manuscript.

It seems self-evident that authors should read the whole of any paper they cite. My reading and refereeing suggest that instead some authors base a citation only on information in the title or abstract, without consulting the original paper, although titles and abstracts alone can be misinterpreted.

More than once, too, I discovered that authors of published papers had cited three or four references that were identical in number and sequence to those used in one of my own reviews—but those references were cited to validate points that were not identical to mine, though almost the same. Some of the citations did not even support the points made, suggesting that they might simply have been copied from my reviews without consulting the original papers!

After the manuscript has been prepared, references and their citations should be one focus of final meticulous verifications (see *Last checks*).

Tables

Tables are a concise way of presenting results, even complex ones, and should be as clear and relevant as possible. Ways to produce tables suitable for publication are recommended in Table 5.

Table 5. Recommendations for tables intended to be published.

Context
Use concise and informative captions
Make each table self-contained for ease of reference
Number
Optimize the number of tables
Delete tables and data that are not essential
Consider transferring the content of simple tables to the text
Delete data repeated in another table
Present data in a table or a figure, not both
Divide a table that is unwieldy
Basic elements
Have table elements read up to down, not across
Use concise, descriptive column titles
Reduce the number of columns by combining if possible
Layout
Avoid elaborate layouts
Use horizontal rules only
Separate subsections with horizontal rules only if necessary (may be useful for text entries)
Put adequate space between cells
Use a legible font and size
Treatment of data
Divide data into easily grasped categories
Put only one value in each cell
Include units
Include error values or statistical significance for most numerical data
Minimize the number of decimal places
Include number of samples if relevant
Make word entries concise
Make entries consistent (e.g., data, formats, parts of speech)
Additions and footnotes
Limit use of abbreviations and symbols
Explain any abbreviation or symbol (by a footnote)
Footnote any abbreviation or symbol in the same table
Cite the reference for any entry from another study

The caption is the first thing the reader sees, so great care should be taken to make it informative and concise. Readers must also be able to interpret all table entries without having to refer to the text, so simple and accurate column headings are particularly important.

Using more tables than necessary encumbers a paper. A table should be used only when it is the best way to present the findings. Indeed, some content can be presented more efficiently in the text than in a table that has few entries. In addition, no data in a table should duplicate data in a figure or another table.

Tables can show information in great detail. However, they aim to show *relevant* findings, not the totality of the available data. Nevertheless, some researchers are prone to include all of their data of every kind, multiplying the number of tables. This habit is potentially useful only for other products with different roles or lower constraints, including some student theses.

Moreover, individual tables that include more detail than is needed to illustrate the points being made are potentially confusing. A few of the entries in some tables do not even relate to messages in the text, a defect easily overlooked by people who are used to working with "data sets" from different experiments.

The main requirement for all tables is clarity. Column headings should contain as few words as possible, and tables should contain as few columns as possible. Columns can sometimes be eliminated by combining many small categories into one (e.g., *less than* or *more than*), provided the analysis of the data reveals that they add little useful content. Tables that are unwieldy, or will exceed the width of a printed page, are easier to grasp if they can be divided into two smaller tables.

Likewise, the overall layout of a table should be as simple as possible. Elaborate tables are difficult to follow[2]. The font for text should be large and plain enough to be readily legible. Placing adequate space between rows and columns enhances legibility.

Data are easier to understand if they are grouped into logical sets, with clear units, and error values when appropriate. The number of decimal places shown should be the smallest that both suits the purpose of the table and does not exceed the precision of the data. The number of samples need be listed only if informative, and when it is both large and similar among treatments can be included under Methods. Entries consisting of words rather than data should be reviewed several times to condense them as much as possible.

Consistency is particularly important. Differences in the parts of speech in column headings, and in the nature of data or text in table entries, are especially noticeable, but nevertheless surprisingly common in submitted manuscripts. They distract the reader and make headings and entries hard to assimilate, which is exactly contrary to the point of a table, to summarize and illustrate content succinctly.

For example, the headings *Individuals* and *No. of responses* should have been replaced by *No. of individuals* and *No. of responses*. The entries *Rapid development* and *Cold hardy* should have read *Rapid development* and *Cold hardiness*.

Footnotes and abbreviations tend to clutter a table and make it more difficult to understand. This pitfall can be avoided by using full terms rather than abbreviations when possible, and by footnoting only information that is critical. However, sources for information taken from other studies should always be cited in table entries or footnotes.

To meet the criterion that a table should be understandable without reference to the text, any abbreviation that is used should be footnoted in the same table. The reader should not have to remember the meaning of an abbreviation, nor look for it elsewhere.

The many items listed in Table 5 and discussed above show that constructing and finalizing a table takes diligent work. Maximum effectiveness stems from repeated refinement and streamlining—just as for the text.

Finally, all table entries (including column headings) should be checked against the text to ensure that the rationale and detail are correct, and that the message being delivered is clear. These checks may have to be repeated if either text or table has been significantly revised.

Figures

Some types of information are best displayed as figures (illustrations), which may summarize data more efficiently than tables or text, and may be easier to understand. The adage “A picture is worth a thousand words” reflects these abilities.

Figures can be adapted precisely, in many ways, by the digital tools available on personal computers. More and more journals accept various kinds of digital images with formats that were not available in the past. Reproductions in colour were once prohibitively expensive, but now they are commonplace, especially for online publications. New terminology has arisen, including “infographic”, which simply means a visual representation of information.

The many kinds of figures that can be prepared for publication, and their main features, are listed in Table 6.

---

[2] A parallel problem arises in some Power Point presentations. That software includes elaborate templates full of coloured distractions that may tempt presenters to use eye-catching tables that are hard to read.

Table 6. The types of figures used in scientific papers.

Type of figure	*Description*
Image	Visual representation of an object
Photograph	Likeness obtained through light, using a film or digital camera
Reproduction	Copy of original or part of it (sometimes adjusted to enhance clarity)
Modification	Photographic image changed to remove distractions or augment understanding
Synthesis	Image created from other photographs
Light scan	Likeness obtained by a scanner that converts light into digital data
Flat scan	Two-dimensional image of a flat object
3D scan	Three-dimensional image generated by computer integration of successive flat scans
Other scan	Image generated other than by light. Examples are:
Electron micrograph	Transmitted or reflected electrons (and resulting X-rays) visualized on film or digital media
X-ray	Transmitted X-rays visualized on film or digital media
Ultrasound	High-frequency-sound-wave echoes integrated by computer to form an image
Art	Drawing produced through creative and artistic skill
Realistic drawing	Copy of an object, usually one difficult to photograph or scan
Creative drawing	Artistic interpretation of a real, hypothesized, or imagined object
Diagram	[Other classifications for diagrams differ from the system used below]
Map	Diagrammatic representation of an area
Representation	Semi-realistic drawing of an object or structure. Examples are:
Sketch	Drawing that approximates the main elements
Exploded view	Components shown slightly separated for clarity
Schematic	Simplified diagram using symbols (or words), and lines. Examples are:
Flow chart	Diagram showing a sequence of activities or happenings
Chemical equation	Representation of a chemical reaction
Chart or data plot	Data set represented graphically [sometimes termed a diagram too]
Many types	See Figures 1–14 and captions for examples

Photographic images can be reproduced untouched (apart from cropping), but digital formats allow the quality of originals to be enhanced. Hue, brightness, saturation, sharpness, and other features can be adjusted with powerful image-editing software. Photoshop (Adobe 2024) and the considerably less expensive Affinity Photo (Serif 2024) are two of the most powerful. Labels can also be placed directly on an image.

Content can also be changed. For example, photographs can be simplified by deleting elements. Structures that are damaged or only partly shown can be digitally repaired to reduce distractions or permit a preferred orientation. Illustrations can even be synthesized by combining several different digital images.

However, attempts to make an image easier to understand can distort as well as enhance it. Great care is needed to ensure that a synthesized or highly modified image is used only to show fully established facts more clearly, and not to promote a theorized or hoped-for result. Even so, although such constructs are useful for some sorts of publication, they may show "falsified" data. Consequently, for a journal paper, a diagrammatic representation or a schematic (see Table 6) is nearly always more appropriate than a created image.

Scanners and other tools use energy such as light, electrons, Xrays, or ultrasound to generate images of various kinds, including images integrated digitally from multiple scans. False colours can be added to clarify some of the outputs, coding them by computer. Although electron micrographs (examining structures smaller than visible wavelengths) are generated only in black and white, ways have been found to introduce colour to these images too. In all instances, the meaning of the colours must be shown in the figure.

Artworks are produced through copying and creative skills. However, artistic impressions may be inaccurate or misleading in a scientific context, and should be treated with care. Realistic drawings are the preferred alternative, but are less common than in the past. New technologies have made it possible to show three-dimensional and other structures that once could not be adequately photographed, and were reproduced with drawings.

Diagrams are commonly used to map geographical areas, explain structures, and show temporal sequences or interrelationships. Cartography and many other techniques use standard conventions and symbols. Brief overviews of some of these conventions are available on the web, and may help in designing figures.

Charts (data plots) are the best way to summarize many kinds of numerical data. They can convey large quantities of data quickly by focussing on differences, relationships, or patterns. However, charting is a complex subject, and more than 100 different ways to visualize data have been established.

Graphing data not only helps to decide how the data might be prepared for publication, but also may reveal new insights. Unusual distributions, outliers, clusters, missing values, and unexpected gaps, for example, are sometimes easier to see on a chart than in the raw data.

Many papers and books have covered the topic of visualizing data. They explain how to illustrate written and oral scientific presentations, how to present data to business customers, and how to ensure the statistical validity of data analyses and the visualizations that stem from them (e.g., Tufte 2001; Briscoe 2012; Healy 2018; Evergreen 2019).

The types of charts most commonly used in scientific publications are characterized in Figures 1–14 and captions. These generalized figures are intended to illustrate the variety of patterns, and so exclude the conspicuous axis labels normally required, and limit the size of axis numbers.

Besides their visual impact, analyses and presentations of data must be correct statistically. Statistical requirements are beyond the scope of this treatment of publications, although a few constraints and misleading choices are mentioned below.

