

A Core Dump on Writing – Brian Junker

“...no idea is fully formed until it can be communicated... the organization required for writing and speaking is part of the thought process that enables one to understand material fully.” The Boyer Report, Carnegie Foundation, 1998.

A writing problem (that is, difficulty thinking what to say or how to say it in a paper) is most often symptomatic of a thinking problem (that is, insufficient ideas, or insufficient effort analyzing the ideas you do have and breaking them down into communicable pieces).

Writing is not the same as doing a “core dump¹” of all the information one has, nor is it the same as producing a “dear diary” account of all of the activities one engaged in, in order, to produce the research. In writing, we must formulate a story and tell it in a focused and compelling fashion, eliminating all the parts (however clever) that are not part of the story, and providing the information that the reader needs, when the reader needs it, to follow the story and get the ideas.

Writing is teaching, no matter who or how accomplished the reader is.

I also find writing to be a very effective way to learn: as I struggle to express an idea in writing, I learn the idea better.

Academic writing is judged on

1. Clarity
2. Novelty
3. Contribution

Of course 2 and 3 are important for your career. But in fact, they are useless, if you cannot be understood. So, in fact, peoples’ judgement of you (and your ideas) will rely more heavily on 1 than 2 or 3. If you write clearly, the novelty and contribution of your ideas will shine through.

There are a few principles.

1. Put the important stuff where a reader in a hurry will see it. [EVERY reader is in a hurry!]
2. Don't make the reader work harder than necessary to understand it.
3. Unlike other forms and purposes of writing, make the writing 'disappear' in the readers' eyes, so that only the ideas remain.

Technical writing (in statistics, or I imagine, many other fields) should be approached at several grain sizes. I will summarize here what I think the main issues are at each grain size.

Sub-sentence level:

¹ The irony of a “core dump” on writing warning against making your writing a “core dump” does not escape me!

1. No choice of words or turns of phrase should surprise or distract the reader in any way. The idea is to meet the reader's expectations so that the sentences fade away and only the ideas remain.
2. To determine whether a sentence meets this criterion, delete part of the sentence (perhaps a "suspect" part) and imagine how a typical reader would complete the sentence. If it is the same as the way you wrote that part, you are fine. If not, change it to how the reader would complete it (without changing the information it contains, of course).
3. Obvious things like misspellings, usage errors, subject-verb-agreement, malapropisms, and so forth, must be attended to. However, there are a couple of other related issues that are unique to technical papers:
 - a. The main goal is to never distract the reader from the *ideas* in the paper by calling attention to the *effort* (successful or not) you took to write the sentence.
 - b. Word choice should never be fussy, clever, or unusual. Choose words the reader expects to see. This does not mean, be boring!
 - c. Be maniacally consistent about technical or quasi-technical terminology. Once a term is used in a technical way to mean a certain thing, *ALWAYS* use the same term to refer to that thing. Never, ever, use multiple equivalent terms to refer to the same technical object. If need be, point out that other terms are in use in the literature, but be maniacally consistent about the term you use in your paper.
4. Sentences do not have to be boring, but they should not be intricate or ornate. They should be constructed, as Gopen and Swan say, so that information is presented when the reader needs it.

Sentence level:

1. A consequence of the ideas in the Gopen and Swan paper is that the information you want the reader to take away should be put in the topic position (beginning of the sentence) or stress position (end of sentence). Everything in the middle of the sentence is just details for the (few) super-interested readers.
2. Of course there is much more to the Gopen and Swan paper, mostly about managing the reader's working memory load by placing information where the reader needs it.

Paragraph and Section level:

1. The Gopen and Swan ideas can be applied to paragraphs and to sections in the same way. When applied to paragraphs, this improves skimmability and helps enforce the idea that a paragraph is an extended discussion of a single idea. When applied to sections also, this really improves the skimmability/readability of the paper.
2. To facilitate skim-readers, there should be enough information in the beginning line (and perhaps ending line) of each paragraph, that a reader who only reads those lines will get a pretty good idea of what's in the paper.

