

Things you need to know if you are writing a scientific paper, thesis or report in Engineering or Science

Diogo Freitas ¹, Fernando Morgado-Dias ², and Adriana Ribeiro ²

¹Madeira Interactive Technologies Institute (ITI/LARSyS/M-ITI)

²Affiliation not available

October 30, 2023

Abstract

This paper focuses on the common mistakes made by students at the early stages of research when writing scientific papers, thesis and reports. It points out mistakes to be avoided and the basic rules that can be inferred by reading a few scientific documents but are not usually clearly written and that we, as professors, end up teaching over and over again. The outcome of this paper will be, hopefully, that we will not need to correct the same common mistakes again and you, as a student, will have a faster lane to publishing. Starting with the abstract: it is composed of a single paragraph, does not contain acronyms or references and describes in short the work, main highlights and points out the results or main conclusions obtained from the work being presented. The abstract is an independent part of the paper and commonly has a character or word limit that you need to respect. It can be read as a “stand-alone” and the paper starts in the introduction, meaning that the introduction is not the sequence of the abstract and it can have some text in common if needed.

THINGS YOU NEED TO KNOW IF YOU ARE WRITING A SCIENTIFIC PAPER, THESIS OR REPORT IN ENGINEERING OR SCIENCE

A PREPRINT

F. Morgado-Dias

Madeira Interactive Technologies Institute (ITI/LARSyS/M-ITI)
University of Madeira
Funchal, Portugal
morgado@uma.pt

Diogo Freitas

Madeira Interactive Technologies Institute (ITI/LARSyS/M-ITI)
University of Madeira
Funchal, Portugal
diogo.freitas@iti.larsys.pt

Adriana Ribeiro

Secretaria Regional da Educação
Madeira, Portugal
adrianarib.prof@gmail.com

August 8, 2020

ABSTRACT

This paper focuses on the common mistakes made by students at the early stages of research when writing scientific papers, thesis and reports. It points out mistakes to be avoided and the basic rules that can be inferred by reading a few scientific documents but are not usually clearly written and that we, as professors, end up teaching over and over again. The outcome of this paper will be, hopefully, that we will not need to correct the same common mistakes again and you, as a student, will have a faster lane to publishing. Starting with the abstract: it is composed of a single paragraph, does not contain acronyms or references and describes in short the work, main highlights and points out the results or main conclusions obtained from the work being presented. The abstract is an independent part of the paper and commonly has a character or word limit that you need to respect. It can be read as a “stand-alone” and the paper starts in the introduction, meaning that the introduction is not the sequence of the abstract and it can have some text in common if needed.

1 Authors

Author of the paper is someone that contributed to the final result reported in the paper. The person who did the experiment, wrote the code, did the state of the art or had the original idea is (most commonly) in the first place. The order in which the authors appear in the paper is an issue that can cause controversy. Common approaches are:

- First author is the one that conducted the experiment;
- Last author is the supervisor or the one who had the original idea.

Some journals request a detailed description of the work developed by each of the authors, the author contribution. Some journals allow writing that some of the authors had an equivalent contribution even though only one can be the first author.

2 Introduction

This is where the paper starts (it is not the sequence of the abstract). Expose what the problem is, what was done previously (this will be the state of the art, although it might be placed in a separate section or subsection) to solve it and why there is still need for something different (the contribution). The authors might also unveil a little about their solution.

The introduction can finish with a resume of what will appear in the next sections, describing very shortly the content of each of the sections that will be presented next.

3 Structure

There are different structures that can be followed and there is some freedom to choose what fits best the work and document. Do note that some journals and conferences have their own predefined structure that has to be followed.

Usually the structure has an introduction, state of art, problem description, results, discussion, conclusion and references. Introduction, results, conclusion and references are almost always mandatory. The others might go under a different name or not exist at all. In most formats, a section and subsection should be preceded by a number. They can follow regular or roman numbering.

In some fields, some structures are more common or almost fixed. Take the example of the IMRAD structure [2]. The IMRAD acronym stands for Introduction, Methods, Results, and Discussion, and it has started in the health field but is becoming more and more used.

