Scientific Writing

Guide of the
Empirical Software Engineering Research Group of the University of Helsinki,
Department of Computer Science

Tomi Männistö, Juha Tiihonen, Fabian Fagerholm
firstname.lastname@cs.helsinki.fi
www.cs.helsinki.fi/group/ese/
Version 1.0, October 21, 2016

Introduction

These guidelines aim to assist students in writing scientific text. The guidelines are applicable for seminar reports, B.Sc. and M.Sc. theses and also when aiming to write your first scientific publication. These guidelines are intended for software engineering and related areas of research. They are not necessarily directly applicable to other fields, e.g., theoretical computer science.

Note that guidelines and instructions vary in many respects between institutes, individual supervisors and scientific forums. However, some principles are more or less generally applicable. Learn first to follow some rules before even considering bending them. With experience, you may reasonably justify not following a particular rule.

Learning scientific writing provides the ability to express one's thoughts with particular clarity and communicate them in a manner that seasoned scientists find easy to follow.

General Notes on Scientific Writing

Some aspects of scientific writing differentiate it from many other forms or styles of writing. These include:

- **Audience.** In all writing, one needs to know the audience. You can assume basic knowledge about the field or area of research. Therefore, you can be very brief with the basics – it is more important to get to the point. This also applies to thesis work: you should not start writing a textbook about your thesis topic.
- **Clear concepts.** Extra care is put to the definition and consistent usage of key concepts. Conceptualising relevant phenomena is an essential ingredient of scientific thinking. This is reflected in writing as well. You must select the concepts, define them and use them consistently in exactly the meaning you defined. Define all concepts when they are used for the first time. However, in some, relatively rare

circumstances a forward reference to a definition is acceptable. For example, if the definition is complex and requires other, yet undefined concepts, it may be better to define the concept accurately later. However, consider the possibility of restructuring the text so that there is no need for forward references and definitions can build on each other.

One rule of thumb is **not** to use synonyms! If you use two closely related concepts in the same article, the reader starts asking what the difference is. Thus, you either need to very carefully clarify the difference, and justify why the distinction is needed, or select either one of the terms.

    - For example, the terms “user” and “customer” can be defined to be clearly separate concepts or they may refer to the same, in which case using only one is less confusing.

- **Claims or arguments.** The overall tone of any claims is cautious. That is, do not make any unjustified claims. Be very careful with “generally known facts”, as it may turn out that in some reader’s context they fail to apply.

  In general, use references and other means of justification that give the reader a clear understanding what is the basis for your claims. When you refer to others’ work as a part of your argumentation, make sure that what you write is what the authors of the source actually wrote and meant – do not put words in their mouth.
- **Your work and that of others’.** It is of utmost importance to make a clear distinction between what is your original contribution and what was known before your work. Apply references to indicate previous knowledge (see below). To be able to give proper credit to the previous work, you need to know it. State clearly what is your own contribution and its novelty.
- **Clear and to the point.** Scientific writing can be considered even dull in the sense that you are not expected to perform creative writing, but to clearly communicate your point with explicit arguments and justification. This also includes the use of certain idioms, expressions and structures, which can be learnt from existing scientific work. In fact, expressing something in a novel way is not necessarily a good idea at all and it is better if you can find an expression used in the scientific literature. That is, you are better off with sticking to the established ways of scientific writing.
- **Storyline and consistency.** A scientific article needs a clear storyline and all the parts must form a consistent whole. Here, storyline does not mean a narrative in the sense of a fairy tale or novel, but a specific writing pattern that takes the reader from an understandable starting point, usually a research problem or question with a basis in previous literature, through the design and execution of the research study, to the results, and finally to the answers to the questions and implications of the answers. The reader should have no trouble following your line of thought. A clear structure helps both you and the reader; see section Structure below for more details.
- **Figures and Tables**. Scientific results are often summarised as figures and tables. Thus they have a central role in the presentation and are not just illustrations. Consider how to show your results in a compact way. Typically this calls for a table or a figure. Tables and figures should be independent in the sense that they can be understood with the help of their caption but without the main text. They should also have a clear supporting role to the text. Explain the table or the figure in text; all elements, e.g., of a figure should be explained.

Writing Process
(intertwined with the research process)

Writing is a personal process. Different people have very different ways of getting the words on paper. You need to learn to know yourself in this sense: what gets you to write even if cleaning windows seems much more interesting and urgent?