For example, changes in a variable for discrete categories such as treatments or taxa (categorical data) (e.g., Figures 1–4) cannot be plotted in a format restricted to showing how a variable changes according to unbroken elements such as time and size (continuous data) (e.g., Figures 5–8). Several types of data can be graphed as dots, but dot plots (Figure 4) are appropriate for categorical data and scatter plots (Figure 12) for continuous data. Showing percentages or other transformations rather than original data violates underlying assumptions in some kinds of plots.

Figures. 1–14. Types of charts commonly used for scientific data.

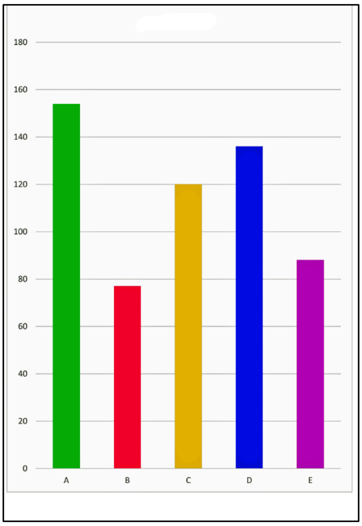

Figure 1. Bar chart.
Vertical or horizontal bars for each category of data, the length of each bar showing the value of the variable for that category.

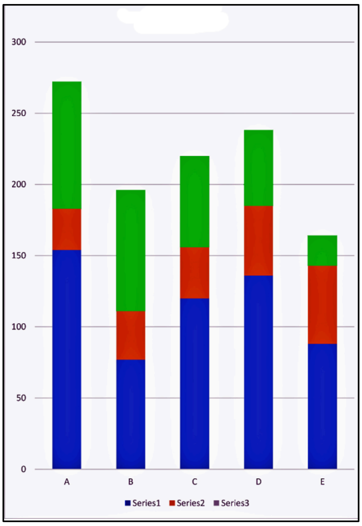

Figure 2. Stacked bar chart.
Bars divided into sections, showing the contribution of each component to a given category.

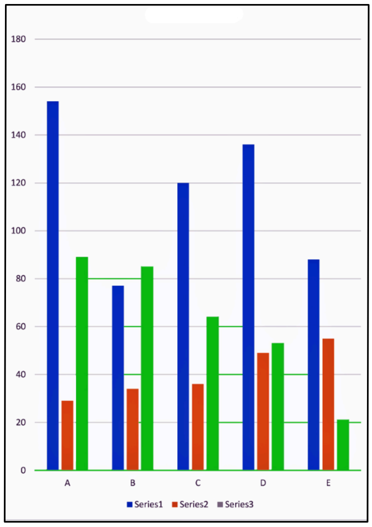

Figure 3. Grouped bar chart.
Bars in clusters, serving to compare values for different components within each category.

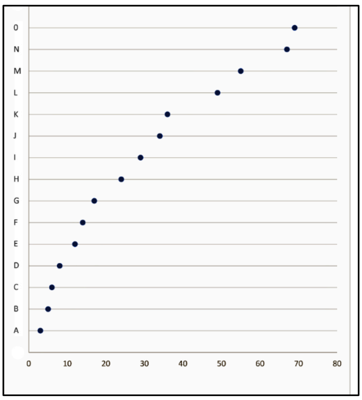

Figure 4. Dot plot.
Dots showing the distribution of a variable by plotting the values for each category. Several types exist; a simple ordered chart with categories on the y-axis is shown here.

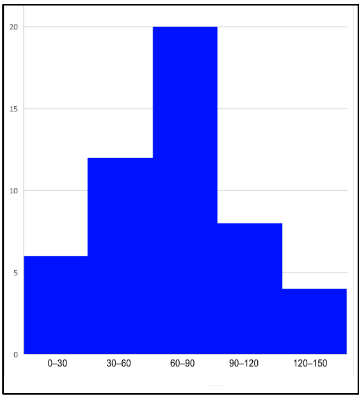

Figure 5. Histogram.
Bar chart showing the distribution of a variable over a continuous numeric range by aggregating data for each chosen interval into a single bar.

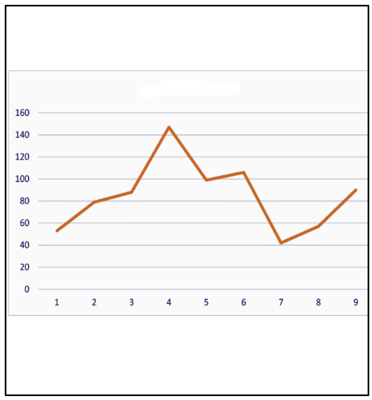

Figure 6. Line chart.
Plot showing patterns of change in the value of a continuous variable, by means of a line that connects the values (values are normally aggregated for chosen intervals, like a histogram).

Apart from their statistical validity, charts should demonstrate the data set to good advantage, be readily understood by the viewer, and impart the intended messages. For example, dot plots (which can summarize data without distortion) make outliers easy to see. Some plots are better than others for showing relatively small variations (e.g., line charts).

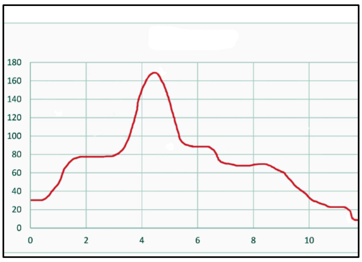

Figure 7. Density curve.
A line that integrates the contribution of each data point into a probability curve, thus smoothing out the data relative to a line chart or a histogram.

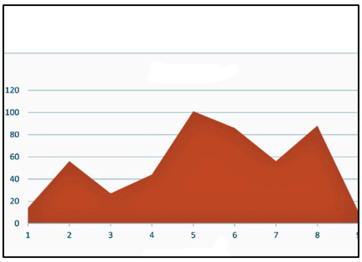

Figure 8. Area chart.
A line chart with shading between the data line and the baseline, emphasizing changes in the key variable.

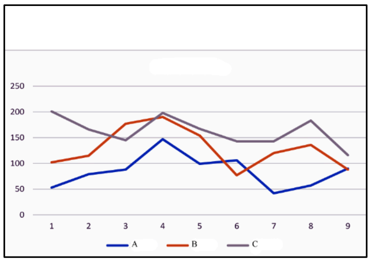

Figure 9. Multiple line chart.
Line charts (for each of several categories) plotted on the same axes, allowing direct comparisons.

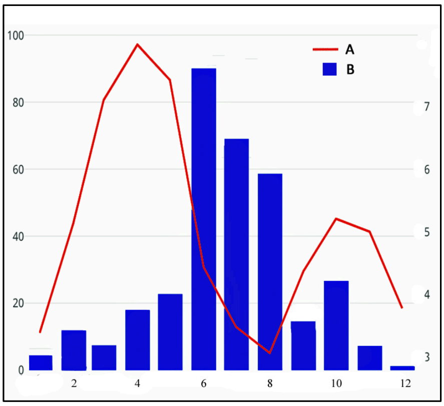

Figure 10 (left). Dual-axis chart.
Two different variables plotted on the same x-axis, demonstrating the relationship between the variables (typically using two different types of plots).

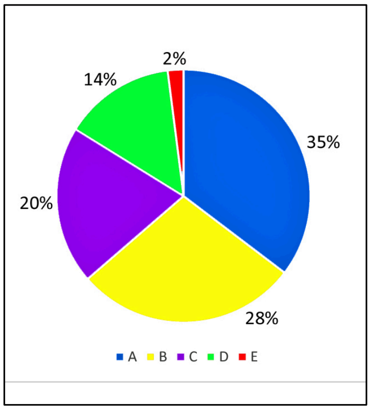

Figure 11 (right). Pie chart.
A circle divided into sectors (like a pie cut into slices), making it easy to visualize the contribution of each category.

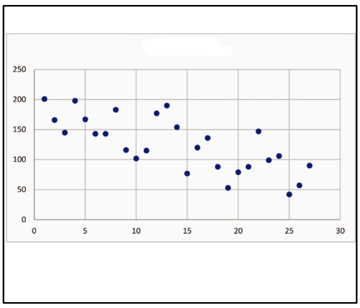

Figure 12. Scatter plot.
Scattered points derived by plotting two continuous variables against each other, showing the relationship between them.

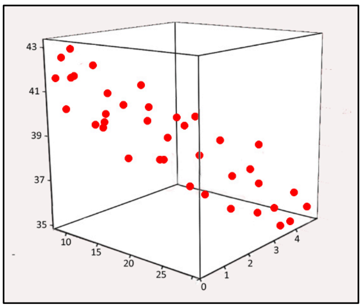

Figure 13. Complex scatter plot.
Scatter plot with data in more than two dimensions. Additional dimensions can be plotted on 3 axes (as here), and by coding data points by size, shape, or colour.

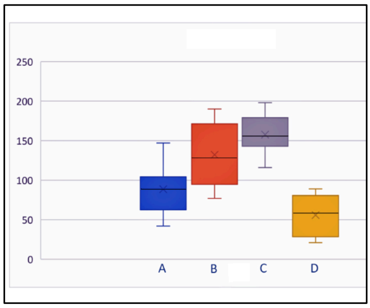

Figure 14. Box plot.
Boxes and lines (for each category) plotted across the range of a variable, typically showing the median, first and third quartiles, and upper and lower limits, thus summarizing the distribution of values. A similar plot using density curves is a *violin plot*.

Plotting most kinds of charts is now relatively easy with available software, even when data are complex. Even so, errors are more likely when charts are produced using digital "cookbooks" without proper understanding of the underlying relationships. Spreadsheet software, for example, generates some kinds of charts automatically from tabulated data: premium versions of Microsoft Excel (Microsoft 2024) are used especially widely. Most other spreadsheet programs are less costly, but not as powerful.

The clearest charts are simple and concise (e.g., vertical bars and pie charts, Figures 1–3, 11). Charts with "normal" (not extreme) aspect ratios, saturated (not pastel) colours, and colours or markings that are strikingly different, can be understood quickly. Marking numerical values on key components, such as the slices of pie charts, is helpful provided it does not clutter the figure.