4 State of Art

In the state of art section it is necessary to review what has been done before this work in the area of the new contribution. Note that:

- If the area already has many contributions, it will be difficult to make a detailed state of art, so the author might want to make an incremental one, presenting only recent work on the field;
- If the area is very new, the state of art might open up to a broader view so that the author points out that there are (almost) no relevant contributions in the area;
- The objective of the state of the art is to identify the knowledge gap that the work is trying to cover with this contribution. Imagine that sleep has been monitored in several mammals and stages have been successfully identified and linked to the length of their life span; sleep has also been monitored in humans and stages have been identified. The proposed work identifies the knowledge gap for making the connection of length of lifespan to sleep stages in humans, and this is where a contribution is presented.

Without a proper state of the art, the reviewer will have a hard time to measure the contribution of the proposed research.

Structured reviews became a hot topic recently. To make the process more systematic and with the increase of means to do searches on-line, PRISMA was created. It stands for Preferred Reporting Items for Systematic Reviews and Meta-Analyses and is illustrated in figure 1.

Usually the process starts by explaining how the query was placed, for instance, using the combination of two keywords: professional AND football. It is necessary to register the number of records (documents) that are found from several sources, which then will go through the steps indicated for: removing duplicates; checking for relevance and inclusion. This process includes reporting everything from the queries to the number of records in each stage.

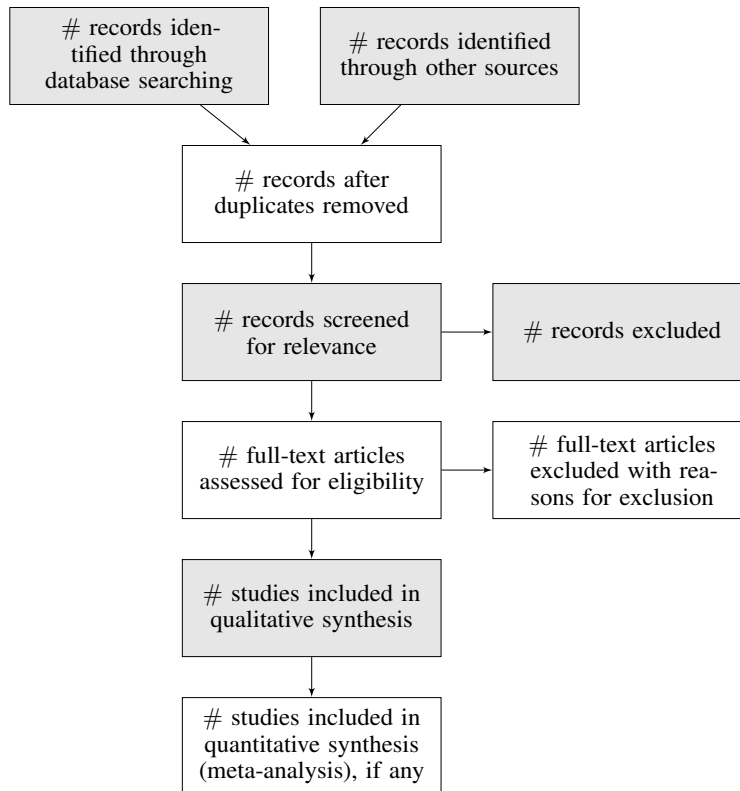


Figure 1: Steps for the PRISMA structured review technique [3].

5 General Writing

There are a few remarks about the general way of writing. They will be presented in two subsections regarding style and acronyms. Do note that, when opening a section with subsections, in general, there should be some text in between.

5.1 Style

A paper to present research is a formal document written using an impersonal style. Although the latter might not be consensual, one should not write: “I/We did the experiment in the following form...”. Consider replacing it with a more impersonal sentence like: “The experiment was conducted in the following form...”.

Note also that the written form (specially in a more formal document) is different from the oral form. For instance, do not use the abbreviated forms like “didn’t” or “it’s”. They can be used in an oral form but they should not be in a formal document.

If the paper is to reach a broad audience, then English is the reference language. Please do not translate word to word from the author’s native language. Try to think in the language being used.

Here are some bad examples: “It was used electricity to power the car (...)”. It should be: “Electricity was used to power the car (...)”. There are other common mistakes, such as: data is plural (“Data were used to calculate the results”).