There are good guidebooks for writing a thesis; look for them. (One example for Finnish readers is Svinhufvud, K. (2015). Gradutakuu. Art House. Also see the accompanying web site at http://www.gradutakuu.fi/)

General phases in the whole process include:

- **Previous work.** Understanding the previous, published scientific work is the bedrock for and distinguishing aspect of scientific research. Simply, read, search for information, and take notes as you go. In writing, the previous work is clearly reflected in the beginning of the text. It is not seldom that too little effort is spent on previous work and something important is noticed late in the process. In the worst case, this may invalidate much of the effort spent or otherwise cause major challenges or rework.
- **Research methods.** Another bedrock is applying proper methods in your research. In order to do that, you need to know the methods and how and when to use them. Then, of course, you need to use the selected method(s) properly. When writing, explain the methods you used and how exactly you did it. In other words, describe your study design including data collection methods and sources, analysis methods, etc. Look for good examples from existing articles.
- **Research questions.** We recommend spending notable effort on thinking and (re)phrasing your research questions. Defining research questions is both an essential part of your research work and also a powerful tool for your writing process. It focuses and scopes your work. Do not underestimate the importance of your research questions or the effort to specify and justify them. There is a section below devoted to research questions.

Starting the writing process is an important activity that has to be performed multiple times, perhaps daily. We recommend writing something every day to keep the process running. It also helps to have a few structures to help you:

- **A logical paper or thesis structure** allows you to immediately write in the right location rather than struggling to find where to put your words. While writing, focus first on getting a few sentences or paragraphs down. Then, move them to the right location if they ended up in the wrong place. Regularly review the structure and think about how it supports the scientific storyline. Restructure if necessary.
- **Writing as you read** related work allows you to capture both the facts from the literature as well as record gaps in the literature, which you could address. Writing

already at this stage helps you to connect different pieces of related work into your discussion about the state of the art and build an initial version of your text.
- **Write, revise, write, revise.** Writing is an iterative process where you alternate between writing new text and revising the text you already wrote. Revising includes restructuring the text but also improving it to make it clearer, more understandable, and removing parts that are not necessary. This involves critically reading your own text, taking the role of a reviewer who checks that what you say holds up and is clearly expressed. If you have trouble getting your writing to flow, read and revise some part of your previous text, perhaps what you wrote the day before. In the best case, you will develop a routine where writing and revising keep you going for the entire writing project.

Research questions

A research question summarizes an issue that you want to investigate in a clear and focused way with exact and carefully considered wording. Good research questions are the basis of your research work. They guide your research work, help to identify data to be collected and analysed, facilitate the construction of logical argumentation, and provide extremely valuable assistance for writing.

A good research questions is:
- **Relevant**: unanswered question originating from literature or practice will be of academic and intellectual interest.
- **Interesting**: At least you should be interested in the answer. A good research question should also be interesting to the research community and potentially to practitioners as well.
- **Clear, simple and conceptually sound:** Formute the research question with exact and carefully considered wording. Include what is essential for sufficient exactness, but do not bloat your questions with unnecessary terms. You should have a good understanding of what the concepts mean in the context of your work. Avoid any ambiguity in the expressions you use.
- **Manageable**: you must have access to relevant data sources. And you must be have the skills, resources and time to provide an answer.
- **Scoped**: A research question must be substantial enough – not trivial. And it must be focused enough so that answering is possible and the question remains manageable.
- **Novel**: Scientific publications and doctoral theses require creating new scientific contributions – a research question should facilitate original research on the topic. Scientific novelty of research is not as strictly required, for example, in a Master's thesis or a seminar work.

Formulating the research questions is a major and difficult task. It tends to require multiple iterations and consideration of many alternatives. The process is thus iterative, converging towards the explication of what your research work about.

As mentioned, a good research question is well scoped. A research question is typically broader than what you can entirely cover in your research; your research is providing a reasonable, although incomplete, answer to the question. On the other hand, the question should not be too broad so that your results hardly even start to answer the question. For example, think about a research question: How to improve feature prioritisation in company XYZ? It may be that the research community is not that interested in what happens to company XYZ, but they may find the more general problem of feature prioritisation interesting. Thus, you may state your question more broadly without the company and give the answer through the lessons from one company, set in the context of existing literature, etc.

The relationship of a hypothesis and a research question: A hypothesis shares the same purpose of being a clear explication of what exactly is being investigated. A hypothesis is a proposed explanation for a phenomenon – a tentative, testable answer or "an educated guess" to a scientific question. A well-formed hypothesis consists of one or more assumptions that can be tested through study or experimentation. For example, an assumption makes it possible to create a setup for an experiment, where it is possible to measure the outcome of a factor that you change in the experiment. A research question, however, also bears the idea of being a question, that is, not primarily assuming, stating or claiming something, but asking.