Even so, simple formats are appropriate only when the number of elements is limited. When there are many bars, slices, or lines, for example, the entities are difficult to distinguish and compare. Dividing the data into subsets, or using a different type of chart, may be preferable.

Complex data can be visualized through multidimensional plots (e.g., Figure 13), although readers unfamiliar with a field may find them hard to grasp. However, increased understanding is unlikely but distortion of data likely when three-dimensional versions of two-dimensional data are created for visual appeal. For example, if histogram columns or pie-chart slices are simply copied from two into three dimensions, the differences between them are exaggerated.

Therefore, charts are misleading if they are unsuitable for the data or carelessly executed (e.g., Nguyen et al. 2021). A few sample errors are shown in Figures 15–18.

Figures 15–18. Examples of errors in the use of charts.

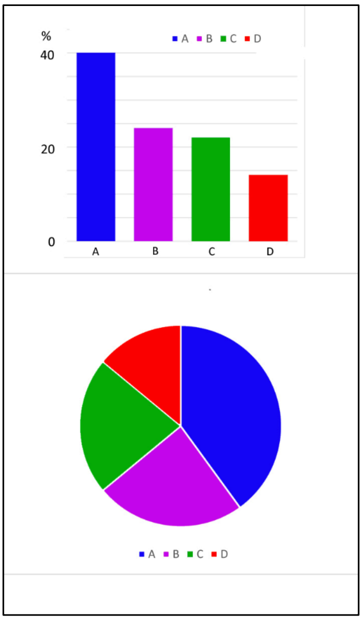

Figure 15. Erroneous use of a bar chart to show percentage composition (top), when a pie chart should be used instead (bottom). The incorrect type of chart makes the data difficult to assimilate.

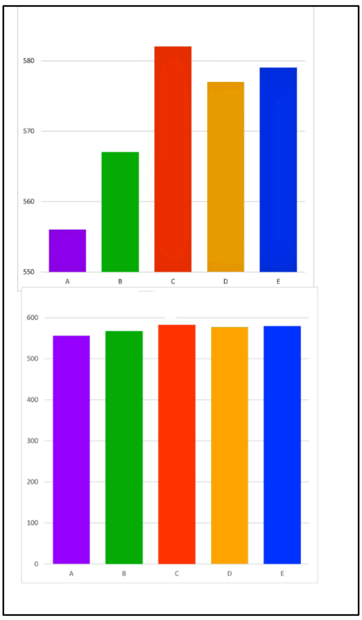

Figure 16. Erroneous use of a bar chart with a y-axis that does not start at zero (top), although a zero baseline is required (bottom). Showing only the tops of the bars exaggerates the differences among categories.

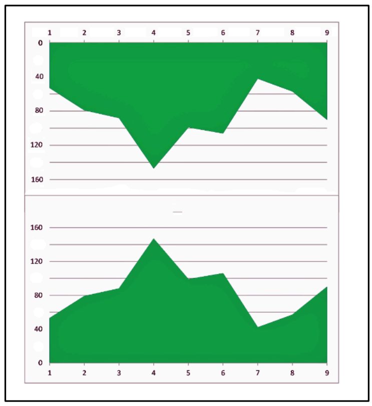

Figure 17. Erroneous use of an area chart with the x-axis along the top (top), instead of the correct orientation (bottom). The inversion gives a false impression.

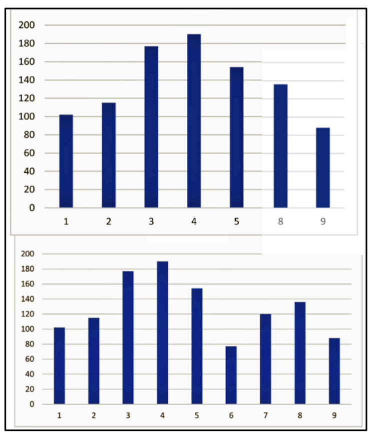

Figure 18 (right). Erroneous use of a chart that eliminates a section of the x-axis (for a variable such as time) (top), instead of one with an unmodified span of data (bottom). The omissions distort the pattern.

One set of errors comes from using the wrong type of chart or the wrong scale. Proportions should not be plotted on a bar graph as though they are data points, but shown as sections of a pie chart (Figure 15).

A relatively common error is truncating the y-axis to show only the tops of bars in a bar chart. The truncation may have been supposed to save space or make the information more visible, but it distorts the data by making the differences among treatments seem greater than they are (Figure 16). Dot plots and line charts, however, need not have a zero baseline because they focus on patterns rather than totals.

Data can be visually misleading if they have the wrong orientation. An area plot with the x-axis along the top gives the opposite sense to a normal plot (Figure 17). Likewise, closing up records when data are missing along a timeline—or other continuous axis—generates a false pattern (e.g., Figure 18).[3]

When the types of figures that will summarize and display the results to best advantage have been decided, publication brings additional requirements (Table 7). Most scientific publishers provide brief guidelines of this sort for their journals.

An important ingredient of all figures is a caption that is informative, yet as short as possible (just as for tables). Several drafts of a caption may be necessary to find the right balance.

A figure should be planned for publication at a size that will show all features adequately. For example, the published size determines what line thickness, font size, and other characteristics (e.g., the width of bars and the distance between them) will guarantee visibility. The aspect ratio of a figure may be dictated by its content, but the desired size in relation to journal dimensions may influence it too.

A hard-copy original is normally prepared larger than the size at which it will be reproduced. Size is not critical for digital publications, because the user can enlarge the display, but even so a figure that will be embedded in the text should be legible at a reasonable text size.

Images can easily be downsized for publication, although too much reduction generates unwanted effects. Digital magnification increases file size only by interpolating data. Accordingly, the intended resolution of an image should be decided early on. In any event, it is wise to err on the side of higher resolution and greater clarity of images if possible, because a copy editor may override the author's suggestions or intentions about size.

Technical considerations help to determine the format of digital image files. Most journals now accept many different file types. However, well-known lossless formats that compress files without deleting original data, such as tiff and png, are much larger, and hence more cumbersome to transmit, than "lossy" formats such as jpg that increase compression by permanently deleting data. Depending on the needs of the journal, they may be necessary only when images must have maximum quality. Several newer lossless formats are more efficient than tiff and png, and hence allow greater compression, but some journals do not yet accept them.

---

[3] Unfortunately, errors of charting in a few manuscripts appear to be deliberate, including tricks like mislabelling or distorting numerical data, and omitting an axis so that data can be lined up to suggest spurious correlations.

Table 7. Recommendations for scientific figures intended to be published.

General	
	Develop clear, concise caption
	Choose appropriate size and aspect ratio
	Use appropriate file format
Images	
	Use images to illustrate a point, not simply be attractive
	Include scale bar, or specify size(s) in the caption
	Label important items if not self-evident
Digitally enhanced or produced images (e.g., scans)	
	Specify meaning of different colours or symbols
Artwork	
	For realistic illustrations, follow guidance for images
	Ensure that artistic impressions are not misleading
	Ensure that relevance of creative illustrations is clear
Maps	
	Include latitude and longitude
	Label important locations and features
	Include scale bar
	Consider magnified inset for clarity (normally with its own scale bar)
Schematics	
	Include essential components only
	Highlight and label key components
	Avoid or limit elaborate icons and hieroglyphics
	Provide key to any symbols
Charts	
	Label all axes
	Use legible font and size
	Specify all units
	Label or identify all curves
	Label or identify all data sets

Other requirements for figures depend on their type, as shown in Tables 6 and 7 and Figures 1–14, and discussed above. However, all of them accord with four general themes.

First, a figure must help to move forward the message of the manuscript by giving pertinent information. Any figure that does not meet this criterion is superfluous, no matter how attractive or creative it is. Eye-catching or emotionally engaging illustrations are useful for some purposes, but not for most journal papers in entomology.

Second, a figure must be self-contained, allowing a reader to understand every component just by looking at the figure itself. Some items may need further explanation in the text, but the reader should not have to search for the identity of curves and bars, the parameters plotted on the axes of charts, the key structures, the meanings of symbols and colours, or the geographical location of a map.

Third, quantitative information must be included. For example, the units for different parameters, the size of specimens, and the scale of maps are called for.

Fourth, figures should be as easy to grasp as possible. Among guidelines already summarized, clarity stems from using the minimum complexity consistent with the purpose, and from elements of design, including the hue and saturation of colours, the thickness of lines, the size(s) and style(s) of fonts, and the aspect ratio of graphs. Testing several different versions of a chart may be needed to optimize its appearance. Potential ways to enhance a lifeless draft can often be evaluated most easily with the basic tools of image-editing software.

Revision and improvement

Once the whole manuscript has been drafted, it will have to be revised. These revisions improve the order of material, clarify the meaning, and streamline the text. For example, they ensure that statements are logical, add linkages, and correct remaining weaknesses of style. They confirm that tables and figures are well integrated and clearly titled. Depending on the nature of the manuscript, checking and editing may take as long as writing the first draft!

Reading the manuscript detects deficiencies of logic and flow, but typically misses detail. Therefore, the text should also be checked sentence by sentence to find defective spelling and grammatical errors. Subject-verb agreements, assignment of objects to the correct verbs, and arrangement of dependent clauses can be verified, for example.

Some sentences will have to be edited several times. An idea that has been noted for inclusion, and written up as part of the flow of facts and ideas in a section (without spending too much time improving the text of an initial draft), can later be revised.

Thus, a sentence from the introduction to this article might have been expanded as follows:

Idea noted*: Also for a thesis.*

Placed in context during an initial draft*: The same elements apply to graduate studies.*

Adjusted for clarity and precision, through one or more revisions: *The same recommendations apply to the preparation of graduate theses.*

Many draft sentences need to be condensed rather than expanded. Apart from improving the style (see *Conciseness* in the section on *Writing*), the content can be pruned.

For example, that same sentence from the Introduction might have been drafted in a longer form: *The same elements apply during graduate studies, which aim to produce a thesis, whether or not that thesis stands alone or comprises journal manuscripts.*

However, in this article the Introduction serves to outline the content as concisely as possible. Therefore, the potential elaboration at the end of the alternative draft [*…whether or not that thesis stands alone or comprises journal manuscripts*] would best be omitted, because it might distract the reader.