A paper is composed of sections, subsections and so on. You can also refer to them as parts and in most cases when making references to other parts it might be interesting to add the number. Example: “As mentioned in subsection 3.1 Hardware Description...”.

5.2 Acronyms

Acronyms are only worth to be introduced if they are used several times throughout the text. It should be written with the full name in the first time, with a capital letter starting each word, followed by the acronym between round brackets.

Example: Field Programmable Gate Array (FPGA). Then the next time mentioning it, use FPGA, not the full name, not the full name followed by the acronym again.

6 Figures, Tables and Images

Figures and Tables follow the journal/conference style for the legend and referencing.

All graphs need to have the axis information plus the corresponding units from the international system, unless there is a very specific reason to do otherwise.

The figure 1 should be Figure 1. If referring to a specific figure, like figure 1, do not use “The”. If referring to figures in general, in an undefined way, use “the”. Example: “The figures must be numbered. Figure 1 is the first”.

Figures and tables must always be mentioned in the text.

A block diagram is different from an image.

When using a figure, table or image that is not originally from the authors, introduce a reference in the subtitle of the figure.

The image should not be too big. The text in the image should not be bigger than the text in the main body.

The image should be, preferably, in vectorized format. Many conferences and journals do a post processing editing of the submission. When they change the images, vectorized images “behave” much better to the adjustment.

7 Results and/or Discussion

This section is where the results are presented.

The results section is meant for presenting the results and the conditions that led to them, so that someone can repeat the experiment, when that can be applicable. For instance, it might be necessary to state the characteristics of the computer if presenting the processing time used to obtain the solution.

The discussion might be part of the results section, at the end, or be a separate section. In either case, this part should be used to discuss the advantages or disadvantages of the proposed solution, mostly compared to other work found in the literature, the state of art. The authors want to convince the reviewers that they have made a progress, that they have taken the research in the area one step further so that the reviewers can give a positive review and recommend the publication of the paper.

8 Conclusions

Do not make it just a summary of the paper and do not state the obvious. The reader needs to know why things happen and not just a confirmation that the authors also have seen what is reported in the data and tables.

Make it clear what the contribution is and why (at least in some aspect) it is better than what was done before by other researchers.

9 Citing

When citing relevant work, let us say a theorem, one must cite the original research, that is, the one work that presented proof of the theorem, not the latest user (or paper) mentioning the said theorem.

Citations are also a form of recognition of the authors’ work. With some restrictions and subject to discussion, the citations the work gets is a form of measuring the impact it had on the readers.

If this was helpful, please cite this paper as: F. Morgado-Dias, D. Freitas and A. Ribeiro, “Things you need to know if you are writing a scientific paper, thesis or report in Engineering or Science”, TechRxiv, 2020.

When inserting the citation in the text, one should do it in such a way that if the reference is removed, the text is still readable. Example: “According to [1], the maintenance cost is estimated between 1–3% of the initial investment per year”. This should be: “According to Morgado-Dias and Freitas [1], the maintenance cost is estimated between 1–3% of the initial investment per year”.

Each journal or conference has its own style for references that need to be followed strictly. This involves the references inside the text, for instance just a number between brackets [1] or using the names: (Morgado-Dias and Ravelo, 2020) or (Morgado-Dias et al., 2020) when there are more than 2 authors, but also the way the references are organized in the reference section at the end of the paper. Please check what is the style the paper should have.

The authors can cite their own work if it is relevant to the current publication and they need to inform the reviewers about the previous published work if it is connected to the current work. Let us say that the current paper (intended to submit to a journal) is an extended version of the paper that was published in a conference, then they should explain this in the paper, probably at the end of the Introduction.

For managing citations, there are a few tools that can be used and that will help save some time. Most authors will end up using a few citations several times across different documents, so it is better to keep them handy. Examples of tools are: Citavi, Mendeley, Zotero, BibTeX (if using a \LaTeX -based tool). These tools can also help switching from one citation style to another without breaking a sweat. They can be very helpful when it is necessary to resubmit a paper that was not accepted at first attempt.

Another key principle is that whenever reading a new paper, thesis or report, notes should be collected (preferable in a digital document that can be searched). Mark the highlights, how it is related to the work and what can be cited (maybe relevant sentences). Doing so will save time when it is necessary to go back and will make it easy for a quick search.

