# Science writing
U. Minn.
Daniel Kersten

## General writing

### Introductory points

There is no shortage of advice on how to write. Take a look at:
http : // www.dcs.qmul.ac.uk / ~ norman / papers / good_writing / general_. principles.html for a simple summary of main points, together with nitty-gritty practical advice. Today's lecture will be a more personal view on good writing. But the principles are the same.

#### ■ Where your teacher is at

I have spent a lot of time critiquing my own writing and that of others. But I'm still learning. It is one thing to know what to do and another to actually do it. It is also one thing to do something well, and quite a different thing to explain how you do it well. I don't claim to always write well or to know how to write the best. But I do write better than I used to. I've learned from my advisors, journal reviewers, and from my own students. I continue to learn from self-evaluation. So I'll summarize my ideas and advice.

Two favorite sources:

Gopen & Swan: Provides an excellent discussion of general principles. Not a cookbook.

A Pocket Style Manual 2nd Edition. Diana Hacker: A good cookbook.

And even though this lecture is about science writing, you can learn much from reading "how to" books by fiction writers. One of my favorites is Stephen King's "On Writing". I think I learned as much from seeing how he managed to turn a non-fiction book into a “page-turner” as I did from his explicit advice on writing.

It helps to find writers that you like, read their work, and then ask yourself what makes their writing work for you. One of my favorite technology writers is David Pogue, formerly of the New York Times. Goofy at times, but he does catch my interest. But more importantly, he does a good job of anticipating my next question. I'll talk about this below in the context of creating and fulfilling expectancies.

And pay attention to your computer's grammar checker ! Jerry Pournelle, a well-known science fiction writer suggests the use of a grammar checker in his advice to beginning writers (http://www.jerrypournelle.com/slowchange/myjob.html). Also, see: http://papyr.com/hypertextbooks/grammar/gramchek.htm

I will talk about science writing, rather than other kinds of writing, for three reasons: 1) This is a science course and you'll get practice doing it; 2) I'm your teacher, and it is the only kind of writing that I know; 3) At a general level, the kind of writing doesn't matter too much, because the main principles generalize to other types of communication. Many of the principles also apply to video presentations. So even if after this course you never write about scientific results again, hopefully you will have learned something about clear communication that will pay off later.

## ■ Why bother trying to write better?

Why bother learning to write well? After all, you might argue, a good scientific result should stand on its own. There are a few cases where just a hint of a major discovery is so motivating that exactly how the information is conveyed is relatively unimportant compared to the result. (The structure of DNA, Einstein's theory of special relativity, the proof of Fermat's last theorem.) But for most of us who write up scientific results, good writing is essential to success. Here are a few reasons.

The vast majority of scientific papers contribute only incrementally to knowledge. Knowledge production is so voluminous that a published paper may get thoroughly read by only 2 or 3 people. On the one hand, this figure is depressingly low. On the other hand, this is an opportunity to gain 2 or 3 advocates. An incremental contribution can be amplified or minimized by the quality of the writing.

While it is certainly true that, at least in general, good writing is altruistic and polite in that it respects the reader, good writing is also selfish--someone else who writes better can get the credit. I know of papers in my field whose (important) results never gained widespread recognition, and were superceded by a later paper which was written better (or is it “better written”?). The author of the second paper is the one that gets remembered in citations and textbooks. It isn't just the clarity of exposition, but often the depth and breadth of the later paper that leads to broader readership. However, this depth and breadth can emerge in part from an effort to clearly communicate thoughts and intuitions that are otherwise to vague to be useful. Thus, probably the best reason to practice writing well is that it will help you think more clearly. And if as a scientist you think more clearly, you may end up discovering more. As Gopen and Swan put it:

> "*Improving the quality of writing actually improves the quality of thought*".

There are many times that I've thought that I've thoroughly understood something, only to be stumped when having to explain it.

As an aside, the same goes for programming. It is one thing to think you have a good idea, and quite another to design a computer program that implements it. There are many apparently "good" theories in cognitive science that stumble at the implementational stage. That is why I teach this course using computers and programming. Producing clear writing is like good programming--they both force you to make your assumptions and theories sufficiently explicit to be communicated and tested.