Once the manuscript has reached a late draft, it profits from further checks for repetition and bad habits. Repetition stems especially from the fact that a writer reaches a subsequent page much more slowly than a reader, even during revision, and is less aware of the fact that similar points may have been made, or the same words used, in many different places.

Most noticeable is the repeated use of relatively rare, distinctive, or elaborate words. One manuscript I reviewed contained the same dramatic word on almost every page. However, eliminating overuse applies only to striking English words, not to necessary scientific terms.

For example, *aegis*, *deft*, *eschew*, *moiety*, *plummet*, *salient*, and *voracious* should not be used multiple times in the same manuscript.

It is even more difficult for an author to notice the repetition of "ordinary" English words. Many authors overuse *some*, *many*, and other words that indicate approximate quantities, even when they are not required.

As these and other overused words build up through the manuscript, the reader may notice one of them and become distracted. Even so, ordinary words are customarily used many times in a normal text, so the process of reducing repetition does not have to be overdone. If a repeated word does seem to be intrusive, though, a word-processor search can be used to find all of the instances. Some can then be substituted with synonyms, or removed by rewriting sentences. To make such substitutions and balance their use during a few of my longer projects, I even compiled a list of synonyms for words that had become my favorites.

For example, there are useful replacements for *especially* and *particularly* [*chiefly, generally, in particular, mainly, mostly, normally, notably, predominantly, primarily, principally, usually*], and *significant* [*critical, essential, necessary, helpful, important, influential,* and others, depending on context].

However, unusual or weakly matched words should not be chosen as synonyms, because their presence would be more disruptive than the repetition they prevent.

For example, *concise* should not be replaced by *breviloquent* or *aphoristic*!

Also, so-called "elegant variation" within a paragraph (using multiple synonyms to avoid repetition or introduce variety) should be avoided if the variation would prove confusing, if pronouns would be more effective, or if simplifying sentences would be preferable.

For example, variation is overdone in the sentence *The caterpillar climbed the stem, and the early instar hung from the top as the hungry larva searched for a leaf of the foodplant.*

Most authors have acquired specific bad habits in addition to overuse of their favourite words and phrases, but familiarity causes them to be overlooked. Authors learn gradually to recognize these habits; referees may identify others! Some such defects can be sought and corrected through word-processor searches.

One common habit is compiling unduly long sentences (see *Style* above), which force the reader to plough through text that is too dense. Inversions are used repeatedly, but—like the instructional parody *Backward ran sentences until reeled the mind*—some are more difficult for readers to grasp than equivalent direct constructions. Unnecessary use of unusual or pretentious words, and overuse of present participles (especially *being*), are other frequent habits.

Revision may change the emphasis, and even the meaning, of a sentence. For that reason, the subfinal draft should be reviewed again with attention to cited references. Unless an author is certain that the changes have created no inconsistency with the content of each reference, it is wise to check back with the originals. It is better to validate everything at this stage than to be thought careless or ignorant by referees or future readers!

At a late stage, other people may be able to assist in tuning up the language. Many supervisors are willing to improve the English expression in manuscripts prepared by their students. Some colleagues will kindly include revisions of English in their private review of papers (see *Last checks* below), especially those prepared by authors working in a second language.

English can also be checked by online programs, which offer basic free versions as well as fee-based premium ones. Grammarly (Grammarly 2024), HemingwayApp (Hemingway Editor 2024), and ProWritingAid (ProWritingAid 2024), for example, flag problems and offer instructive hints as to how to remedy them. Unfortunately, many such programs (but not ProWritingAid) require users to grant them wide rights to the content entered for checking.

Grammar-checking programs use language "rules" to identify faults. However, the rigid rules are not always applicable, and most experienced writers judiciously break the rules anyway.

Default levels of these programs test whether the language is suitable for popular and advertising text aimed at a grade 5 level. Journalists rushing to meet daily deadlines are likely to find them useful. The academic or technical levels seem less helpful when evaluating detailed scientific texts. Typical entomological discussions and reviews include paragraphs with relatively long sentences that contain scientific and specific terms, and multiple citations. Automatic language checkers may judge them harshly!

Some programs will rewrite the author's text, using Artificial Intelligence. However, the design of typical programs seems to undervalue the fact that scientific papers seek to maximize rigour and precision. Rewritten versions characteristically include text in which accuracy and nuance—perhaps achieved with great effort by the author—have been reduced.

Manuscripts can be corrected by online companies that offer editing services for a fee (for example, 15 cents per word for rewriting). Some of them, typically with higher fees, specialize in scientific texts. I do not know the quality of the rewritten manuscripts. However, they are unlikely to be as precise as text prepared by someone familiar with the organisms and the data.

Minutiae

Checking the text for small imprecisions ensures general consistency and meets editorial requirements. Some typographical errors are not recognized by spellcheckers, including the automatic systems that operate in word-processing software. Typically overlooked are alternative spellings, some kinds of errors in grammatical number, almost identical words with different meanings, and other inconsistencies.

For example, inconsistent or wrong alternatives might remain in the manuscript for *fulfill* or *fulfil*; *grey* or *gray*; *larva* or *larvae; statistics* or *statistic*; *prescribe* or *proscribe*; and *opposite* or *apposite*.

Abbreviations for metric units follow international standards (Bureau International des Poids et Mésures 2022). Most journals leave a space between the number and the abbreviated unit. Statistical terms follow universal standards (e.g., ANOVA for analysis of variance). Some publishers provide a list of units and other preferred abbreviations.

Abbreviations used frequently as part of the text, and accepted by most publications, are listed in Table 8. The table is intended to give a sense of the minutely focussed work required to standardize abbreviations, rather than to offer a comprehensive list.

Most journals permit acronyms and uncommon abbreviations provided they are defined on first use. Other journals spell out all abbreviations in full, leading to cumbersome text despite efforts by authors to reduce repetition[4].

Many standardizations in the list of references are dictated by journal style. Reference-management programs (see *References*) help to generate these lists in appropriate formats.

For example, specified citation styles dictate whether multiple initials have a space between them or not, whether initials precede or follow the names when there are multiple authors, whether dates are enclosed in parentheses or followed by a period, whether every word of the title of a book or paper is capitalized or only some of the words, whether journal names are or are not italicized, abbreviated in set ways, or spelled out in full, and whether edited works use *Edited by*, *(Editor)*, or *(Ed.)*.

[4] I once prepared a manuscript about the Biological Survey of Canada that necessarily referred many times to that organization and to the Entomological Society of Canada. The first reference to each organization introduced the abbreviations BSC and ESC, which were then used for the rest of the paper. Subsequent copy editing spelled out each use in full, following the journal's normal standard. The scores of changes not only dramatically clogged the text, but also increased the published length. Fortunately, the Editor agreed with my assertion that the changes were counterproductive, and the requirement was waived for the special issue in which the paper appeared.

Table 8. A few of the common abbreviations and other text elements that are standardized in published work to ensure consistency. They are shown here in the form preferred by most journals.

three [not 3] for numbers 1–9	e.g., [not e.g.]†	nos [not nos.] for numbers
13 [not thirteen] for numbers 10 or greater	et al. [not et al]	p.m. [not pm]
2,000 *or* 2 000 [not 2000, except for dates]	et seq. [not et seq]	PhD [not Ph.D.]
BCE [not B.C.E.]*	i.e., [not i.e.]	Prof. [not Prof]
cf. [not cf]	MSc [not M.Sc.]	q.v. [not q.v.,]
Dr [not Dr.]	Mr [not Mr.]	viz. [not viz.,]
	no. [not no] for number	

*A few journals still use an older style that retains periods in such abbreviations as B.C.E. and M.Sc.

†Abbreviations of Latin terms (e.g., et al., et seq., i.e., q.v., and viz.) were italicized at one time, but this is no longer done. Few modern authors use other Latin terms that were once common, such as passim, ibid., and loc. cit.

Checking minutiae requires extraordinary attention to detail. However, trying to standardize them is the author's responsibility, despite the widespread belief that "There are people to do that." The more people striving for consistency the better. These tedious checks are facilitated by thorough reading of journal guidelines, by word-processor searches of the text, and by ensuring that citations were meticulously verified when first acquired or consulted (see *Managing references* above, and *Last checks* below).

Final title, and keywords

The title should be reconsidered when the manuscript is complete in case it can be improved after the content is fully known (see *Title* above for general information). Most titles are also considered in the context of the *keywords* called for by major publishers.

Keywords—together with titles and abstracts—are used by internet search engines to index content, and thus ensure that searchers will find titles of interest to them. They reflect the content of the paper more broadly than the title alone. Most but not all publishers suggest limiting overlap with title words. Short phrases are often more useful than single words, because they are more specific and overlap less with keywords in other papers.

Information on how keywords can be chosen parallels the extensive literature on choosing search terms for references (see *Reference searches* above). Commercial publishers, some university websites, and many companies that provide editing and related services also offer suggestions.

Search Engine Optimization or SEO (Enge et al. 2023), used especially by businesses to increase website traffic, includes related analyses, with an emphasis on marketing. SEO techniques seek ways to ensure that website content will be found by search engines such as Google, and to increase rankings, which determine the order in which results are listed. However, the algorithms used by search engines are modified frequently to improve the relevance of the results and to offset misleading or manipulative techniques used by some webmasters to increase traffic. Publishers are concerned with SEO (including keywords) in operating their websites, but most authors select keywords for papers simply by using subject knowledge and common sense.

Choices for any given paper are sharpened by first preparing a large pool of possible keywords for evaluation, including common synonyms of significant words in the paper. Word-cloud generators are also helpful. They analyze word frequency and relevance in a given text or texts (such as the manuscript and related articles). Words are then sized and arranged in a "word-cloud" display to reflect their potential importance and linkages. Many web-based word-cloud generators are free of charge (e.g., Tag Crowd 2024, Word Cloud Plus 2024, and many others). However, computer programs designed to select the actual keywords are untrustworthy, because they cannot integrate scientific information.