10 Answer to Reviewers

When answering the reviewers, it should be done in a separate file. Copy their comments and answer below. In general, try to avoid disagreeing and answer in a polite way. If it is necessary to refute an argument, try to do it in such a way that does not cause resentment. Remember, if a certain point was explained but the reader did not understand it, most likely the explanation can be improved.

Also note that some answers can be only to the reviewer, while others need to be in the paper but also adding a comment to the reviewer so that he knows what was done.

11 Letter to the Editor

In most journals, it is possible to include an optional letter to the editor. This letter can be used to explain some particular detail of the paper in case it exists or simply to make it clear that the paper is original and is only being submitted to that particular journal or why it is a good fit for that journal.

12 Unwritten Rules

Some journals have rules that are not clearly written. Some journals only accept papers under certain conditions. The authors need to collect that information before trying to submit. This can sometimes be inferred by reading most recent published papers in that journal.

13 Acknowledgements

Acknowledgements are not meant to say that the authors are thankful to their girlfriend/boyfriend or parents for their support. They are meant to acknowledge funding entities (see example below) and collaborators with a smaller role in the work that meant they are not co-authors.

Examples are:

- Acknowledgements to the Portuguese Foundation for Science and Technology for their support through Projeto Estratégico LA 9 – UID/EEA/50009/2013;
- Acknowledgement to ARDITI – Agência Regional para o Desenvolvimento e Tecnologia under the scope of the Project M1420-09-5369-000001 – PhD Studentship;
- Also acknowledged the Funding Program + Conhecimento II: Incentive System to Research and Technological Development and Innovation of Madeira Region II, through the project “Smart Solar” – MADFDR-01-0190 – FEDER-000015.

14 Choosing Where to Publish

When choosing where to publish one should consider that a conference is usually the choice for a smaller contribution or to get feedback for a more meaningful contribution that will be followed by a journal publication. There are, nevertheless, conferences where it is very difficult to publish but in most areas the journal publications will be considered more relevant or in the upper part of the pyramid.

14.1 Indexing and Impact Factor

The Impact Factor (IF) is a ratio of citations per paper that is assigned to each journal. The highest the impact factor, the more prestigious the journal should be.

Currently, there are mainly two indexes in use to measure this:

- Web of Knowledge (or ISI or Clarivate), published in the Journal of Citation Reports (JCR);
- Scimago.

The first is considered as the main reference and considers only citations from other journals indexed in their own index, but has a paid access; the second because it is open and covers more journals and conferences is having an increasing use.

Google Scholar, that is probably the best source to search for scientific documents, also has indexes but they are less specific in terms of areas and mix together conferences and journals so they are less used.

Because there are so many journals nowadays, they are frequently divided in Quartiles and the Quartiles are used as a simple reference to the journal’s importance.

To understand quartiles and to make it very simple, let us say there are 100 journals in a given area. The 25 ones that have the highest IF would compose the First Quartile, Q1 for short. From the 26th to the 50th (ordered based on the IF) would be in Q2, then 51st to 75th would make it to Q3 and the remaining ones would be on Q4. This is independent of the values of the IF. And these are very area dependent.

Now, publishing in a Q1 or Q2 is a lot more valuable.

14.2 Practical Examples

Now let us see an example. Consider the case of the Entropy journal, which is a journal that covers a wide set of topics. On their page we can find this information:

- **CiteScore** (2019 Scopus data): **3.7**, which equals rank 8/62 (Q1) in “Mathematical Physics”, 14/54 (Q2) in “Physics and Astronomy”, 210/670 (Q2) in “Electrical and Electronic Engineering”, and 104/300 (Q2) in “Information Systems”.

Although this information is a bit confusing since the first value belongs to Scopus (another index) but the following ones belong to JCR.

The information related to the same journal on Scimago is copied in figure 2.

Where yellow represents Q2, orange Q3 and red Q4. Green would have been used for Q1.