Utilizing research questions as tools

First, the research questions are a means for scoping your research. On the basis of research questions, you can determine the required extent and scope of background and previous work. You can also determine whether you actually have or will have something to write (i.e., results) about the actual question(s) you are putting forward.

Second, the research questions provide a structural backbone to the text, as the background (motivation and literature on previous work) should clearly connect with them. Thereafter, your research methods and study design should be suitable for providing answers to the kind of question you have. The results, then, should give those answers. Your answers are not necessarily exhaustive, but should provide something meaningful at the very least. In Discussion, you separately address how good answers your research was able to produce to the research questions.

Examples and suggestions to improve

RQ1: Can test automation be effective?

- Consider the question word you use. This kind of question expects a Boolean answer, which may justifiable if feasibility is the interesting question. However, the interest may actually be in how to make something happen. So, you may consider:

RQ1': How to make test automation effective?

- Now, the answer could be based on literature and synthesising empirical evidence or may come from the understanding your research has gained from a case company or companies.

- With this question, you would need to have a good understanding what you mean by something being "effective" in your research. It is time to think about alternatives such as "productive", "economically feasible" or "support continuous integration".
- It is also good to think, whether the "making" is the point, or would it be more about the factors behind the effectiveness. Or maybe even focussing a bit more, such as into test cases.

RQ1'': What makes test automation effective?
RQ1''':What affects the maintenance of test cases?
RQ1'''': ...and so on.

Avoid superlatives

RQ2: What is the best way to utilise sentiment analysis in requirements elicitation?

- Extreme words, such as best, optimal, maximal, etc., are difficult. You often already have hard time defining what "best" means, not to mention measuring it.
- Finding the best, optimal or maximal something is probably even not the real question, although formulating such a question is tempting. Often, finding something "good" could be more appropriate. So, you may consider, for example:

RQ2': How to improve requirement elicitation by sentiment analysis? or just
RQ2'': How sentiment analysis can support requirements elicitation?

Be neutral

One particular and important aspect when asking a research question is trying to be neutral. That is, do not state the question so that it is leading towards particular answers. E.g.,
RQ3: What are the biggest flaws in feature models?

- Well, again "the biggest".
- When looking for flaws, you easily find flaws (even if not that remarkable). For example, if this reflects on how you interview developers or analyse the models, you may be (mis)lead to digging only for challenges. Therefore, think carefully why not to consider the both sides, e.g., benefits and challenges. This helps to gain a more balanced view and a context for the severity of the identified challenges. This is certainly not a rule, but merely a note for making you to think bit harder. The following might work for you as well:

RQ3': What are the benefits and challenges of using feature models?

- Already hinting that the interest would be in use of feature models in practice rather than, say analysis at conceptual level.
- Or, consider splitting into two subquestions and making the view of practice more clear:

RQ3'': How feature models are used in practice?
RQ3''.1: What are the benefits of using feature models?
RQ3''.2: What are the challenges in using feature models?

Avoid coupling research questions

As above, it is sometimes acceptable to have two questions in one, but as a general rule that is better avoided. Particularly, when the questions are not such a natural pair or are more clearly a distinct pair. E.g.,
RQ4: How to improve architecture descriptions and architecture assessment?

- There are clearly two questions, although closely related, but it is probably better to treat them separately, e.g.,

RQ4': How to improve architecture description for architecture assessment?
RQ5: How to make architecture assessment lighter?

Structure

These guidelines adopt the so-called IMRAD structure of scientific articles as the backbone. With some modifications, IMRAD structure is applicable to a thesis as well. IMRAD is an acronym from the letters of the main headings a scientific article should have: Introduction, Methods, Results and Discussion.

