

FLORIAN HARTIG, UNIVERSITY OF FREIBURG

RESEARCH SKILLS

AN INTRODUCTION TO THE CRAFTS OF A SCIENTIST

Contents

1	<i>Logic and philosophy of science</i>	4
2	<i>On top of the published literature, and finding a good scientific question</i>	6
3	<i>From empirical data to scientific conclusions - experimental design and validity</i>	7
4	<i>Communication principles in science</i>	8
5	<i>Scientific writing</i>	9
6	<i>A working scientists</i>	18

Preface

Lecture notes for the MSc course "Research Skills". Work in progress.

Please send comments and suggestions to <http://florianhartig.wordpress.com/>

Acknowledgments: Severin Hauenstein for contributions to the chapter on scientific writing

1

Logic and philosophy of science

1.1 *What is science?*

The Oxford dictionary defines science as

Science /ˈsɪəns/ [mass noun] - the intellectual and practical activity encompassing the systematic study of the structure and behaviour of the physical and natural world through observation and experiment

Hence, science is a method to find out how the world works.¹ In fact, "a method" is an understatement. In the last centuries, the scientific method has been the reason for an unprecedented increase in knowledge, opportunity and well-being. In the western world at least, science is the current dominant answer to what one can know and how to find it out (epistemology).

But science is also a human endeavor.² If one says "science", it may be that the reference relates to the history of science, the state of scientific knowledge, or the science as a social institution. Regarding those latter aspects of science, things become less formally defined - if we look at the history of science, we see a complex of individuals that are somewhat unified by a common cause, but at the same time subject to their national and cultural influences as well as by personal vanity, pride or madness. As a result, the history of science, as well as its present state, is complicated by personal preferences, power games and random events, in the same way as political and economic history is.

In the center of science, however, is a set of simple assumptions and ideals, often referred to as the "scientific method". Understanding how this method works will be the focus of the remaining part of this chapter.

¹ Epistemology, from Greek *epistēmē* = knowledge and *logos* = word is the branch of philosophy concerned with what knowledge is, how it can be acquired, and what can be known

² Many famous scientists engaged in personal feuds that, looking back, seem to be fueled by ego more than by scientific evidence. A famous example is the doubtlessly ingenious Isaac Newton. Brian Gottesman notes [here](#): "Newton was a man of great ego and great temper, and had few close friends. His dispute with German mathematician Gottfried Leibniz over the invention of infinitesimal calculus is the stuff of legend, but Newton's less famous academic feuds were both bitter and many. His fellow scientists John Flamsteed, Robert Hooke, and Henry Oldenberg were just a few of those who at times felt the sting of Newton's viciousness. Sir Isaac's most famous quotation may well have been an exercise in sarcastic, spiteful anger. In February 1676 Newton wrote to Hooke 'if I have seen further it is by standing on the shoulders of Giants.' Often taken as a sign of Newton's great humility, this famed quote was almost certainly intended as an insult to Hooke, who was hunch-backed and may have suffered from a form of dwarfism."

1.2 *Philosophical foundations of scientific thinking*

Induction and deduction

Empiricism

Rationalism

In one of the foundational works of modern science, "Le Discours de la Méthode (1637)"³, René Descartes famously observed:

Good sense is, of all things among men, the most equally distributed; for every one thinks himself so abundantly provided with it, that those even who are the most difficult to satisfy in everything else, do not usually desire a larger measure of this quality than they already possess.

³ Full title "Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth" (?)

These words of one of the founding fathers of what we today consider Science are often read and cited as a sarcastic observation of human vanity. However, if one continues reading, it becomes clear that Descartes may, after all, have been

And in this it is not likely that all are mistaken the conviction is rather to be held as testifying that the power of judging aright and of distinguishing truth from error, which is properly what is called good sense or reason, is by nature equal in all men; and that the diversity of our opinions, consequently, does not arise from some being endowed with a larger share of reason than others, but solely from this, that we conduct our thoughts along different ways, and do not fix our attention on the same objects.

For to be possessed of a vigorous mind is not enough; the prime requisite is rightly to apply it. The greatest minds, as they are capable of the highest excellences, are open likewise to the greatest aberrations; and those who travel very slowly may yet make far greater progress, provided they keep always to the straight road, than those who, while they run, forsake it.