Now it is interesting to note that both indexes (JCR and Scimago) place this journal in different areas, which makes sense since it is a multidisciplinary journal. In case of Scimago only for 2019 it has been indexed also in the Mathematical

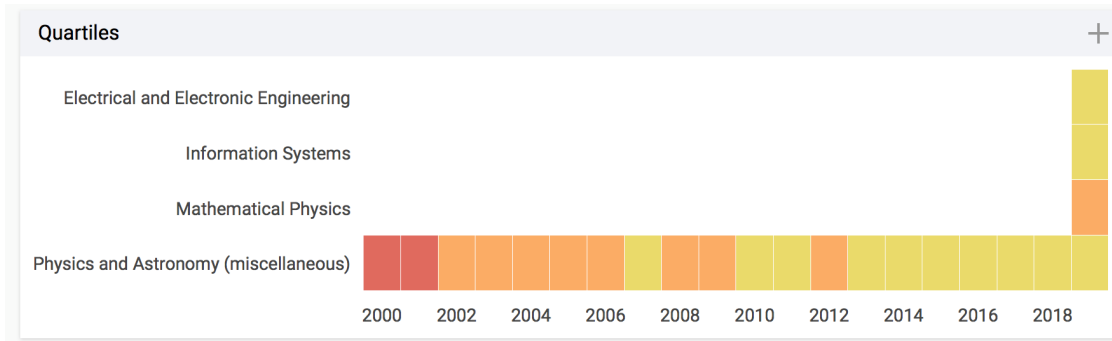


Figure 2: Data for Entropy journal in Scimago database [4]

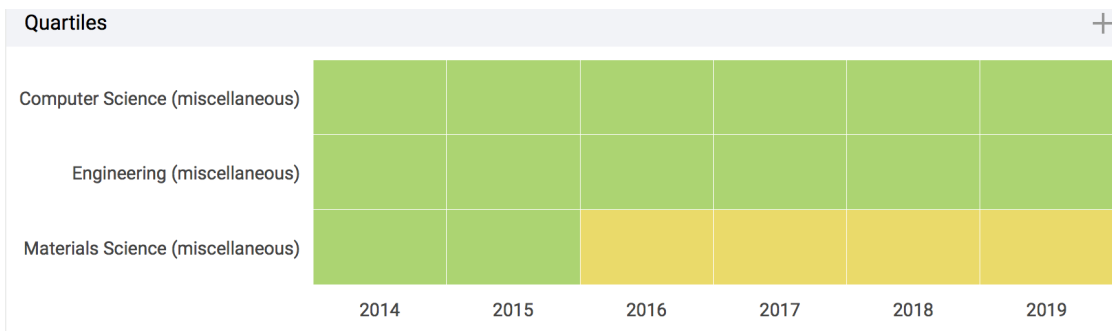


Figure 3: Data for IEEE Access journal in Scimago database [4]

Physics, Information Systems and Electrical and Electronic Engineering. Also interesting is the fact that the classification areas do not have the same names.

The same happens with another multidisciplinary journal, IEEE Access whose data can be seen in figure 3. This journal is classified in 3 different areas.

Looking back to the Entropy example, one should note that it is Q1 in Mathematics (according to JCR) and Q3 in Mathematical Physics in Scimago. They are not exactly the same area but they could be and this could still happen.

The explanation lies in the fact that the citations are counted in a different group of journals since the indexes do not cover exactly the same journals.

It would be like asking two different groups of people to rank the best soccer players in the world and expect the exact same order.

In general, the citations for each journal are the same for each year and so is the IF, but when compared within each group, the same IF of 2.494 for Entropy can result in Q1 or Q3.

One should also note that there are areas that reach much higher IF than others. This is the case of Health and Energy (as an example) while others keep lower values, such as Electronics. This is related not only to the quality of the journals and papers but with the number of journals in the area, the number of researchers, the interest it has for the public, among other factors.

15 Open Access

Most of the scientific journals are connected to publishing houses that work to obtain profit. Being so, although the large majority of the work is done freely by researchers (as authors, reviewers and editors), in the end, the authors either pay to access the paper, or pay to have it open for the public, the open access concept. The two journals mentioned in the previous examples are open access journals and that means any reader can download and read the paper for free. This is very interesting for the paper since it can reach more people and have more recognition (citations). The drawback is that the authors will have to pay from a few hundreds to a few thousands dollars or euros.

