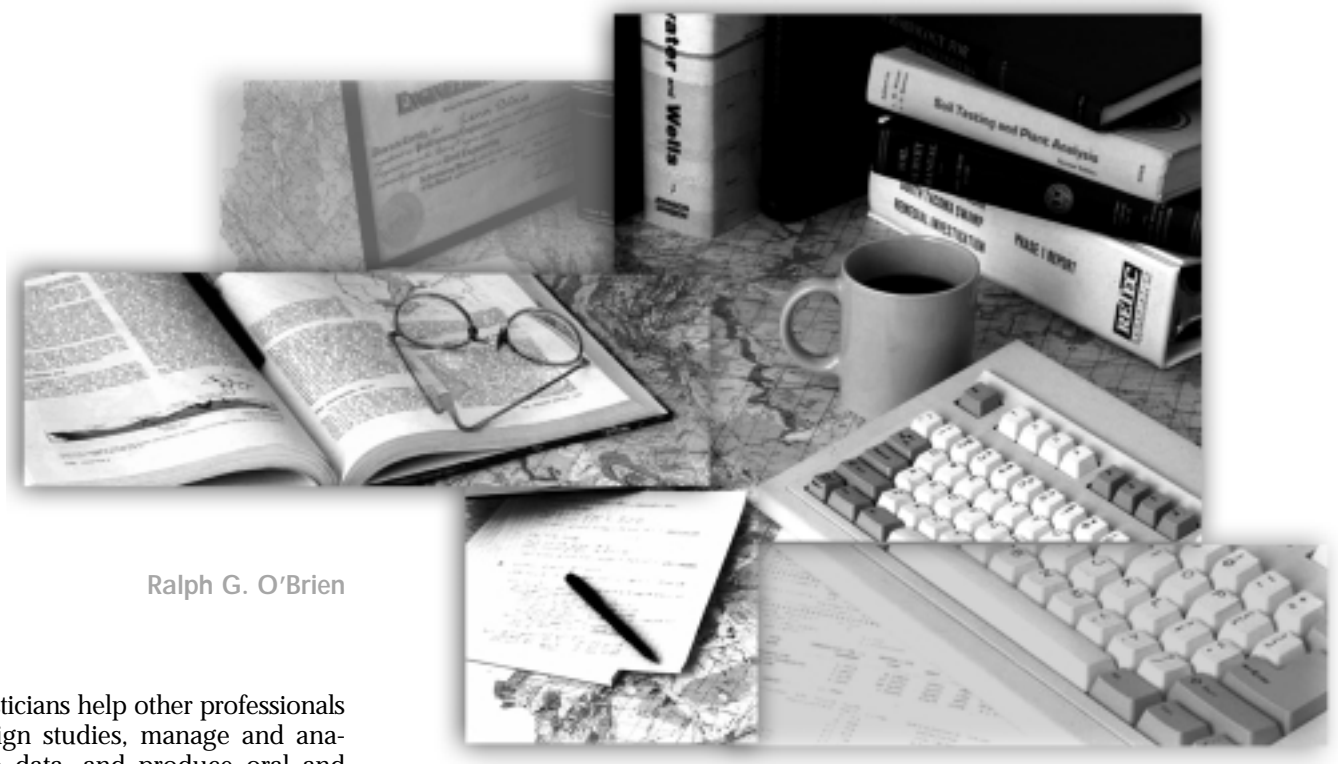




Discover the Science of Technical Writing



Ralph G. O'Brien

Statisticians help other professionals design studies, manage and analyze data, and produce oral and written reports. Degree programs in statistics concentrate on building the mathematical and computational skills required in the profession, and they do this well. But what about building communication skills? Almost every facet of our profession requires us to write clearly, yet what we produce is often onerous to read. This is why position notices advertise for statisticians who can “demonstrate excellent oral and written communication skills.” Let’s face it: most academic statistics departments give short shrift to developing skills in technical writing. To become a first-rate statistician, you may have to fill that void on your own.

Writing at a professional level takes substantial time and discipline, even after years of experience. Only the rare person possesses a natural gift for it. Moreover, technical writing is fundamentally different from writing in the humanities. W. Earl Britton stated the problem boldly when he titled his 1974 essay “The Trouble with Technical Writing is Freshman English.” Even if you earned high marks in college for penning creative essays, you may need to develop new skills.