1.3 *Popper and the scientific method*

1.4 *Kuhn and variations of the scientific method*

1.5 *Summary: the principles of science*

2

On top of the published literature, and finding a good scientific question

2.1 Literature - standing on the shoulders of giants

Citation rules

<http://www.nature.com/ng/journal/v47/n4/full/ng.3271.html>
Greenberg-Howcitationdistortions-2009

2.2 What makes a good scientific questions?

3

From empirical data to scientific conclusions - experimental design and validity

Assuming you have your question, how can we find an answer to this.

3.1 Validity

Construct validity

Are the things

Bias

(?)

Internal validity

Statistical validity

External validity

3.2 Experimental design

4

Communication principles in science

4.1 *Principles of communication*

4.2 *Oral presentation*

4.3 *Scientific visualisation*

Kelleher-Tenguidelineseffective-2011

4.4 *Logical fallacies*

5

Scientific writing

5.1 The purpose and the art of writing

Writing is an essential skill for a professional academic. Scientific progress relies on the communication of (original) scientific results(?). If you are not swayed by the greater good - also your personal recognition, your career, or simply the mark for a report or an essay relies on being able to communicate your work in writing. These things are unlikely to change soon - even in the world of technology, writing will likely remain the most important communication channel in science and many professions beyond.

Now that we established why scientific writing is important to us, we may be interested to know: What is good science writing and how can we improve our skills?

Writing is a subtle art. As for any art or craft, you will not acquire it by learning the "laws of writing" by heart and following them to the word. To become a skilled writer, you need to understand the art of writing by observation, practice and reflection. Read widely, and think about why some writers appeal to you more. Look critically over your writing, and think about the things that you can improve. This is a process that takes time and requires experience.

Nevertheless, there are some hints that can be given to be more effective in writing, in particular in academic writing. Here are the seven most important ones:

Start writing early Writing structures your thinking, and you will find flaws in your logic that you didn't see before. "The time to begin writing an article is when you have finished it to your satisfaction. By that time you begin to clearly and logically perceive what it is that you really want to say." (Mark Twain)

Know your audience Who is going to read what you write? This mostly depends on the journal you choose to publish your work. You

need to adjust your text to the knowledge and background of your audience. As ? state "it matters only whether a large majority of the reading audience accurately perceives what the author had in mind".

Know what you want to say . What is the main message you want to tell people with your paper? Sounds trivial, but in my experience it isn't. Remember: Writing a good paper is not about telling everything you know or have done about a topic – it is about telling a story and selecting the relevant details. Remember, "if you don't know where you are going, you'll end up someplace else" (Yogi Berra). Before spending hours on writing the wrong text, try to explain a friend in 5-10 min the essential points of piece, and note the structure. If you fail with this, think again about your story.

Mindset Make your goal, above all, clarity of thought and expression, and impervious logic of your argument. Read ?.

Respect conventions Most readers are familiar with the conventions of science writing and expect to find arguments, sections and statements in relatively fix positions. That is to say, if the writer respects these structural standards the readability is greatly improved. You certainly know the general layout of a scientific article (at least in the (environmental) natural sciences). It is divided in a title, abstract, introduction, materials and methods, results and discussion. Within these structural elements there are further recognizable patterns, which lead to a better understanding of the article's messages (??).

Modesty Think of your text or thesis as a piece of craftsmanship. You are an apprentice and you are asked to make a chair. What is expected of you is a simple piece of work that respects the rules you have been taught, and that is, above all, solid. You are not expected to add fancy ornamentation that might be used by an ingenious master craftsmen, and you are not asked to produce a entire set of furniture that would suffice to fit out a house. You would gain admiration and praise if you could manage to achieve one of the latter, but likely you will fail to even produce the very thing you were asked for while trying ? a simple, solid chair. Bottomline ? stay modest in what you want to achieve, but stay ambitions in how you achieve it!

Style Writing is difficult. Good writing is incredibly hard. If you think writing is no problem for you, you probably haven't even realized the extent of your writing problems. Read

In the rest of the chapter, we will look at two aspects of this hints in more detail: writing style (which covers style as well as clarity),

and the structure of a scientific paper.

