Some Introductory Notes on Writing

1 Introduction

In your career as a graduate student of statistics, and later as a professional statistician, you will do lots of technical writing. Examples of some forms you can look forward to (or have already tried) include:

- Research Reports for Stat Practice (your 'last' MS project at Carnegie Mellon)
- Final Papers for Advanced Data Analysis (your 'first' PhD project at Carnegie Mellon)
- Dissertation Chapters (and for that matter, the whole PhD Dissertation)
- Journal Articles
- Class Notes if you are teaching
- Homework Assignments
- Homework Solutions
- Etc.

…and along the way you will also need to write professional (as opposed to personal) email, interim progress reports, slide presentations, letters of recommendation, job application cover letters, etc. etc.

Each form (or, as the English teachers would say, each 'genre') has its own conventions, and you will need to learn those conventions as you encounter each form in your work. The good news is you don't have to learn them all at once. The bad news is you do have to learn them all eventually.

I have long felt that paying attention to writing conventions is essential to being a good scientific or technical writer. I have come up with a number of aphorisms about writing. If you ask a student who has had me for a class that required writing, he or she will tell you that I alway say things like

> **Technical writing should be invisible. In songs, poetry, short fiction, or novel writing, it can be wonderful to write memorable phrases and sentences. However, after I'm done reading your technical writing, I don't want to remember any particular word or phrase or organizational device you used. I only want to remember the ideas you were trying to communicate.**

or perhaps

> **Meet the expectations of the reader. If you meet those expectations, the reader will be free to concentrate on the ideas you are trying to communicate. If you do not meet those expectations, the reader will be distracted by the disappointment, and will not be concentrating on your ideas.**

If you ask one of my PhD students (past or present) they will tell you I always say

> **Technical writing—even an *Annals of Statistics* paper—must tell a story. Therefore it needs a strong plot. It needs heros and villains. It needs dramatic conflict that the reader aches to see resolved, and it needs to (at least partly) resolve that conflict by the end of the paper.**

These aphorisms don't always communicate as clearly as I would like, and so I have looked for books and articles that say what I am trying to say about writing in a more dispassionate, less idiosyncratic way. One of the best is an undergraduate technical writing text,

Alley, M. (1996). *The Craft of Scientific Writing, 3rd ed.* New York: Springer-Verlag. Web resources at http://www.writing.eng.vt.edu and http://www.me.vt.edu/people/faculty/alley.html.

Even though it is an undergraduate text, it has been used for graduate courses in technical writing, short courses in technical writing for professionals, etc. There are guidelines for everything from cover letters to PhD dissertations based on this book. It is becoming the Strunk & White[1] of technical writing.

I freely borrow from the Alley text (and the websites above) below in discussing writing.

It is already hard for a native speaker of English to figure out what makes English technical writing good. It is even harder for a speaker of English as a second language (ESL). I have not been happy with past resources I have found for ESL writers, though recently I stumbled on the following book which has some material aimed at ESL writers that may be helpful:

Hacker, D. (1998). *A Writer's Reference. 6^{th} Ed.* NY: St. Martin's Press. Web resources at http://www.dianahacker.com/writersref/.

(I haven't had a chance to look at the book or webpage myself yet. I'd be interested in what you think of them.)

[1] Strunk, W. Jr. & White, E.B. (2000). *The Elements of Style. 4th Ed.* BN Publishing. [Editions going back to 1918 are available on-line. This is the classic guide to style in general English writing.]

2 Approaching Writing

There are two major writing considerations in approaching any writing task. *Writing constraints* are the external conditions within which your document must 'live'. *Writing characteristics* result from the choices you make in writing; they affect how readers assess and interpret your writing.

2.1 Writing Constraints

Writing constraints are the external conditions under which your document must live. Who will read it? When? Why? What do you want to accomplish with the reader(s)? Etc. There are three major writing constraints:

- *Audience*
  - Who are they
  - What they know
  - Why they will read
  - How they will read
- *Purpose*
  - To inform
  - To persuade
- *Occasion*
  - Format
  - Formality
  - Politics and ethics
  - Process and deadlines

Whether you are writing email or a PhD dissertation, you have a better chance of writing efficiently and effectively if you stop and consider these constraints—early and often—as part of your writing task.

Some writing constraints are immediately obvious from the context in which you are writing. For example: an *Annals of Statistics* paper has a certain format; a project paper for this class has two primary audiences—me, and the outside investigator you are working with. But other constraints have to be worked out each time you write—and even these examples are not as obvious as they seem—and there are really no simple formulas to relieve you of thinking about them.

2.2 Writing Characteristics

The characteristics of your writing—the ways people would describe your writing—determine how readers assess and interpret your writing. If your writing can be described as sloppy, readers will tend to think your ideas are sloppy. If your writing can be described as stiff or uninspired, readers will tend to think of your ideas that way as well. Sloppy or uninspired—either way readers will tend to discount your ideas before you've had a chance to communicate them fully. If your writing is clear, precise and alive, then readers will tend to think of your ideas that way as well.

Your writing acquires its unique characteristics—its personality—from the choices you make as a writer. These choices fall into three broad categories:

- *Content*
  Content is the subject matter of your writing. It is the thing that distinguishes scientific writing from newspaper, magazine, fiction and poetry writing. In a scientific paper the content is usually a highly constrained and specialized area of your work: theorems and proofs, data analyses, developing, testing or illustrating methodology, scientific experimentation, etc.