The basic procedure to choose keywords from the pool of possibilities (within parameters suggested by journals) is to focus on the main topic—including methodology if significant—and to be specific. However, undue specificity may fail to reach a potentially interested target audience beyond fellow specialists. The selections can then be optimized by searching in Google Scholar and reference databases (see *Reference searches*) to verify that they yield mainly papers relevant to the subject. The same tests can be applied to words in the abstract and final title.

Last checks

A final attentive reading of the manuscript before submission allows the text to be screened one last time for errors, ambiguities, and repetition, and checks that tables and figures are fully integrated. Authors should also take the time to verify any sentence, assumption, inference, conclusion, citation, or applicability of a reference that they are uncertain about. By this stage, there ought to be few remaining items of that sort.

The completeness and accuracy of citations should be confirmed. First, all text citations must accord with entries in the list of references, which should not include items that have not been cited. This requirement can be met through a page-by-page review of the manuscript to add conspicuous temporary marks to the text and the list of references as each citation is matched with its reference. After that review, the absence of a mark will reveal if any uncited reference is listed.

Second, if any citations were not verified during acquisition or reading, they should be checked with the originals to ensure that all details (authors' names, dates, titles, journal names, and volume, issue, and page numbers) are accurate. These checks are time-consuming when there are many references, but they prevent the unwelcome and potentially misleading errors of this sort that occur in the majority of manuscripts. In the past, many such errors were caused by unclear handwriting in transcribed citations. Most prevalent nowadays are keyboarding errors, and failure to check the citations listed by other authors.

It is wise whenever possible to set aside the manuscript for a week or two before these last checks. Revisiting it with fresh eyes may find defects that were missed, ranging from faulty logic to typographical errors.

It is also wise to ask two colleagues to review the "final" version of a journal paper before submission. These private reviewers need not be specialists, and should not be expected to rewrite English, nor to attend to other matters that are the responsibility of the author. However, even non-experts are able to detect faulty techniques, imprecisions of logic, expression, or syntax, and possible omissions. Authors should take full notice of their comments, as explained under *Responding to reviews* below.

Choice of journal

Scientists publish nearly all original research in scientific journals. An entomologist can choose to publish in any one of hundreds of different journals. The significance of that choice depends chiefly on the subject area and on the characteristics of the journal, as listed in Table 9. However, the importance of any given factor depends partly on circumstances.

Given the many variables, a likely journal for submission is best chosen early. The orientation of the study may suggest the most suitable journal, but alternatively the nature of a journal may affect the orientation of the paper. In any event, journal choice governs many details of format during preparation of the manuscript.

Table 9. General differences among scientific journals, influencing authors' choices for publication.

Component	*Potential elements*
Language	English, French, English and French, other
Region	International, national, regional
Scientific field by taxon	Science in general, biology, entomology, insect order or other restriction
Scientific field by subject	Target (basic or applied) Major arena (e.g., taxonomy, ecology, physiology) Limited subject or methodology (e.g., endocrinology, molecular biology)
Quality of scientific content	Editorial strategy (selection and judgment of papers) Perceived quality (e.g., specialist opinion, use by colleagues, citation-index ranking, other derived statistics)
Quality of production	Editorial diligence (e.g., consistency, absence of typographical errors) Production standard (e.g., paper and print quality, resolution of figures, permitted figure formats)
Audience	Small to large, general to specialized
Availability	Hard copy, digital, or both Circulation (tens to unlimited copies)
Approach of journal	Quantitative (mainly experimental and analytical) Qualitative (mainly descriptive and observational) Qualitative and quantitative Clarity of author instructions
Competition	Proportion of submitted papers that are accepted
Speed of publication	Average time after submission to acceptance or rejection (days to months) Average time after submission to publication (weeks to months)
Cost	Cost to publish (none to considerable) Cost to readers (none to considerable)

Journal characteristics appear in the instructions on journal pages or websites, and can be verified by studying individual issues. Some instructions include not only rules for submission, but also helpful general statements about the types of papers the journal prefers or insists on. In turn, these features influence the number and nature of readers.

Ideally, the results appear where they are most likely to be of service to science, reaching interested audiences, especially those with the greatest interest in the subject. Some possibilities are suggested by the choices made by colleagues studying the same topics. Other considerations reflect the location of the research, and the novelty, importance, or breadth of application of the results. Another set of determinants includes the scientific and production qualities of the journal, and its circulation and accessibility.

Several factors modify the choices. Coverage restricted to particular taxa or subjects limits the number of suitable outlets. The approach of a journal dictates the way that information has to be presented. Some journals publish short, highly focussed submissions (and tend to edit papers severely); others favour or allow extended treatments with intricate results or conclusions. Some insist on highly quantified data that meet rigorous statistical tests, or on investigations that address explicit narrow hypotheses; others accept observational or synthetic contributions. A few will publish basic data meant to support future research, including records of species collected in a particular geographical area.

The selection of an otherwise ideal outlet is tempered by the likelihood that the paper will be accepted. Publication of rejected papers is likely to be delayed, because the papers will have to be reformatted, and even rewritten, for submission elsewhere, and the review process will start again.

In most instances, the likelihood of acceptance depends on the subject matter and on the quality and conformity of the submission. However, competition to publish in the most widely distributed and highly rated journals is high, and they reject a substantial proportion of submitted papers. Furthermore, in highly sought-after and broad-ranging journals that receive more manuscripts of adequate quality than they can publish, acceptance depends on editorial judgments about timely, preferred, or important subjects, and on competition for space with other submissions received during the same period. These judgments change with time, and normally are not known to authors.

Selection of a journal is also affected by the typical interval between submission and potential publication. In some fast-moving research fields, timely publication is particularly desirable. Most online journals review and publish items more rapidly than those produced in printed form only. Nevertheless, some journals—of all kinds—take a long time to review and publish submissions. A typical interval could once be compared among journals because the dates of receipt and acceptance were published with each paper (documenting the earliest report of a particular finding), but many journals now show only when papers were first published online.

The costs of publication, and the costs to readers, have to be taken into account. More and more journals are being published only in digital form (without a printed version) and have no page charges. However, publishers may charge institutions and scientists high fees for access to the content, thereby limiting or modifying the readership. Others will allow open access to articles, but charge authors for this service. The impact on individuals of these costs for publication and access depends partly on the funding available through grants and employers, and on institutional edicts and agreements.

Journals to be studiously avoided by all authors are produced by predatory publishers. These publishers earn revenue by establishing substandard open-access publications, funded by the misled or misguided scientists who publish there. Most predatory publishers have several common traits, including aggressive marketing, concealing identity, falsifying or obscuring location and academic standing, and publishing papers of low quality without serious review.

For an introduction to predatory publishing, see Floate et al. (2013) and online overviews such as Madigan Library (2024), which cite references that identify in detail the traits likely to indicate a predatory publisher[5]. However, attempts to characterize these practices, and especially to identify individual perpetrators, have attracted push-back from publishers and spurred controversies about open-access publishing (see Beall 2017).

Moreover, these days, some scientists choose even fully reputable journals for reasons unrelated to the science being published, to the quality of the journal, or to attempts to reach the widest appropriate audience. Their behaviour reflects the degree to which careers can be constrained or advanced by the way scientific publications are evaluated to determine employment, promotion, and funding.

Some evaluations are made by administrators who simply measure the number rather than the quality of publications. Institutional restrictions that emphasize or mandate particular outlets or languages are another adverse influence. Judgments may rely unduly on indexes that purport to measure quality through “impact factors”, by comparing the number of times papers in a given journal were cited over the past two years. The data are used to generate a ranking for each journal (cf. Science Citation Index-Expanded, formerly Science Citation Index, and Journal Citation Reports: Clarivate 2024b, c)—but the rankings are for the journal as a whole, rather than for any individual paper or author.

This methodology leads to unfair judgments of scientific output (e.g., Seglen 1997; Curry 2018; Triggle et al. 2022). Seglen (1997) explained more than 25 years ago how the impact factors of journals are influenced by technicalities unrelated to the scientific quality of their articles.

In other words, some journal choices have been distorted in response to unfair evaluations of scientific quality, and by political pressures. The results are considered under *Ethical and legal considerations*.

Ethical and legal considerations

Introduction

Authors have a responsibility to science, and to their fellow scientists. It goes without saying, for example, that no result should ever be fabricated. For the same reason, every effort should be made to ensure that proper scientific methods (likely to yield trustworthy results) are used, and that a researcher is competent to do the work in question.

[5] For example, the contact pages for a predatory publisher may include only a web form, which does not reveal the publisher’s location. No genuine academic information may be provided about the editor or review-board members, and countries or scientists with no connection to the journal may be named in journal titles or listed on editorial boards. The journals are marketed to scientists at the same time as their true nature is concealed. Potential authors receive spam invitations to submit papers (often including compliments about their earlier genuine publications), but information about fees charged to authors may be insufficient or hidden. Moreover, most such publications merely post manuscripts as submitted, with little or no review or copy editing. They may include non-academic papers, and falsely claim impact factors or invented measures that feign international standing.

In contrast, some scientific work is published by people not qualified for the task. A portion of the so-called "grey literature" belongs to that category. The grey literature consists of items that have not gone through peer review, and includes reports by consultants, newsletters, working papers and plans, official and unofficial government documents, and policy statements. When people cannot distinguish the species under study nor apply valid statistical tests, for example, their results are rife with errors of identification and analysis. Other faulty and unreviewed work is published by predatory publishers (see *Choice of journal*), and by vanity presses. These products hinder scientific enquiry, because even conscientious authors might read and cite them without sufficient caution.

Another responsibility is to publish scientific results in a way that will be most helpful for future studies. Unfortunately, a pre-occupation with increasing the number of publications and their perceived value, in order to impress administrators and funding agencies (see *Choice of journal*), has led some scientists to offset this responsibility.

For example, some authors spread their results over as many papers as possible ("Least Publishable Units"), involving considerable repetition, although they would be the ones best placed to integrate the information properly in a smaller number of papers. Integration then has to be done, with more difficulty, by future authors.