5.2 *Style*

Sentence Structure

The battle for good writing is won sentence by sentence! A good sentence is: short, has the subject and verb together, has an active verb, has the points of emphasis at the beginning and end, and moves the reader along from a familiar launch point at the start to the new information at the end. (Brian McGill, on the blog Dynamic Ecology)

Paragraph structure

A paragraph is a logical unit of text.

The first sentence provides the topic of the paragraph. The last sentence should, ideally, provide a closing

<http://www.economist.com/styleguide/introduction>

Logic

Clarity

Keep it simple. ¹

¹ Mark Twain is famous for his effective writing. A tip from him: "As to the Adjective: when in doubt, strike it out."

5.3 *The structure of a scientific article*

On the largest unit of discourse the research paper comprises a title, abstract, keywords, and an introduction, methods, results and discussion section (and where required appendices). This structure somewhat forces the writer to repeat the story of the paper over and over again. In the end it should be told roughly $3\frac{3}{4}$ times – once with the title, once with the abstract, a quarter in the introduction, half in the discussion/ conclusion, and again once with the whole article.

Another structural element concerning the general layout is the "hourglass structure". This means the writer should start off with general information, becoming then more and more specific and completing with a link to the general again. Yet, it only applies to the abstract and (more flexibly) to the whole article.

Title Zooming in to the next smaller unit of discourse we start off with the title. Its main purpose is to interest potential readers since 97 % of them only read the title. To keep them reading the title should signal the topic. A good tip is to make it precise and informative with a hint at the findings – it should reflect the main message of the article's story.

Keywords The author is asked to list some keywords, which are usually located right below the abstract. These terms allow people to find the article when searching with ISI web of knowledge. Yet, since Google scans whole documents the keywords are probably not as important as they used to be. Nevertheless, a good strategy is to choose keywords complementary to title and abstract, e.g. for important technical terms that you did not want to include in the abstract nor title.

Abstract The abstract is the second time you will tell your story; in fact it can be considered as a mini paper. It starts off by setting the scene with a very short and general introduction. This usually links to the second part: Raise the problem. You can emphasize the research question by indicator words like "however". Here you tell why your work is necessary. Third, you briefly delineate your approach. The fourth subsection serves to state the results. Finally, you present your conclusions and discuss the broader significance.

To make sure you follow this structure it may be helpful to copy a template into your manuscript.

Introduction The first main section of the article is the introduction. It comprises four parts: Start off with one or two paragraphs outlining the broad topic and what work has been done before. Here you usually cite a lot of literature showing that you know the current state of knowledge concerning the topic of your study. Use another one or two paragraphs to explain the problem, i.e. the gap of knowledge. Here, you may sell your idea and why the knowledge gap needs to be filled. Then, delineate your approach in one paragraph: What are your methods; what are the data you used? Finally, you may list your specific research questions. This subsection does not need to be clearly separated from the approach subsection. It should however, provide a link to the result sections by outlining the work flow of the analysis.

Go from the general
...to the specific
...but not back to the general.

Make use of indicator words to clarify the role of the individual paragraph:
however; problem/ challenge; here we used; we applied; we asked; we tested.

Methods In the methods section the reader learns precisely what has been done. Even though details may be important this section should not take up too many words, since the total amount for the paper is usually fairly limited by the publishing journal. You will have to make a selection here; technical information, which is necessary to replicate the study (Software version, etc.) can be moved to an (online) appendix. This may even increase the readability. If helpful the methods section may also be divided in subsections such as study area, data, statistical analysis/ model, analysis, etc. Another helpful tip is to add figures. They can greatly enhance the ability of the reader to understand your approach. Furthermore, if there were any problems with your methods, do not discuss them here – use the

Keyword is reproducibility.

discussion for that.

Results Now that you have explained your methods and the course of analysis, present your results in the next section. Here you will have to make a selection once again. The results you show need to be relevant to your story, and need to match the specific research questions you lay out in the abstract and more importantly at the end of the introduction. To clarify your findings it is helpful to use figures. Remember the most important aspect about this section: Your results need to be presented as neutral facts. Do only state what is obvious, but leave speculation for the discussion. You may however give an interpretation and cross links if they are obvious.