Fortunately, there is a science to technical writing based on a set of prin-

ciples. If this intrigues you, read George Gopen and Judith Swan’s seminal article, “The Science of Scientific Writing” (*American Scientist*, 1990, 78:550-558), available through the Web site of the ASA’s Section on Teaching Statistics in the Health Sciences at www.bio.ri.ccf.org/ASA_TSHS. Gopen and Swan argue that complex technical matter should not lead to impenetrable writing. Technical writers need to understand how readers read, and they must write with the reader’s perspective in mind.

Gopen and Swan's principles

Many lessons on writing begin by stressing the need to 'know your audience.' We statisticians do indeed write for a variety of readers, and this shapes our decisions on technical levels, vocabulary and jargon, formality of discourse, and examples and analogies. But knowing your audience does not tell you how to write.

Gopen and Swan's overarching principle is:

Place each unit of information where readers expect to see it.

A "unit of information" can be a word within a phrase, a phrase within a sentence, a sentence within a paragraph, a paragraph within a section, a section within a chapter or article, or a chapter within a book. It can be a table or graphic.

Look at each unit and ask 'Is this the best place to put this?' When readers find information where they expect it, they are better able to focus on the substance of the piece, rather than its style. Gopen and Swan delineate specific principles on how to fulfill readers' expectations.



Take advantage of what goes in the topic and stress positions. Readers pay greatest attention to the words at the beginning and the end of a sentence. The beginning should usually hold the topical subject, which sets the context for the reader to process the rest of sentence. Compare:

1a. Bootstrapping can also be used to estimate standard errors.

1b. Standard errors can also be estimated using bootstrapping.

Both impart the same information, but the content of their topic positions tells the reader that the first sentence is about *bootstrapping* and the second is about standard errors. The topical subject is typically 'old information' to the reader; in fact it often links back to the previous sentence. By the same principle, readers expect the first sentence in a paragraph to be the thesis sentence. Likewise, a new section should open with an introductory paragraph.

The end of a sentence should stress 'new information' that propels the discourse forward. Which one of the following sentences would be most effective if written for the *Amstat News*?

2a. Starting salaries are good and getting better in statistics.

2b. Starting salaries in statistics are good and getting better.

The first reads flatly because "in statistics" occupies the stress position, yet for *Amstat News* readers this is old information. In the second sentence, placing the stress on "good and getting better" satisfies the reader's desire for something new and interesting at this point. Again, the same principle applies to other information units. Thus, a paragraph's final sentence should stress new information that pushes the reader's interest into the next paragraph.

Of course, not all text appears in the topical and stress positions. Ancillary text is needed to expand on the topical subject and provide contextual support for what is being stressed. Each sentence should have a single mission and each word or phrase in it should support that mission. Likewise each sentence should support its paragraph and each paragraph should support its section. If text is not supporting the mission of its unit, it should usually be excised.

Place subjects and their verbs near each other, and use verbs to locate the action. Consider the following:

3a. The propensity score approach for bias reduction in treatment comparisons in non-randomized studies has led to another round of debates about the use of correlational analyses to make inferences about causality.

This is grammatically correct, but the reader's attention fades quickly when verbs are used so poorly. Here it is again with the two long supporting phrases obscured:

3a* The propensity score approach [A] has led to another round of debates [B].

This shows how the reader must trudge in limbo through the nine words in [A] to get from the subject, "propensity score

Before You Begin...

Read the publication and its style guide. ASA publications use the *Chicago Manual of Style*. All references cited in the article must be included in the footnotes. Examples of properly documented references are:

Dixon, W. J. (ed.) (1983), *BMDP Statistical Software* (Vol. 1, 3rd ed.), Berkeley, CA: University of California Press.

Efron, B., and Morris, C. (1977), Comment on "A Simulation Study of Alternative to Least Squares," www/publications/styleguide.html by H. Clark and T. Schwisow, *The American Statistician*, 72, 102-109.

■ Check the grammar, spelling, punctuation, and continuity. Also have someone else proofread the document.

■ Avoid abbreviations and technical jargon.