- *Style*
  Style is the way you communicate to your audience. It usually consists of three aspects:

  - <u>Structure</u>
    Organizational signposts for the reader. Documents, Sections, Subsections, Paragraphs and Sentences all have: Beginning, Middle, End.
  - <u>Language</u>
    Also called 'Voice'. Some positive adjectives for Voice in technical writing: Precise & Clear, Forthright & Familiar, Concise & Fluid.
  - <u>Illustration</u>
    Figures, Tables, Photographs, Drawings, Diagrams.

- *Form*
  Form is how you put marks on the page, and how you use those marks to build up units of meaning:

  - <u>Format</u>
    - *Typography:* Size and style of type (letters, numbers, symbols).
    - *Layout:* Physical appearance of paragraphs, sections, pages, etc.
  - <u>Mechanics</u>
    - *Grammar:* The syntax of English.
    - *Usage:* The semantics of English.
    - *Punctuation:* Traffic signs in a sentence.
    - *Spelling:* The bricks and mortar of meaning.

3 The Writing Process

Here are some typical stages in a writing project.

- Getting in the Mood
- Writing the First Draft
- Revising, Revising, Revising
- Finishing

Writing takes a certain thinking style, just like doing proofs, writing R code, or holding office hours require their own styles. Writers need to be thinking about their subject matter, they need to be imagining how their audience will read their work, and they need to be self critical, *both* at the fussy, anal level of finding spelling and grammar errors, *and* at the higher level of uncovering flaws in their reasoning, logic, or evidence as they write. Yet they must also have the courage to write confidently, to engage the reader in their subject matter. The way to get good at getting in the mood to write is to do a lot of writing, so you get a lot of practice at adopting the thinking style required for writing.

It is very easy to underestimate how long it will take to write the first draft. Aside from the 'blank piece of paper' syndrome, there is usually more than one way to organize a piece of writing, and it may take time—and a few discarded attempts—before you arrive at an organization that you are comfortable with. Even after you have decided how to organize your writing, it will take a long time to hammer out the first draft. For example I thought this document would take me about 1.5 hrs to knock off: I know my audience, I know how I want to organize it, and this is mostly stuff that I can write "out of my head" because I have been thinking about it lately… *So far it has taken over twice that long (in first draft!), and I am not done yet.*

It is also easy to underestimate how much revision, and how many revisions, are required after the first draft. In the first place, you wrote it the way you wanted to say it the first time, why should you have to change it? Getting someone else to read your writing and give you unvarnished feedback about it will quickly convince you that your first draft is not enough to communicate with many other people. In the second place, after the third or fourth revision, you will hardly be able to read your own work anymore. Either you will be bored stiff of it and unable to concentrate anymore, or you will be so familiar with what you are trying to say that you can no longer see missing or misspelled words, poor organizational choices, etc. This is a good time to put the manuscript away for a while, or have someone with a fresh set of eyes read it and comment.

Because it is so easy to underestimate the time a first draft takes, and the amount of revision that is required, writing for a large project, like a data analysis project report, should usually start early—long before the project is done. Problems uncovered in the writing process may suggest additional data analysis avenues to explore, and new data analysis discoveries may suggest a different way to write the report. Ideally, writing and research interact with and improve each other.

Finally, there is the problem of finishing what you write. If you take the revising task seriously, you will find there is always one more revision to make: one more mis-spelled word, one more graph or figure to convince the reader, one more late-breaking result from data analyses, one more way to re-organize the discussion section at the end of the paper, etc. At some point you have to say, "done, finished... this is the best I can do and now it should be somebody else's problem". Not deciding the writing is "finished" is a primary reason for delay of dissertation defenses, lack of publications for junior faculty, etc.[2]

4 Resources

I mentioned the Alley book above. It is my favorite; you should log on to `amazon.com` right after class and buy it. You might also want to buy the companion volume *The Craft of Editing*.

Alley's main resource page, `http://www.writing.eng.vt.edu/`, has examples, exercises, course notes, and Powerpoint slides for teachers. The slides elaborate on most of the topics in this introduction:

```
01-introduction.ppt
02-formatting.ppt
03-mechanics.ppt
04-editing.ppt
05-structure.ppt
06-language.ppt
07-illustration.ppt
```

Alley's other webpage, `http://www.me.vt.edu/people/faculty/alley.html` contains additional examples and course notes. The books contain much more.

Although they are not as comprehensive as a textbook (nor are they intended to be!), I like the following two articles a lot:

Bem, D. (2002). Writing the empirical journal article. In Darley, J. M., Zanna, M. P., & Roediger III, H. L. (Eds.) (2002). *The Compleat Academic: A Career Guide.* Washington, DC: American Psychological Association.

Gopen, George D. and Swan, Judith A. (1990) The science of scientific writing. *American Scientist*, **78**, 550–558.

Bem discusses the overall organization of an article, what should go where, and why. There are also brief discussions of style, grammar and the like. Gopen & Swan clearly and precisely describe a challenge of writing that is "above" spelling and grammar but "below" overall organization of a paper or paper section: organizing sentences to help readers concentrate on meaning by meeting readers' expectations about where different kinds of information should be in the sentence. (Yes, I know, this sounds very dry. But it is not! [People say the same thing about statistics, and we know better, right?])

[2] Sometimes this is called a "file drawer problem": "Smedley has a lot of good papers, it's just that they're all in a file drawer in his office. He never declares them done and sends them out for publication!"