Other authors indulge in "Journal chasing". Journals even remotely likely to accept the work are listed according to "impact factor". The paper is submitted first to the highest ranked journal. If rejected, it is submitted to the next journal on the list, and this process is repeated until the paper is accepted. However, each journal may require a different format or even a different approach. The strategy therefore not only engenders delays, but may force the results into a final form that is not optimal for future use, into an outlet that is seldom consulted by the scientists most interested in the topic—and sometimes, ironically, even into a journal that falls short of the quality implied by its ranking (see *Choice of journal*).

Responsibilities to other scientists also include never copying the work of other people without crediting them (see *Plagiarism*), ensuring proper assignment of credit for publications (see *Authorship and acknowledgement*), and taking note of ethical and legal obligations related to *Copyright*.

Plagiarism

Plagiarism is presenting as your own work the product or ideas of someone else. Even reproducing your own material without citation is usually regarded as plagiarism too, and recent definitions add the way in which Artificial Intelligence can be used to generate uncredited outputs based on other people's work.

Scientific responsibility therefore prohibits the theft of intellectual property. Plagiarism is a serious offence. Papers with plagiarized material are rejected outright by journals. Harsh punishments are handed out for plagiarism at most academic institutions, including failing grades and even expulsion.

Ideas should not be taken without permission or acknowledgement. Nonetheless, new ideas included in oral presentations or discussions by other scientists have sometimes been used to pre-empt their work, by rapidly following up the ideas in better funded laboratories. As a result, no-one in rapidly advancing fields now divulges interesting but unpublished information at conferences, in order to limit this sort of unethical behaviour.

How plagiarism is dealt with depends partly on its significance. A colleague once used a large section of one of my papers, verbatim and without acknowledgement, in his own article for an international newsletter. When I discovered it, his excuse was that he had no need to change the text because the subject could not have been written up any better! If the item had been a journal paper rather than a newsletter article, the transgression would have forced a formal response.

Plagiarism has been reduced since the digital revolution because software now exists to screen for it. Many plagiarism checkers are available for free download, although some are quite limited. The value of plagiarism checkers depends on whether they scan an extensive database beyond the web, and on other features, especially ease of use. Two of the most powerful are Unicheck (Unicheck 2024), used by many institutions, and Scribbr (Scribbr 2024). Some software suites used by journals to handle the submission and review of papers (see *The publishing process for journals*) include a plagiarism checker.

Authorship and acknowledgement

Authorship reflects creativity, and should be confined to participants who have contributed skill, judgment, analysis, and insight. Those who provide significant ancillary but not creative help, such as skillful technical assistance or comments on a draft manuscript, might be acknowledged but not credited with authorship.

Several difficulties in applying these strictures stem from funding patterns. For example, some people (including some managers and faculty members) expect to be included as authors on any publication from a project they helped to fund, even if they had little or no scientific input. This wish stems largely from the importance of authorship for promotion and further funding (see *Choice of journal*).

Other supervisors (as did mine) favour sole authorship by graduate students who have been almost entirely responsible for the conception and execution of their own projects. However, significant intellectual input by a supervisor should be recognized by authorship.

The establishment and order of co-authorship is normally assigned according to relative creative input. One ecological paper that I reviewed contained an independent taxonomic description of the study species that had been prepared by a specialist. However, the manuscript was set up in such a way that the author of the ecological section would not only be a co-author for taxonomic work to which he had not contributed, but would be the first-named authority for the specific name. This format, which would have resulted in the permanent registration of both names in the taxonomic literature, might have been adopted inadvertently.

Not every co-author acts appropriately. A colleague of mine once asked my advice because, to his surprise, his single co-author had added two technicians as co-authors on a final draft. Those people had played no role in the development of the project, nor of the manuscript. I advised him to challenge that inclusion, and later he told me that his co-author offered no objection and admitted that he was merely hoping to give unwarranted credit to his assistants and his laboratory!

In another instance known to me, an author submitted a lengthy manuscript for a proposed set of invited major syntheses, but the project faltered. Many years later, a similar project was launched by a different editor, who asked the author to update the work. He was too busy to attempt the substantial update required, and so the editor enlisted a second entomologist. The original author passed on detailed suggestions for revision to that colleague, who later reportedly withdrew without providing any feedback or revised drafts.

A year or two passed … and suddenly the original author received a revised manuscript, which had been submitted for publication before he had a chance to review it. Little of the content (including the core data) had been updated, despite some additions to the text and references—but the manuscript now showed the "reviser" as lead author, reversing the alphabetical order. The sequence of names was changed back only after objections were registered with the "reviser" and the editor.

These examples for authorship, and for plagiarism in the previous subsection, suggest that unexpected diligence is required to minimize potentially disruptive behaviours by other scientists.

Copyright

Copyright exists automatically in original works, published or unpublished, once they are in a fixed form. All items produced by skill and judgment, but not copied from another work, are deemed as original. They encompass written, artistic, and other products, including images and videos—and entomological papers. However, raw facts and ideas (but not certain expressions of ideas) and some other items are excluded.

The original creators of copyrighted products, or someone they have authorized, are the only ones with the right to reproduce them. These and many other aspects of copyright are governed by the Canadian Copyright Act (Copyright Act 2024).

However, fair-dealing exemptions allow limited use of copyrighted materials without the holder's permission, under certain circumstances. First, the use must be for an allowed purpose, which includes research, private study, education, and several other categories. Second, the use must be fair, an evaluation for each particular instance based on the criteria of purpose, character, amount, nature, effect, and alternative possibilities.

Authors share copyright when individual contributions to a joint work have not been distinguished. In most circumstances, one of the authors cannot exclusively license or assign the copyright, but has to deal with the other author(s)[6].

The prohibitions imposed by copyright are strictly applied by many copyright holders when use or potential use is for profit. However, some authors allow reproduction for purposes beyond fair dealing, and may include such specific permissions with the work itself.

In particular, images and other items are made available as part of the Creative Commons. This international non-profit organization enables sharing and reuse of creativity and knowledge through free legal tools (Creative Commons 2024a). Allowed uses, with authors credited, are specified by various licenses, which must be cited (see Creative Commons 2024b). Other images are free of copyright [CC0] and can be used in any way without acknowledgement. The fact that creators make their work available in this way is very helpful for some types of publication.

The copyright laws mean that publishers of journals and books insist on formal documents giving permission to publish, or even on transfer of copyright, before papers or chapters are published there. In addition, publishers typically require letters of permission from the

[6] However, Canadian law on this matter is not entirely settled.

copyright holder for any copyrighted material that has been included in a submitted manuscript.

The Copyright Act has been amended several times, and some significant copyright protections or assignments were changed. Some reflect new modes of digital production and digitally enabled copyright infringement. Especially relevant for scientific papers is that copyright for images belonged to the owner or commissioner prior to 2012, but the photographer is now deemed to be the creator. (However, copyright normally still belongs to the employer of people acting under a contract of employment.)

Although damages can be recovered for copyright infringement (and repeated and deliberate infringement risks criminal liability), pursuing remedies through the courts might take years and cost many thousands of dollars. However, authors risk personal or institutional liability if they mislead journals into publishing unauthorized materials, for example. Consequently, legal as well as professional and ethical constraints limit infringements by entomologists.

The publishing process for journals

Introduction

A paper is ready for submission only when the exacting process outlined above has been completed. The document then follows a peer review process, led by a scientific editor and implemented by volunteer reviewers (referees). The comments of referees may lead the editor to reject a paper outright, or to accept it, typically after the author has revised the manuscript in response to those comments. A representative process is outlined by Voight and Hoogenboom (2012).

Most major journals organize the stages of submission and review through computer software. These programs arrange for online submission, manage referee assignments and invitations, track manuscripts as they progress through review and revision, and facilitate the correspondence of authors, editors, and referees. Some programs include means to evaluate submissions for plagiarism and other defects, and provide automatic updates to authors and editors about the status of submissions.

Programs vary in capability, efficiency, ease of use, and price. Many programs claim—not unexpectedly—to be the best! Editorial Manager (Aries Systems 2024), ScholarOne (Clarivate 2024d), and Scholastica (Scholastica 2024), are among the programs most widely used.

All of these systems rely heavily on unthinking automatic computer processes rather than on individual human judgments. Therefore, authors can avoid potentially time-consuming missteps by careful review of journal instructions.

Scientific review

Normally, scientific editors or subject editors assign submitted manuscripts to referees for peer review. Several journals invite referee suggestions from authors, who are likely to know other specialists in their fields. Some of these suggestions may be accepted.

The editor or the publisher approaches potential referees to assess their willingness to review a particular paper. Editors learn which referees tend to accept assignments readily, which ones provide thorough reviews, which of them complete reviews in a timely manner, and sometimes (but not always) which specialists treat potentially competing manuscripts in their field unfairly.

Nevertheless, the process of assigning referees is inexact. Reviewers may be difficult to identify for a particular manuscript, especially one that encompasses several subjects. A few subject areas are studied by hardly any scientists. Potential referees with specific expertise may be too busy to accept the assignment.

Some journals rely heavily on a master list of names, addresses, and areas of interest, which they use to identify potential referees. Incomplete or outdated information in such databases generates capricious results. Therefore, not all of those who review a submitted paper are familiar with its field of research, even though most can detect general flaws.

The chief role of referees is to evaluate the conduct of the research, the validity of the results, and the way the data have been interpreted and reported. However, most referees do not have the time to correct large numbers of deficiencies of style or grammar, but only those that are confusing. Hence it is up to the author to ensure that a submitted manuscript has been well written. Otherwise, readers will have a low opinion of the work … or the manuscript may be returned by the editor for improvement before scientific review.

Responding to reviews

Many of the faults commonly detected by referees (e.g., Hoogenboom and Manske 2012) apply to the core research or the way it has been interpreted. They include use of inappropriate, suboptimal, or insufficiently described populations or instruments; small or biased samples; statistics that are inadequately chosen, applied, or described; and over-interpretation of results. Researchers who have conducted all stages of the research with care should see few comments of that sort.