## ■ Determine your audience

Before you put your fingers on the keyboard, think about your potential audience. I find it useful to think of a “specialty gradient” going from general to specific audiences:

general public, kids, undergraduates, graduate students, scientific layperson, scientific professional, scientific colleague in your field, scientific colleague in your speciality.

Where does your audience fit on the speciality gradient? It isn't always easy to figure this out, and it may take some research and discipline.

Because of specialization, there are significant differences amoung members of the category of professions that know and use science:

Social scientists, physical scientists, biological scientists, scientific professionals, engineers, grant reviewers, scientists in industry, business users of science.

You may want to write to reach two audiences. There are good and bad ways of doing this too.

*Tip: Write as if to someone you personally know who represents your audience. If you have a chance, have that person* *read a draft. Encourage the person to be frank about what is not understood.*

As Gopen and Swan put it:

"*If the reader is to grasp what the writer means, the writer must understand what the reader needs*"

***So to be clear, I will tell you who your audience is for this course. In your final paper, your audience is not me! The audience is your peers!***

### ■ Good writing is hard work for most of us

Good writing is hard because it is the result of clear thinking, and clear thinking takes work. Good writing is also hard because it requires thinking about expectancies and logical flow at multiple levels of abstraction.

If there is one "take home" message, it is that you do learn best by doing, but only if you get feedback. In this course, you'll get feedback from me and your peers. But ultimately, the most important feedback is that which you generate yourself by reading and revising what you've written, often many times.

Let's understand what is meant by expectancy, logical flow, and levels of abstraction.

## Generating and fulfilling reader expectations

The 18th century English historian, Edward Gibbon wrote:

"*The powers of instruction are seldom of much efficacy except in those happy dispositions for which they are almost superfluous*"

There is a general principle here that applies beyond teaching to wide range of human activity, to humor, creativity in art, music, and fiction, non-fiction, and science writing. Effective communication means that the recipient needs sufficient context to understand what is being said, and needs to have an expectation for the kind of phrase, sentence, paragraph or section that will come next. What comes next provides closure.

A joke is a caricature of what you want to accomplish when communicating. And to get nerdy, what we've learned about predictive coding by the nervous system, also provides a good analogy.

Avoid spurious information. But "spurious" can be a function of the audience. If details obstruct understanding for most of the audience,  but are important to a minority, put the details elsewhere (e.g. mathematical details are often put in an appendix). Determining your audience helps determine what goes where.

To use Gopen and Swan's terminology, the *opening* (of a sentence, paragraph, section, or paper) sets context,  motivation and general expectation. The c*losing* (stress position) brings fulfillment to expectation, i.e. closure.

When you write up your final projects, you will follow a particular section form, not because "that's just the way it is", but rather because the form makes explicit the role of expectation. Expectation is why there is often a standard form in science papers: Intro, Methods, Results, Discussion. I'll say more about the Abstract below.

Violating the section order of a paper can disregard the tacit structural knowledge of our audience.

Many scientific journals have changed their format to put the detailed Methods at the end. Why?

## "Logical flow" & expectancy

### ■ The importance of smooth "transitions"

Any piece of writing should carry the reader along without extraneous mental intrusions into the flow of the description. This is true of non-fiction and fiction writing. The maxim in fiction is to avoid intrusions that interrupt the "suspension of disbelief". Bad grammar is an intrusion that can distract the reader, even when the sentence meaning is clear.

For science writing, the way to do this is aim for logical flow.

What is "logical flow"?

From Gopen & Swan: "Put in the topic position the old information that links backward; put in the stress position the new information you want the reader to emphasize".

The topic position is typically at the beginning to set context. The stress position is usually at the end to provide the new information. Old to new.

*Tip: Watch for ambiguous referents!! Especially at the beginning of a new paragraph. Here's an example of ambiguous referents*

"Put the old information in the topic position. The topic position is typically at the beginning. This will help you write better."

Ambiguous referents often crop up during revisions, when you decide to insert a new sentence. If you were paying attention, you might have noticed the above transition to: “For science writing, the way to do this is to aim for logical flow.” The flow was fine until I inserted an additional sentence: “Bad grammar is an intrusion that can distract the reader, even when the sentence meaning is clear.”