■ Use a colon to introduce a list of more than three items coming at the end of a sentence, to introduce an example or explanation related to something just mentioned, or to introduce a quotation over three lines. Do not use a colon after such as or including, or after any form of the verb be.

■ For a list of three items, place a comma after the first and second items. For example, "Statisticians work in industry, academia, and government."

■ Submit an article to a journal in manuscript format, which means a single column, double-spaced, with 1" margins.

■ "[G]raphics should be reserved for the richer, more complex, more difficult statistical material (Tufte, 1983)." Use text to support and explain the tables, not to simply reiterate the data. For more information on visualizing information, read Edward R. Tufte's books *Envisioning Information* and *The Visual Display of Quantitative Information*.

approach,” to the verb, “has led.” Readers expect to find verbs close to their subjects.

Readers also like verbs to locate the action. The only verbs in 3a, “has led” and “to make,” are lifeless. The only action comes from nouns such as “reduction” and “comparisons.” Now try this revision:

- 3b. The propensity score approach offers a way to reduce bias when comparing treatments in non-randomized studies. However, researchers are debating yet again about using correlational analyses to infer causality.

Note, for example, how the passive phrase “bias reduction in treatment comparisons” became the active phrase “to reduce bias when comparing treatments.” Can you find similar changes? Can you improve it?

Link all units of information; avoid logical gaps. Each unit of information should have clear links to closely neighboring units. Subtle links that seem obvious to the writer may not be evident to the reader. When the chain of discourse is broken, readers become confused and frustrated. Their reading speed slows and their understanding suffers. Consider:

- 4a. Career opportunities are strong in statistics and starting salaries are good and getting better. Statisticians who have strong oral and written communication skills are the most successful.

Did you stumble a bit at the gap between the sentences? In fact, the second sentence has no specific tie to the first unless the reader pauses to build one. Now try:

- 4b. Career opportunities are strong in statistics and starting salaries are good and getting better. This is especially true for new statisticians who have strong oral and written communication skills.

The simple word “this” reaches back to the first sentence, thus preserving the chain.

Principles, not rules

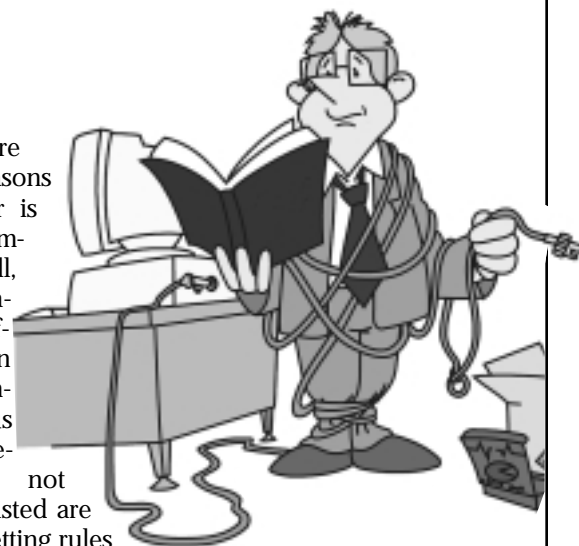
Readers become confused, even if only for a split second, whenever the writer violates these and other principles of good writing. However, these are not rules to be followed blindly. Not every sentence can have important material in its stress position. You may even decide at times to use less active verbs, because you need a sentence to be as bland as possible. Sometimes you may choose to be ambiguous or obscure. When skilled writers violate these principles, they do so intentionally and sparingly. Unskilled writers violate these principles unknowingly and excessively.

Learning to write professionally

You can learn to write well. After studying the Gopen-Swan article, treat yourself to William Strunk and E. B. White’s *The Elements of Style* (4th ed., 1999, \$9). This timeless classic captures the essence of clear writing, but it lacks detail and exercises. Joseph Williams’ *Style: Ten Lessons in Clarity and Grace* (6th ed., 1999, \$31) covers the same principles as the Gopen-Swan article and *The Elements of Style*, and also provides the necessary instructional depth and exercises to enable you to master these skills. For a handy reference tool, Alfred, Brusaw, and Oliu’s *Handbook of Technical Writing*
continued on page 35

Why Is the Computer not Like a Typewriter?