Facts are better conveyed short and simple!

Discussion/ Conclusion Finally, we need to wrap up the article. The discussion/ conclusion section comes commonly at last but with extensive functionality. It contains a summary of the main findings. Remember here, many of your readers only read this section (and the title). It is common to summarize by restating the problem, significance and the methods in the beginning. Secondly, this section is the platform to discuss the results, speculate about reasons and theories and for the connection to other literature. As was mentioned before, this section is the place where you may tell the reader about the limitations of your approach and possibly the reasons why the main findings are still credible. However, rather establish the domain under which your results are valid; do not question your approach in general. Finally, the reader expects you to provide conclusions, recommendations, applications and questions for further research.

Move from the specific
...to the general.

Only discuss the limitations that question/limit the credibility of your results.

5.4 *Showcase: A commented paper*

Tropical forests: Finding the real reason for species richness

Florian Hartig and Severin Hauenstein

*Department of Biometry and Environmental System Analysis, University of
Freiburg, 79106 Freiburg, Germany*

October 2, 2015

Abstract

Tropical forests are some of the most species-rich ecosystems of the

Set the scene

world. The reason for this, **however**, is still widely debated. Hypotheses range from processes related to productivity over environmental stability to the historical changes in geography. **Here**, we contrasted these different hypotheses by using data from ... together with ... (fancy new method). **We find that** hypothesis X seems to be significantly better supported by our data than all alternatives we tested. Specifically ... **In conclusion**, our study supports the hypothesis that species diversity in the tropics is mainly driven by higher productivity. These results challenge some long held ideas about geographical stability being the main reason for global diversity patterns. They also have important practical applications for mitigation of climate change, as ...

Keywords: Biodiversity – Primary production – Floristic structure ...

Introduction

Tropical forests provide a suitable habitat to a vast variety of species. According to XY (19XX) tropical forests make up $x\%$ of the planet's biodiversity.

...

Many ecologists have aimed at explaining this phenomena and a large number of hypotheses has been published in recent years (e.g. ABC, 19XX; D&E, 20XX; FG, 20XX). However, not all of whom have made convincing arguments. To order the state of knowledge validation and comparison of the studies is desperately needed.

...

In this short note, we compare a set of pre-selected hypotheses published by ABC (19XX), D&E (20XX), FG (20XX) and ... and present a validation of the used approaches. For the latter we used X data provided by XYZ (2014)

...

We first test X in order to find the difference in Y. As an alternative approach we choose to look at Z for the purpose of XY. Specifically we are interested in species Y. Furthermore, the validation of the hypotheses will provide information about ...

...

In the following we briefly outline the concepts of ... and present our validation approach.

...

Methods

Results

Discussion

In this study, we used method X to test whether prevalence of A is associated to factor B. Our main result is that there is a significant correlation of B with A. This correlation, however, was only found when factor C was also present.

Raise the problem

Introduce your approach

State your results

Give your conclusions and discuss wider significance

Our findings support earlier findings of (X, 2005) and (Y, 2002) who also found a positive correlation between A and B. The fact that this correlation was only present in samples under conditions C may also explain previous opposite results such as (ZZ, 2006,2008), as neither of these previous studies controlled for C. The reason for this positive association and the fact that it is affected by C is still unknown. We speculate that mechanism R could be a reason for this. This idea is given further support by the observation that the correlation between A and B is affected by C. Such an effect would be expected if R is really the cause of this Correlation, as R is dependent on C (XAY, 1968).

5.5 *A list of common writing problems*

(A) Strategic writing problems

- (a) Do not make statements that are unnecessary for your argument, but could be questioned. Example: "The two most classic games to simulate the behavior of animals are the Hawk-Dove game and the Prisoner's Dilemma." → Others might disagree. Is it really important to be so definite here? Why not say: "Two classic games..."?
- (b) Passive/ Active: See the comments in the lecture notes on scientific writing. In general you should write in an active voice, unless 1) Sentence structure C₃ takes precedence, or 2) the thing you talk about is impersonal, was done by another person. More details in the lecture notes. In particular, opinions of your own SHOULD be active.
- (c) Be definite - remember Yoda: There is no try, do it. Phrases like: "we attempted to find whether there is" or "this study aims at" sound weak and defensive. Write "we studied whether ...". Also: "to the best of our knowledge ..." This is a borderline case. You will see this expression in scientific articles, and it's OK to use it. However, it does seem insecure. Better to indicate that you made some effort to find out together with such an expression.