Other common comments pertain to messages that are unclear, text that is written carelessly or with faulty grammar, manuscripts with marked inconsistencies, key papers that have been overlooked, tables and figures that are difficult to follow, and reference lists that are incomplete or full of errors. This article recommends ways to avoid such defects.

Even when text is generally well written and easy to understand, however, referees can help authors to identify potential problems of logic or expression. Typically, such comments arise because a section, paragraph, or sentence has inadequate content, structure, or grammar—even though the defect might not be the one supposed by the reviewer.

Authors are well-advised to park their egos and assess all of these comments objectively; most referees are knowledgeable and try to be helpful rather than unduly critical. Even if a reviewer (or editor) seems to have missed the point or appears to be unfamiliar with the subject, there may be some ambiguity of phrasing that could be improved.

By the same token, an author's response to the editor should not brush off or ignore comments even from referees who have misunderstood, or appear to be ignorant or misinformed. The editor expects to see a clear explanation or justification for non-acceptance of apparently substantive comments. Such responses must be unwaveringly polite, with no implied criticism of referees!

Reviewers do not always agree about an item. I once received two comments on the same paragraph, one of which read "I would be careful with this idea", and the other "Let us hear more about this idea". Little change seemed necessary…

Not all reviews are detailed. The shortest review I ever received, for a manuscript of about 50 typescript pages, comprised only three words, with no additional comments. That positive statement made for an easy response to the referee's assessment, but actually authors should welcome more diligent input. No papers are perfect.

Valuable examples of how to prepare a manuscript come from reading published papers that are well prepared. Authors who have structured the information clearly, written concisely, used optimal syntax, and chosen precise words are worth emulating.

Scientists are also likely to profit from their own roles as reviewer (or editor), because errors in procedures, analysis, and expression are particularly noticeable when they have been made by other people! One of the most common writing defects in my experience is a weak topic sentence to begin a paragraph (see *Writing paragraphs*). Readers then have to work out for themselves the relevance of the detailed content.

Referees are not always rewarded for their efforts. I was once asked to review a long manuscript in which the scientific content was adequate, but the English was seriously flawed. Rewriting a paper, as opposed to suggesting local improvements, is not the reviewer's job (the editor ought to have returned the manuscript to the author for rewriting before scientific review), but I corrected the English in the first few pages of the manuscript anyway and explained the corrections, stating that such changes should be applied throughout. In the version eventually published, all of my suggested changes on the first few pages had been faithfully transcribed—but the error-filled remainder was untouched, leaving countless repeated faults for which both author and editor bore responsibility.

In another instance, I reviewed a manuscript submitted to a leading journal by a recognized scientist who was named as first co-author. The second co-author was a graduate student. The scientific content was satisfactory, but the English was strikingly flawed. Certain that this scientist knew better, I corrected just one of the double-spaced pages, which had more than 20 errors of grammar alone! In this instance, the whole paper was improved before publication. Perhaps the senior author's deserved embarrassment contributed to that effort…

My review would have shown the authors that their sloppiness was now known to the editor—and to at least one specialist in their field. They might wonder if that knowledge could influence the fate of future manuscripts, grant proposals, symposium invitations, or student recommendations, for example. Such a possibility reinforces the lesson of this article, that no manuscript should be submitted until it has been carefully prepared and fully checked

Proofs

The arrival of proofs gives the author a final chance to review the paper, so proofs should be examined thoroughly. Even the most diligent examiners tend to begin concentrated effort at the start of the text, however, and therefore frequently overlook mistakes in header material, in the title, and even in the author's name!

Two sorts of checks are necessary. Reading through reveals general errors. Meticulous sentence-by-sentence and word-by-word comparison of the proofs with the original manuscript finds typographical and other specific errors. Without this last step, errors are easily overlooked, because authors tend to concentrate on the sense. They read fluently and "see" the words they expect. My experience as an editor (albeit much of it when additional proof errors were introduced by typesetting[7]) showed that most authors miss at least half of the errors in their proofs—and some miss nearly all of them.

Attentive proof-reading is important even for manuscripts that have been processed digitally. Mistakes may have been made when text was formatted for publication, when tables and figures were inserted, or when captions were applied. In particular, some tables submitted in digital form are keyboarded again or reorganized by the printer, potentially introducing typographical errors and misalignments.

Detailed review of the proofs is also necessary because copy editors who do not understand the content may "adjust" the text. Surprisingly, after acceptance following review, one of my papers for a well-known journal (*Annual Review of Entomology*) was subjected by the copy editor to numerous minor, and apparently pointless, text edits that changed the meaning—but those changes were not disclosed to me until the page proofs arrived. It was time-consuming to find and correct the many errors that had been introduced. It would have been nearly impossible for an author working in their second language to detect and correct them before the short deadline that had been imposed.

Copy editors may also introduce errors by applying "standardizations", especially in the reference list. For example, issue numbers are essential when each issue in a volume is paginated separately[8]. However, many journals that normally omit issue numbers from cited references remove essential issue numbers too, making the citations ambiguous. In one of my papers with references in German, the capital first letters of German nouns in titles were changed to lower case to accord with the format of the journal for English titles; but such capitalization is part of the German language!

A few copy editors remove (and some printers distort) foreign diacritic marks, even when there are generally accepted transliterations. Such changes may butcher names or introduce misunderstandings.

These examples confirm the need for extreme diligence during the proof stage. Previously overlooked errors might be noticed because time has elapsed since submission. New errors might have come from editorial or printing processes.

---

[7] Before manuscripts were transmitted digitally and modified electronically, they were retyped during typesetting.

[8] Publishers should always number pages in the same annual volume serially, rather than separately for each issue, but not all of them do so.

Summary

A journal paper or a graduate thesis requires careful attention to every component of its development: planning and conducting the research; analyzing the results and placing them in context; writing the manuscript; and bringing it to publication.

Worthwhile scientific content is the core of the work, but presentation is critical too. The necessary steps are shown in Table 10. They were discussed in detail in the sections above, which are indexed in the Table of Contents.[9]

[9] Further information is available on the web, but should be treated with caution because some of the facts and recommendations there are incomplete or incorrect.

Table 10. Summary of steps recommended to complete a paper in a scientific journal.

Steps	*Key requirements*
Planning	
Nature of enquiry	Define subject area, questions, and hypotheses
Design	Identify potential methods for experiments, other investigations, and required analyses
Projected outputs	Decide potential subjects for papers, as well as popular articles or other products
Retained material	Preserve necessary records and specimens
Preliminary considerations	
Potential journal	Evaluate audience, production quality, likelihood of acceptance, and other features
Authorship	Recognize creative contributions
Copyright, plagiarism	Follow ethical and legal constraints
Execution of research	
Procedures	Conduct sampling, experiments, and other work meticulously
Notes	Record methods, and general observations, in writing
Interpretation	Ensure appropriate analysis and follow up
Manuscript design	
Title	Establish initial working title
Draft components	Divide into logical sections and subsections
Writing	
Message	Lead readers towards a clear conclusion
Structure	Optimize flow of subsections, and balance them as far as possible.
Language use	Write paragraphs and sentences that are well structured, grammatically correct with clear style, and correctly worded
Scientific literature	
Initial assessment	Use preliminary knowledge to plan search
Reference search	Make comprehensive search
Management	Establish reference database
Use	Record content accurately and use it appropriately
Correctness	Verify the content of cited references, and their transcriptions

Coninued...

Table 10 *(continued)*. Summary of steps recommended to complete a paper in a scientific journal.

Steps	*Key requirements*
Data presentation	Develop potential tables, and types of figures
Tables	Many detailed considerations, including cogent captions (see Table 5)
Figures	Many detailed considerations, including cogent captions (see Tables 6–7 and Figures 1–18)
Text revision	
Content, structure, paragraphs, sentences, word use	Verify accuracy, optimize order, balance sections, improve English, make more concise, increase precision, correct minor errors
Linkage to data and illustrations	Ensure that tables and figures are well integrated with the text
Other checks	Remove distractions, such as undue repetition
Standardization	Ensure consistency of minor text elements, and follow journal standards
Final title and keywords	Finalize title; choose keywords
Last checks	Re-read completed manuscript with care
Private review	Seek and incorporate input from colleagues
Submission	
Journal review	Incorporate input from referees
Proofs	Check proofs in detail

The first step towards publication is to prepare a focussed plan for the research, rather than gathering scattered pieces of useful information that may not result in coherent products. Thoughtful design aims to generate data that will both address the topic successfully and underpin manuscripts treating questions of interest. At an early stage, too, a plan is needed to ensure preservation of records and specimens for the work itself, as well as for future use.

Several other decisions are best made early. Choosing a likely journal helps to structure the manuscript. That choice depends on the nature of the manuscript, the type and number of scientific readers, the location and production quality of the journal, the likelihood that a paper on a given topic will be accepted, and other factors (see Table 9). Other preliminary decisions decide authorship (only people making creative contributions of intellectual property), try to ensure that the paper will properly serve science, and arrange appropriate credit for material published by others.

Sampling, experimentation, and synthesis must be done carefully and consistently during the research. Contemporaneous notes about methods, and written general observations, support preparation of the manuscript as well as future interpretation of the results. Analyzing numerical data depends on statistically appropriate procedures. It is assisted by considering features of the organisms, not just the numbers.

Preparing the manuscript involves writing the text, using the scientific literature, and deciding how to present data in tables and figures. A draft title and a working outline should first be developed. Useful titles are as concise as possible, but tell potential readers the exact nature and significance of the results. The components of the manuscript (sections and subsections and their headings, and the order in which they appear) are intended to guide readers as effectively as possible towards the conclusions of the paper. Developing an appropriate structure is most likely when the author already knows the field of research thoroughly.

The role of the actual writing is to impart the messages of the manuscript as clearly as possible. To do this, the flow of the content has to be optimized, and the language impeccably crafted (see Tables 3 and 4). Essential are structured paragraphs that begin with topic sentences, and well-ordered sentences that are grammatically correct, smoothly styled, clear, concise, and precisely worded. Therefore, writing is a time-consuming process of diligent drafting and repeated revision.