What did "this" refer to? Don't assume the referent is as obvious to your reader as it is to you. If your reader requires even a brief pause and re-reading of the previous lines to sort out the referent, rework the sentences.

### ■ Logical flow at multiple levels

Logical flow operates within and between units (a unit can be a sentence, paragraph, or section). Use the lead subject word to link with previous sentence or paragraph. Or you could set up a list of expectations at the beginning of a unit, and then follow through with the new information in the same sequence. At a coarser scale of abstraction, a paragraph or section plays the role of a unit. You use the opening concept or paragraph to link backwards to presumed earlier knowledge of the reader, or the previous section of the paper. Use the closing paragraph to emphasize/summarize the new information.

For another nerdy metaphor relevant to a topic in this course, think about self-similarity across scales. The self-similarity in writing lies in the abstract notion of following the old with the new across levels, from sentence to paragraph to section.

An outliner can help to make the logical flow at multiple levels explicit. For example, if I was practicing what I preach, this subsection should have a logical flow, and if you close the cells, the section should also have a logical flow. You can also use parts of an outliner like a scaffolding that gets removed once the structure of the text has been built. Sometimes I use more outliner levels while writing than I need for the final draft. It helps me organize my thoughts at multiple levels. But once I've figured out the structure, the text stands on its own without an overdose of subsubsections, subsections, and sections. Too many of these can disrupt the flow. Whether I decide to use an outliner depends on how hard I have to think about large scale structure organization.

# Writing a scientific paper

## Paper structure

### Title

The title should be as informative about the content and ideally, the conclusion of the paper. And the title should entice readers.

E.g.

**Good:**

Activity in primary visual cortex predicts performance in a visual detection task

Why is snow so bright?

"Shape perception reduces activity in human primary visual cortex" (one of my papers)

**Bad:**

"Aspects of phase" (sadly, also one of mine)

But sometimes one can get away with "vagueness", e.g. "Pursuing commitments" (Nature Neuroscience, Shadlen, 2002). This sounds like a "must read". Why?

### Abstract

Mini-version of the whole paper. This is still a unit of discourse with logical order: set context & motivation, raise question, give answer. Minimize details from the methods, unless the method is a primary contribution.

### Introduction

***Motivation, motivation, motivation.*** Articulate the question for the target audience. You need to tell your readers what is known and what is not known in a way that motivates them to continue reading. Bu the end of the Introduction, the reader should be thinking: “Hey, this is an interesting question. I wonder what they discoverd and how.”

**The "Funnel" principle**

Start the introduction by motivating with the big picture, and gradually focus in on the scientific question/hypothesis being tested. The metaphor is that your paper should provide a funnel into which readers with diverse backgrounds can be brought into an appreciation and understanding of the focused scientific question of the paper. Adjust the "width" of opening of the funnel to the diversity of the audience. Usually the length of the "funnel"--the allotted writing space--is fixed, determined by scientific journal convention.

Here is another way of thinking about the "funnel principle". As we noted above, the average number of readers per scientific journal article is small. Further, the number that actually make it all the way to the end is even smaller. The introduction can serve to motivate and reach an opening audience that might be different than your closing audience.

By the end of the introduction, the funnel is narrow--it is where you, the expert on the material, are at.

**Example**

Here’s an example of a very rapid funnel constriction that works for the right audience:

"Soon after I started physiological research, I was lucky enough to make an interesting, but as it turned out unoriginal discovery (Barlow, 1950). I was repeating some experiments that Hartline (1938, 1940) had done on the frog's retina, with the idea that the very large receptive fields he had discovered might not be simple spatial integrators of light, as he had suggested, but might have some form of pattern selectivity. The experiment was to measure the threshold for eliciting impluses from a retinal ganglion cell as a function of the area of the stimulus spot. If spatial integration occurred, and the sensitivity over the receptive field was uniform, ..."

From: Barlow, H.B. (1982), "Perception: What quantiative laws govern the acquisition of knowledge from the senses?"