(B) Logical problems

- (a) The stated implication does not follow the argument logically. Example: "Because both models follow different pay-off schemes, it is important to know how a simulation model reacts to different settings." → Unclear why it is important to know how simulation reacts to different settings if two models are different.

- (b) "Micrologic": Logical indicators such as "thus", "hence", "however", "nevertheless" have a clear and distinct meaning. "thus", "hence" indicate an implication. "however", "nevertheless" indicate contradicting statements or evidence to the previous proposition.

(C) Style

- (a) Drop redundant words: Any of the following **words** can be erased: "a **total of** n subjects", "the simulations were run with the **exact** same parameter set as before"
- (b) Pompous/ pontificating style: If something can be said in simple words, it should be said in simple words, i.e. if the only purpose of a sentence is to "sound scientific", it should not be said at all. Examples: The data **featured/comprised** a pixel size of 10 (had). We **utilized** (use).² We **used** the model in an integrated framework to **facilitate** a holistic forecast" (predicting from the model). Tip: Read ?.
- (c) Topic position/ Stress position not properly occupied: The first part of a sentence should give the context. The last part of the sentence should provide the crucial new information, and should relate to what is further discussed. Tip: Read ?.
- (d) A paragraph structures the text into topical units. Use them! Each paragraph should introduce its topic in the beginning, and ideally lead to some conclusion or summary at the end.
- (e) Unnecessary adjectives. Mark Twain: "As to the adjective: when in doubt, strike it out." Especially adjectives such as "very", "particularly" are often not necessary. Applies also to larger constructs such as "Evolutionary Game Theory is an important tool to investigate(s) ..." – if you erase the underlined text, the sentence remains fine.
- (f) Unclear/ ambiguous expressions: Write as clear and as informative as possible. Avoid ambiguous words like "tends to", "mostly" etc. if you can make a definite statement. Example: "the frequency of cooperators always tends to be zero" → What does that mean? 1) It was always zero or 2) It was zero in 95 % of the cases or 3) something else.

² Styleguides advice against complicated words such as utilize if there are simpler equivalent alternatives, although others have argued that there is a different between use and utilize - but could you name it? See discussion here [here](#).

(D) Grammar and spelling

- (a) English commas have different rules than in German and other languages <http://www.grammarbook.com/punctuation/commas.asp>

(b) Wrong use of **which and that** Let's say that we have 10 samples. Compare the two sentences: 1) the samples, which were expose to radiation the day before, were analysed. 2) The samples that were exposed to radiation the day before were analyzed. The important difference is that 1) says all samples were exposed to radiation and then analysed, while 2) implies that not all samples were exposed to radiation, and only those that were were analysed. 1) is called a non-restrictive clause, and 2) is called a restrictive clause. "that" always implies a restrictive meaning. Additional confusion, however, is created by the fact that "which" can in fact be used for both, with a minor modification: if you want to use which restrictive, you have to write 3) The samples which were exposed to radiation the day before were analyzed. Notice that there is a tiny difference to 1) - there is no comma before which. This is the only indicator that allows us to know whether the restrictive or the non-restrictive meaning is implied. Many people are not aware of this difference though. In prose, the distinction may matter little, but for science, my opinion is that clarity goes before style, and I therefore recommend to strictly stay with 1) which for non-restrictive and 2) that for the restrictive meaning.

5.6 Examples

In the following, I highlighted redundant words and other problems

- A subsequent ANOVA analysis **enabled a quantification of** the impacts of the varied factors (quantified)
- There are a number of records in the literature focusing on comparisons between sets of modeling approaches while predicting biomass at plot scale (A number of previous studies has compared modeling approaches to predict biomass at the plot scale)
- As such, the explicit findings of our two experiments showed that a total of 9 samples ...

6

A working scientists

6.1 Good scientific practice

The term "good scientific practice"

6.2 Social aspects

Collaborations

Conflicts

Teams

6.3 Science as a career