There is always an unsupervised version like the place where this paper was downloaded: the `techrxiv.org` but in some areas, it will not get the same kind of recognition as with a published paper in a journal with IF.

16 Disclaimer

This document is meant for students at their early research stages.

We are aware that different papers can be presented in different style and the authors have done so in many occasions. Nevertheless, an attempt was made to present here the most common case for papers. Naturally, it will not fit every paper's situation but can help to reduce a number of common mistakes that are usually done when one starts to write scientific papers.

References

References is a section which is usually not numbered, but its content is in itemized style and fully numbered. Usually, to save space it also uses a lower case size. Any reference needs to have, at least, the following: authors, title, type of document, pages (except for books and thesis), publisher, date and eventually location.

Citing web based documents, while not desirable, is sometimes unavoidable. It should not be the main part of the references. When using it, one must place the address and date of access in the reference. Example: Wikipedia, "Tesla Powerwall," Wikipedia, the free encyclopedia, 2015. [Online]. Available: https://en.wikipedia.org/wiki/Tesla_Powerwall. [Accessed: 15-Feb-2020].

The most common reference style (do not forget to check the place where the work will be submitted) will look like this:

[1] First name initial, Family_name (First author), First name initial, Family_name (Second author),... (following authors), "Title of the paper in full", name of the journal or conference, volume, page interval, year, publisher and eventually location (usually for a conference).

[2] Luciana B Sollaci and Mauricio G Pereira. The introduction, methods, results, and discussion (IMRAD) structure: A fifty-year survey. *Journal of the Medical Library Association*, 92(3):364, 2004.

[3] Wikipedia contributors. Preferred reporting items for systematic reviews and meta-analyses — Wikipedia, the free encyclopedia, 2020. [Online; accessed 18-July-2020].

[4] SCImago, (n.d.). SJR — SCImago Journal & Country Rank, 2020. [Online; accessed 18-July-2020].

Please note that the order used for the references also changes with the style adopted by the conference/journal selected. Some organize it in an alphabetical order, some according to what appears first in the text and this needs to be respected for the submission.

Biographies

Several journals ask to submit short biographies of all the authors. These are usually very fact based and not meant for dispersion from the facts. Usually, it reports the degrees, present and past affiliations and research interests and it is enough. Sometimes a few more facts can be added, like previous functions. Some ask for a picture that should be close to a passport type picture.



Fernando Morgado-Dias was born in Coimbra, Portugal, in 1971. He received the B.S. degree from the University of Aveiro, Portugal, in 1994, the M.S. degree from University Joseph Fourier, Grenoble, France, in 1995, and the Ph.D. degree from the University of Aveiro in 2005.

He was a Lecturer with the Technical University of Setúbal, Portugal. He is currently an Assistant Professor with the University of Madeira and Researcher with the ITI/Larsys/Madeira Interactive Technologies Institute. His research interests include sleep monitoring, renewable energy, artificial neural networks and FPGA implementations.

Dr. Morgado-Dias is currently member of the Fiscal Council of the Portuguese Association of Automatic Control (APCA) and Director of PhD degree in Electrical Engineering. He was Pro-Rector and President and Vice-President of APCA. He is also author of around 150 documents among conference and journal papers, technical reports and thesis.



Diogo Freitas was born in Madeira, Portugal, in 1996. He received the B.S. degree in Informatics Engineering from the Faculty of Exact Sciences and Engineering, University of Madeira, Portugal, and is currently finishing an M.S. degree in Mathematics, Statistics and Applications at the same institution.

He is a Research Assistant with the Madeira Interactive Technologies Institute (ITI/LARSyS/M-ITI), Portugal. His research interests include the development of new nature-inspired algorithms for solving some hard numerical problems.



Adriana Ribeiro was born in Esposende, Portugal, in 1979. She received the bachelor's degree from the University of Minho, Portugal, in 2002, the postgraduate studies from Institute Jean Piaget, Porto, Portugal, in 2014.

She was an adult trainer of language and communication, writing and Portuguese for foreigners. She is currently teaching in the Portuguese public educational system. Her research interests include pedagogy, writing, literature, linguistics and grammar. Adriana Ribeiro is currently vice president of the Institute of Knowledge and Development and organized scientific events in Madeira and Azores.