The importance of using scientific literature to steer the work and set its findings in context cannot be overstated. Existing knowledge can be integrated properly only by a comprehensive search for relevant references, and by tracking and recording their content in an organized way. It is also necessary to match citations in the manuscript with entries in the list of references, and to verify that all details in that list are correct.

Many different types of potential tables and figures can be used to present the results. Which ones display the findings most clearly are suggested by detailed consideration of all of the collected data, and by tabulating or plotting data in different ways to see what they might show or where they might lead. How data in the same field have been presented in other publications is also informative.

Preparing tables and figures for publication therefore entails many decisions about relevance, format, explanation, precision, treatment of data including statistical validity, and other features shown in Tables 5–7 and Figures 1–18 above. Concise and informative captions make the content and objective of each table and figure clear to readers.

Once the manuscript and its supporting elements (tables, figures, and references) have been drafted, successive efforts can be made to improve them. The goals are to check that the intentions for the initial draft have been met, remove errors of presentation or analysis, sharpen the language, and correct minor items. In terms of content, they verify that the text accords with the data in tables and figures, that material is in the correct order for the reader to follow the message, and that information from the literature has been integrated. In terms of language, the revisions try—just like the original writing—to ensure that the English is grammatically correct, easy to understand, succinct, and logical, and that individual words have been optimized for clarity and nuance of meaning.

Revisions also seek to eliminate undue repetition and other reader distractions that stem especially from accustomed habits of the author. Standardizing all minutiae, such as abbreviations, reference formats, and alternative spellings tidies up the manuscript.

When the manuscript is complete, the title can be finalized. At that time, keywords can be chosen. They are required by most journals, and are intended to maximize the notice a paper will receive from internet searches. Identifying the most appropriate keywords stems from careful analysis of the content of the paper, including potential synonyms, and from considering the words used in allied publications.

The next step is to re-read the whole manuscript. By now the work will be so familiar that remaining faults are easy to overlook, so this task calls for combing slowly and attentively through the text. A private review by colleagues at this time is especially useful, because they may see things that were missed.

After submission, the comments of journal referees provide another opportunity to improve the manuscript. All of their feedback should be considered in detail, because at least some of the comments are likely to suggest how to refine the presentation of results or clarify the language.

Last but not least, the proofs of an accepted paper have to be checked thoroughly. Reading through and word-by-word examinations are both advisable, even for manuscripts submitted digitally, because errors may have been introduced during editing or printing.

In essence, completing a manuscript of good quality requires planning, diligence, rigour, and concentration through multiple stages. Although all activities profit from that sort of detailed effort, careful work and repeated search for improvement are particularly important for journal papers. Those documents, and the theses on which some of them are based, serve as the main foundation for future scientific work. Moreover, they are visible to others—and attributed to authors—in perpetuity.

List of references

[All web links were verified on 16 July 2024]

Adobe. 2024. Photoshop. [Image editor] https://www.adobe.com/ca/products/photoshop.html

Aries Systems. 2024. em. Editorial Manager. [Journal manager] https://www.ariessys.com/solutions/editorial-manager/

Beall, J. 2017. What I learned from predatory publishers. *Biochemia Medica* 27(2): 273–278. https://doi.org/10.11613/BM.2017.029 [Beall's archived list of potential predatory journals and publishers is at https://beallslist.net]

Briscoe, M.H. 2012. Preparing scientific illustrations: A guide to better posters, presentations, and publications. 2nd edition. Springer Science and Business Media. 203 pp.

Bureau International des Poids et Mésures. 2022. The International System of Units (SI), vol. 2. 9th edition. Available at https://www.bipm.org/en/publications/si-brochure

Clarivate. 2024a. Web of Science platform. [Reference database] https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/

Clarivate. 2024b. Science Citation Index-Expanded. https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/web-of-science-core-collection/science-citation-index-expanded/

Clarivate. 2024c. Journal Citation Reports. https://clarivate.com/products/scientific-and-academic-research/research-analytics-evaluation-and-management-solutions/journal-citation-reports/

Clarivate. 2024d. ScholarOne. Comprehensive workflow management for scholarly journals. [Journal manager] https://clarivate.com/products/scientific-and-academic-research/research-publishing-solutions/scholarone/

Copyright Act. 2024. [Canadian Copyright Act 1985 as amended]. https://laws-lois.justice.gc.ca/eng/acts/C-42/Index.html

Corbet, P.S., and H.V. Danks. 1973. Seasonal emergence and activity of mosquitoes in a high arctic locality. *Canadian Entomologist* 105(6): 837–872.

Creative Commons. 2024a. (CC) Creative Commons: better sharing, brighter future. creativecommons.org

Creative Commons. 2024b. (CC) Creative Commons: about CC licenses. creativecommons.org/share-your-work/cclicenses/

Curry S. 2018. Let's move beyond the rhetoric: it's time to change how we judge research. *Nature* 554(7691): 147.

Danks, H.V. 1981. Arctic arthropods. A review of systematics and ecology with particular reference to the North American fauna. Entomological Society of Canada, Ottawa. 608 pp.

Danks, H.V. 2006. Short life cycles in insects and mites. *Canadian Entomologist* 138(4): 407–463.

Danks, H.V., and D.R. Oliver. 1972. Seasonal emergence of some high arctic Chironomidae (Diptera). *Canadian Entomologist* 104(5): 661–686.

EBSCO. 2024. EBSCOhost. Reliable information for all kinds of research. [Reference database] http://www.ebsco.com
Elsevier. 2024. Scopus: Comprehensive, multidisciplinary, trusted abstract and citation database. [Reference database] https://www.elsevier.com/products/scopus
EndNote. 2024. EndNote. [Reference manager] https://www.endnote.com/
Enge, E., S. Spencer, and J. Stricchiola. 2023. The art of SEO: mastering Search Engine Optimization. 4th edition. O'Reilly media, Boston. 741 pp.
Evergreen, S.D.H. 2019. Effective data visualization: the right chart for the right data. Sage, Thousand Oaks, California. 352 pp.
Floate, K., S. Hoover, and C. Cutler. 2013. Open access, predatory publishers, The Canadian Entomologist, and you. *Bulletin of the Entomological Society of Canada* 45(3) 131–137.
Google Scholar. 2024. [Web search engine] https://scholar.google.ca/
Grammarly. 2024. Grammarly. Responsible AI that ensures your writing and reputation shine. [Grammar checker] https://www.grammarly.com
Healy, K. 2018. Data visualization: a practical introduction. Princeton University Press, Princeton. 296 pp.
Hemingway Editor. 2024. Hemingway App makes your writing concise and correct. [Grammar checker] https://www.hemingwayapp.com
Hoogenboom, B.J., and R.C. Manske. 2012. How to write a scientific article. *International Journal of Sports Physical Therapy* 7(5): 512–517.
Madigan Library. 2024. Predatory publishing. Pennsylvania College of Technology, Madigan Library (Research Guides-Predatory Publishing). https://pct.libguides.com/c.php?g=998550&p=7228380
Mendeley. 2024. Mendeley. [Reference manager] https://www.mendeley.com/
Microsoft. 2024. Microsoft Excel. Turn data into insights with free and premium spreadsheets. https://www.microsoft.com/en-ca/microsoft-365/excel
Nguyen, V.T., K. Jung, and V. Gupta. 2021. Examining data visualization pitfalls in scientific publications. *Visual Computing for Industry, Biomedicine, and Art* 2021: 4–27. https://doi.org/10.1186/s42492-021-00092-y
Paperpile. 2024. The top list of academic research databases. https://paperpile.com/g/academic-research-databases/
ProWritingAid. 2024. Publish without self doubt. [Grammar checker] prowritingaid.com
Reed, J.G., and P.M. Baxter. 2009. Using reference databases. pp. 73–101 *in* Cooper, H., L.V. Hedges, and J.C. Valentine (Eds), The handbook of research synthesis and meta-analysis. Russell Sage Foundation, New York. 614 pp.
Scholastica. 2024. Scholastica. Modernize your journal publishing process. Further your mission. [Journal manager] https://scholasticahq.com/
Scribbr. 2024. Scribbr. Plagiarism checker (premium version). https://www.scribbr.com/plagiarism-checker/
Seglen, P. O. 1997. Why the impact factor of journals should not be used for evaluating research. *BMJ* 314 (7079): 498–502.
Serif. 2024. Affinity Photo 2. The photo editor you've been dreaming of. [Image editor] https://affinity.serif.com/en-us/photo/
Swan, M. 2016. Practical English Usage. 4th edition. Oxford University Press, Oxford. 635 pp.
Tag Crowd. 2024. [Word cloud generator.] tagcrowd.com
Technical University of Munich University Library. 2022. Reference management software comparison—9th update. https://mediatum.ub.tum.de/doc/1320978/1320978.pdf
Triggle, C.R., R. MacDonald, D.J. Triggle, and D. Grierson. 2022. Requiem for impact factors and high publication charges. *Accountability in Research* 29(3): 133–164.
Tufte, E.R. 2001. The visual display of quantitative information. 2nd edition. Graphics Press, Cheshire, CT. 197 pp.
Unicheck. 2024. Unicheck. Plagiarism checker that prefers results over numbers. https://unicheck.com/
Voight, M.L., and B.J. Hoogenboom. 2012. Publishing your work in a journal: understanding the peer review process. *International Journal of Sports Physical Therapy* 7(5): 452–460.
Wanyama, S.B., R.W. McQuaid, and M. Kittler. 2022. Where you search determines what you find: the effects of bibliographic databases on systematic reviews. *International Journal of Social Research Methodology* 25(3): 409–422.

White, H.D. 2009. Scientific communication and literature retrieval. pp. 51–72 *in* Cooper, H., L.V. Hedges, and J.C. Valentine (Eds), The handbook of research synthesis and meta-analysis. 2nd edition. Russell Sage Foundation, New York. 614 pp.
Word Cloud Plus. 2024. WordCloud+: Word cloud generator. https://wordcloudplus.com/?ref=cms.wordcloudplus.com
Zotero. 2024. Zotero. Your personal research assistant. [Reference manager] https://www.zotero.org

List of Tables

List of Figures