**The punch-line of the introduction**

Put the context at the beginning and the "punch-line" at the end--the stress point of the section. The "punch-line" here is the crucial, exciting scientific question, i.e. end the introduction section with a clear statement of the hypothesis to be tested. By the end of the introduction, your readers should be waiting with bated breath to find out how you've answered the question, and what your answer is. Don't give away the whole story. But there are no hard and fast rules. Sometimes a titillating preview of the result can be a good idea--the equivalent of a movie trailer. But a preview is already in the Abstract, and you may not need to repeat it. Again, consider the diversity of the audience. And whether you want to allow for the readers that will quit after the Introduction because they just wanted to get the gist anyway.

A particularly informative Introduction describes the question in such a way that the reader can anticipate how a graph of the results will look if the hypothesis is confirmed or falsified. I personally like papers that show graphs of possible results. (For example, see Figure 2 in: http://www.cell.com/current-biology/retrieve/pii/S0960982208012517)

## ■ Methods

Explain *how* you answer the question posed in the Introduction. A brief overview of this explanation may go in the Introduction, but the details go in the Methods.

## ■ Results

In this section, the context is the data and analysis. The new information is the interpretation in the light of the question(s) posed at the end of the Introduction. Many good results sections have clear figures that carry and summarize most if not all of the results, and the writing centers around explaining the figures.

*Tip: Work on making sure that the scientific conclusions can be drawn naturally and clearly from the figures and the figure captions. Then writing the Results section will come naturally. As discussed above, the reader should have an expectation of what to see in the data or figures if the hypothesis is true vs. false.*

One could, and I probably should devote a whole lecture just to good figures. Instead, I highly recommend the books by Tufte (see References).

## ■ Discussion

Describe the broader implications of your results. This is the place for bold predictions, and for humble speculation. It is also a good place to anticipate the criticisms of reviewers, and thus of your ultimate readers.

## ■ General principles: Economy, symmetry and elegance

Principle of (translational) symmetry. If points 1,2,3 are raised and highlighted in the Introduction, those points should be followed up in the same order in the experiment, results, and discussion. This is a special case of creating and fulfilling reader expectations in a logical sequence.

*Tip: When introducing a key concept with a set of descriptive words, use the same descriptive words when referencing it later. (oops, ..when referencing the* ***concept*** *later)*

If jargon is necessary, pick one word per concept and stick with it. Don't use multiple jargon words that mean the same thing. Sometimes we think that substituting synonyms may help the reader to better understand the terms. It usually doesn't work. For highly familiar everyday words, do use synonyms to avoid echoes which themselves are distracting. Above I've used "expectancy", "logical flow", "transitions", "closure", "punch line", and "stress point". They don't all mean the same thing, but could I have simplified my terminology? If so, would it have helped?

## Sentence and paragraph structure

Let's look at how to achieve logical flow at the finer-grain level of sentences and paragraphs.

### ■ Rhetorical principles from Gopen & Swan:

Subjects should be followed as soon as possible by their verbs, otherwise the reader loses the logical flow within the sentence.

Information to be emphasized should be placed at "points of syntactic closure". E.g. stress positions in a sentence are typically at the end.

i.e. "save the best for last"

Place the person or thing whose "story" a sentence is telling at the beginning of the sentence, in the topic position.

Place context or "old information" at the beginning for linkage backward, and contextualization forward. I.e. work towards good transitions.

A unit of discourse (e.g. sentence, but regardless of size) should serve a single function or make a single point. This applies to paragraphs too.

*Tip. You may have several points you want to make in a paper. But it is rare that all points have equal importance. Determine the priorities. There should always be a clear "take home message".*

Readers expect the action of a sentence to be articulated by the verb. Use informative verbs (as in information theory. "is, "are presumed to be", "has" are low information words.) Good writing is in a mysterious place between low and high entropy.

Again provide context before asking reader to consider anything new ("motivation, motivation, motivation")

Try to ensure that the relative emphases of the substance coincide with the relative expectations for emphasis raised by the structure.

(Goodness, I just re-read the above sentence for the first time in a year or two. Took me two readings to understand it. How about you?)