There are obvious reasons a typewriter is not like a computer. Still, some fundamental differences in using a computer versus using a typewriter are not practiced. Listed are some typesetting rules that will help produce a professional-looking manuscript on a computer, and not one that looks like it was typed on an old typewriter.



■ Leave only one space after periods, question marks, exclamation points and colons, instead of the customary two spaces. On a typewriter each letter takes up the same amount of space. On a wordprocessor, each letter is proportional to the space it needs. In other words, the ‘m’ takes up more space than the ‘i’. To keep all of the copy proportional, a typesetter will delete the double spaces after punctuation.

■ Use smart quotes rather than inch marks. Smart quotes look like this, “”. Quote marks look like (“”), and are often mistaken for inches. Some wordprocessors automatically change quote marks into smart quotes, but double checking the paper for smart quotes before sending it, is a good habit to develop.

■ Don’t underline. Underlining is for typewriters; italicize text instead. Not underlining text is very important when submitting items for Web reading. Underlined text on the Internet infers a link to another Web site.

■ Avoid using ALL CAPITAL LETTERS IN A SENTENCE. It slows the reader down. It also takes up more space.

■ Stay away from using more than two fonts, and use a serif font, (such as *Times Roman*) for text and a sans serif font (such as *Helvetica*) for the headlines.

Writing is hard work. Give your interesting and well written papers the greatest chance of success by making them look well.

–Graphic Designer, Megan M. Murphy

Resource: *The Mac Is Not a Typewriter* and *The PC Is Not a Typewriter* by Robin Williams

“ As H. G. Wells said, “No passion in the world is equal to the passion to alter someone else’s draft.” ”

(6th ed., 2000, \$36) contains over 700 pages of detail on grammar and punctuation, etc. Unfortunately, it devotes only two pages to mathematical material.

Do you want material more specific to mathematics and statistics? At long last, the American Mathematical Society has updated Ellen Swanson’s *Mathematics into Type*, (1999, \$24). It is excellent and modern, even discussing T_EX and L^AT_EX. A chapter titled “Mathematics in Type” appears in the *Chicago Manual of Style* (14th ed. 1993, \$45), which the ASA recommends to authors preparing papers for its journals. Biostatisticians especially will profit from consulting Tom Lang and Michelle Secic’s *How to Report Statistics in Medicine* (1997, \$40), which Nadine Martin praised in her March 1998 review in the *Journal of the American Statistical Association*.

Most of what we write is ‘published’ only by us printing it from our computer or creating a Web page. Thus, today’s most skillful writers must learn how to do their own page layout work—choose the fonts, set the page styles, and position the tables and figures. Here, too, there are principles to guide us. Robin Williams (not the actor) and her various co-authors cover the basics in several wonderful books and articles; see www.ratz.com/robin/books.html.

Editing

As H. G. Wells said, “No passion in the world is equal to the passion to alter someone else’s draft.” Accomplished technical writers have learned to ruthlessly edit their own material, and they should not back off when they edit for others. Serious authors know that tough editing always improves their material. Good editors know how to avoid being overbearing. The collaboration succeeds when both parties set their egos aside, and when the editor accepts that authors who write well deserve the final say.

The ‘complete and sufficient statistician’

Many of us were attracted to statistics because we liked courses in mathematics, science, and computing—all fields that are based on theory and principles but that do not develop writing skills in their curricula. In fact, you may have purposely avoided humanities courses because they were more ‘subjective’ and required so many essays and term papers. Writing at a professional level is difficult and frustrating for many of us. Nevertheless, to be a ‘complete and sufficient statistician,’ you must become a skilled writer. This task will be easier if you understand that technical writing is not part of the humanities, but rather is a science based on theory and principles. ■

Dr. O’Brien directs the Collaborative Biostatistics Center within the Department of Biostatistics at the Cleveland Clinic Foundation. He thanks his daughter, Jamie, for her merciless editing of this piece.

1947

Statistics faculty members (L to R) Paul G. Homeyer, Oscar Kempthorne, Raymond J. Jessen, and George W. Snedecor. (Iowa State College Parks Library Special Collection) *George Snedecor was President of ASA in 1948.*