Storytelling and teaching:

1. Regardless of what you may think, readers are human beings, and human beings respond best to storytelling. So pick a story to tell. The best stories have heroes and villains, dramatic conflict that is introduced at the beginning, and that is resolved by the end of the paper. This doesn't have to be complicated – the “dramatic conflict” could be that there is a hole in the literature, and the “resolution” could be that you have filled the hole. It also shouldn't detract from the ideas in the paper – like everything else it should be a setting against which the ideas glitter. But it greatly helps you to write the paper, and helps the reader to be motivated to read it.
2. A paper is not a core dump of everything you know, or a ‘dear diary’ retelling of all the steps and wrong turns you took, to do the research. Once you have picked a story, anything that does not contribute to advancing the story should be eliminated from the paper. (Kill your darlings!)
3. Most of the writing you do in college (especially undergraduate) is “writing up” to professors who already know or mostly know what you are talking about. Most of the writing you do as a graduate student and in life is “writing down”, to peers who do not know as much about your subject as you, and/or to managers and other consumers who know even less. This means that you should practice writing as teaching – all good technical writing has a pedagogical component, to help the reader get up to speed quickly (and often transparently) to a level at which the ideas in the paper can be understood.

Paper format and organization:

1. These things vary from journal to journal (for academic papers) and from situation to situation (for other kinds of papers).
2. What is important about paper formats is the same thing as is important about constructing sentences, paragraphs, and sections. They mark where information is so that the reader can find the information he/she needs to understand the paper – with a suitable organizing structure or format, the reader does not need to work harder than necessary to find the information he/she needs in the paper.
3. Paper formats, such as IMRaD, are easy to state, easy to teach, and easy to follow. It is important to know a few formats, and to know where each is appropriate, but they are not really difficult to master. Similarly, bibliography formats vary depending on the venue for the paper, and they can be tedious, but they are basically mechanical things that are relatively easy to master. Writing software such as LaTeX, that generalize content across formats, can be very helpful in managing both.
4. A typical skim reader will read the paper by examining the following elements of the paper, in more or less this order: title, abstract, conclusion, skim the paper (first lines of paragraphs), read the paper in detail. If/when the reader loses interest, the remaining elements will be ignored (e.g. if the reader loses interest after the abstract, he/she won't read the conclusion, won't skim the paper, and won't read the paper in detail). So it is very important that each of these parts of a paper contain a compelling summary of the paper, to attract the reader's attention and to accurately represent what is in the paper.

If you have great technical ideas and you write a paper about them using the above guidelines, the ideas will “pop out” and the reader will take away what you want – the ideas! If you have mediocre technical ideas, or even no novel ideas, and you write a paper about them using the above guidelines, the paper will be stilted and boring. So writing this way can also serve as a way to judge how exciting your ideas are.

Like all rules of social convention, and unlike mathematical theorems or laws of nature, the above rules and guidelines can be broken if you have a sufficiently good reason rooted in helping the reader to retain your ideas. But you should learn to write by the rules first, and then experiment with breaking them when you think you need to. (you will find you don’t often need to!)

One might ask whether such a tight prescription makes creativity or individual style impossible. I would respond in two ways. First, read the following two sonnets:

From fairest creatures we desire increase,
That thereby beauty's rose might never die,
But as the ripener should by time decease,
His tender heir might bear his memory:

But thou contracted to thine own bright eyes,
Feed'st thy light's flame with self-substantial fuel,
Making a famine where abundance lies,
Thy self thy foe, to thy sweet self too cruel:

Thou that art now the world's fresh ornament,
And only herald to the gaudy spring,
Within thine own bud buriest thy content,
And, tender churl, mak'st waste in niggarding:

Pity the world, or else this glutton be,
To eat the world's due, by the grave and thee.

- William Shakespeare

Welcome to the endless high-school
Reunion. Welcome to past friends
And lovers, however kind or cruel.
Let's undervalue and unmend

The present. Why can't we pretend
Every stage of life is the same?
Let's exhume, resume, and extend
Childhood. Let's all play the games

That occupy the young. Let fame
And shame intertwine. Let one's search
For God become public domain.
Let church.com become our church.

Let's sign up, sign in, and confess
Here at the altar of loneliness.

- Sherman Alexie

Second, I have taught graduate students and postdocs with very different individual styles to use guidelines like these, and you can still see the individual style of the writer, even when he/she is following the guidelines. You can also see this in the peer reviewing you've done in this class!