In other words, avoid the problem **The Foundations** sang about in 1968: "Why do you build me up ( build me up) Buttercup, baby. Just to let me down (let me down ) and mess me around"

### ■ Economy, "redundancy reduction"

I often review scientific journal submissions that are too wordy. One of the best ways of avoiding unnecessary verbiage is to follow the principle of logical flow at multiple levels. There are too many sentences (or paragraphs or subsections) that just seem to be elaborating on what went before, or that are tangential. In other words, the article fails to achieve a good logical flow. One could advocate a kind of minimax rule: "maximize information transfer with the minimum number of units", but I don't think advice is that useful, any more than a cost function is useful without some algorithm to minimize it. But I do think something like a minimax rule is the end result of working towards logical flow.

But, don't over-do-it! I.e. avoid telegraphic writing.

Great quick overview here:

A Pocket Style Manual 2nd Edition. Diana Hacker.

# Are there exceptions to the rules?

Judging writing quality is certainly subjective, but the desirability of clear smooth communication is universal. When you have a peer critique your writing, you might get ticked off -- “How could they be so stupid! My point was perfectly clear”. But like advice given to retailers regarding the opinions of the customer, think “the reviewer is always right”. Of course, the reviewer isn’t always right, but the fact that they didn’t understand something provides you with information that can improve your writing. When you submit your final draft, you will also submit a “Cover Letter”, that describes your response to the comments of your peer reviewer and your instructor (me). I will play the role of an editor. Avoid defensive cover letters. When a reviewer didn’t understand a point or was wrong about one, use your cover letter to explain how you revised the paper to avoid similar confusions by future readers. Sometimes a reviewer missed a clearly written explanation. Point that out to the editor.

One of the main reasons to have someone else read your drafts is because your own knowledge of the topic is too good. When I read a draft of a paper written I wrote a year or two earlier, I’m sometimes aghast at how opaque the writing is. The reason is that I made the mistake of failing to realize how much “back story” I had at the time that even an expert in my field may not have. After a year or two, I’ve lost my own “back story” and read the draft from a more representative point of view. We can get immersed in the topic during research and writing, and for some odd reason believe that everyone else is right with us.

Are there exceptions to striving for clear, transparent transitions? From a pedagogical perspective, there can be good reason to leave some of the work of unraveling the content to the student. Mathematics and physics textbooks are good examples. But you still want to strive for a logical sequence, avoiding spurious information, even if your reader has to think hard to fill in the details.

# Good and bad writing

## Examples of good writing:

Scientific audience:

Some of my favorite writers in cognitive and neuroscience are: Horace Barlow, Gordon Legge, Anya Hurlbert, John Hopfield, Denis Pelli, Alan Yuille.

And for the scientific layperson: George Gamow, Freeman Dyson, Steve Pinker, and Matt Ridley. I especially liked Bill

Bryson's "A Short History of Nearly Everything", 2004. Bryson is particularly good at local, i.e. sentence to sentence, transitions. Almost too good--I sometimes lose the big picture.

### Examples of bad writing:

Gopen and Swan provide some examples of bad writing, and it is well worth the read.

Also take a look at:

http://www.bulwer-lytton.com/about.html

### For more on writing, both good and bad:

http : // kimberlychapman.com / essay / badwriting.html

http : // www.developsense.com / GuidelinesForBetterWriting.html

http : // www.dcs.qmul.ac.uk / ~ norman / papers / good_writing / general_. principles.html

And as mentioned above, my favorite is: **"On writing: A memoir of the craft" by Stephen King**

- **And for excellent practical advice from one of our colleagues see:**

http://psych.nyu.edu/pelli/style.html

### Some class examples

Examples shown in class

## Next week's topics

- **Perceptual integration**
- **Object recognition**

## References

Hacker, Diane. A Pocket Style Manual 2nd Edition. (Third edition from Amazon)

A good cookbook. Short, handy, good examples, with the most important and useful stuff at the beginning.

Gopen, G.D., & Swan, J.A. (1990). The Science of Scientific Writing. *American Scientist, 78*, 550-558. (See American Scientist online, html)

Provides an excellent discussion of general principles. Not a cookbook.

Tufte, E.R. (1983). The visual display of quantitative information. (p. 197). Cheshire, Conn. (Box 430, Cheshire 06410): Graphics Press.

Tufte, E.R. (1990). Envisioning information. (p. 126). Cheshire, Conn.: Graphics Press.

Tufte, E.R. (1997). Visual explanations : images and quantities, evidence and narrative. (p. 156). Cheshire, Conn.: Graphics Press.