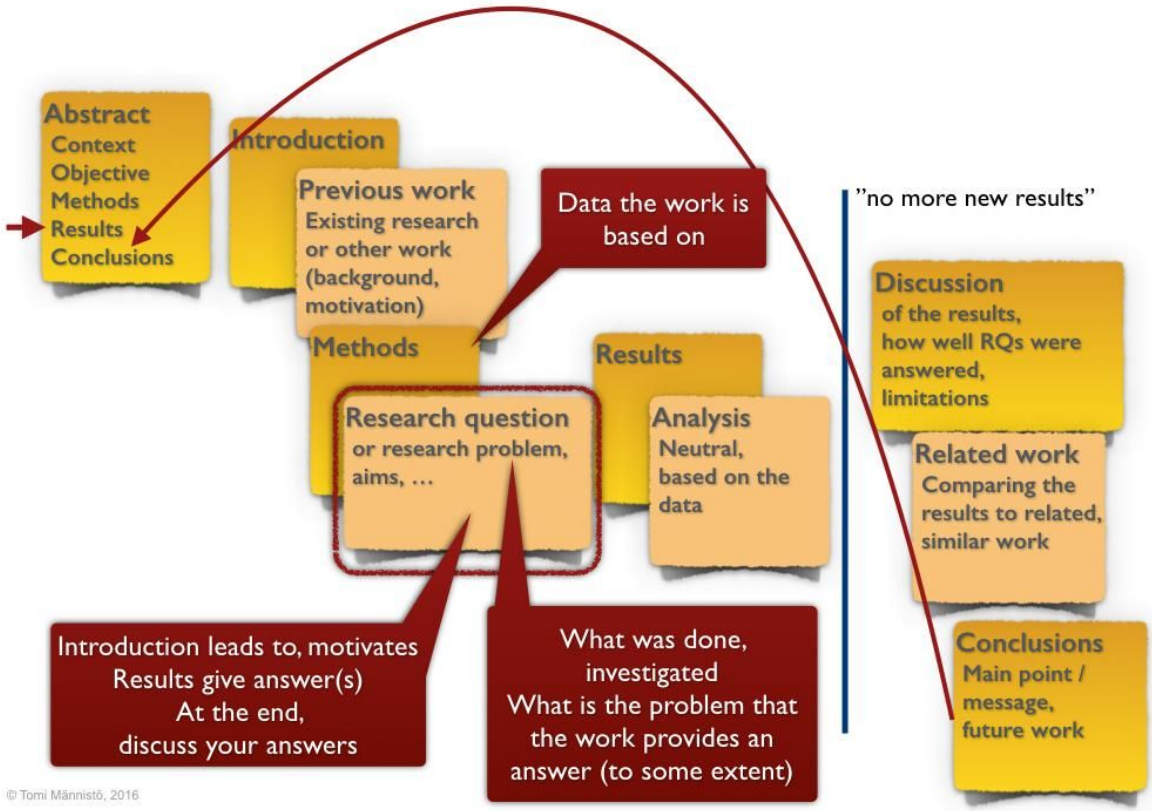

Figure 1. Overview of the structure of a scientific article.

A brief definition of IMRAD taken from Wikipedia: Original research articles are typically structured in this basic order:

- **Introduction** – Why was the study undertaken? What was the research question, the tested hypothesis or the purpose of the research?
- **Methods** – When, where, and how was the study done? What materials were used or who was included in the study groups (patients, etc.)?
- **Results** – What answer was found to the research question; what did the study find?
- **Discussion** – What might the answer imply and why does it matter? How does it fit in with what other researchers have found? What are the perspectives for future research?

In addition to the major IMRAD sections, a research article also has other elements. An article begins with:

- **Title**
- list of authors, their affiliations and contact information
- **Abstract**

and an article ends with:

- **Conclusions**
- **Acknowledgements** of help in doing the research, e.g., funding
- **References**
- **Appendices** (if any)

There may also be other sections, such as:

- **Previous work** – describes the literature that is used as a basis for the work; such background may be incorporated in Introduction, particularly in shorter articles
- **Research questions** or **Aims** – the research questions or aims of the study are sometimes presented as a top-level section of their own, particularly in a thesis
- **Related work** – which often comes after Results relating the results described to what similar, related work exists; may be a subsection of Discussion
- **Limitations** or **Threats to validity** – sometimes a specific section is devoted, or required, for discussing the validity; may also be a subsection of Discussion

Next, we detail the IMRAD and conclusions sections. Figure 1 gives further details.

Introduction

One of the main purposes for Introduction can be stated as giving the reader a good reason to read on. The introduction should set the context for the problem, maybe with the help of a few key references or by describing the problem setting and situation in reality, if your work is more empirical in nature. You should also give an overview of what is known of the problem, mainly to motivate why doing your research is meaningful. You should also give an idea what you have done and how you have approached the problem with proper justifications so that the reader gets the idea that given the above situation and problem what you have done is interesting and worth reading more about.

The last section of the introduction is typically a summarising paragraph on the structure of the article or thesis. It is not supposed to be just a list of empty sentences mentioning all the main sections, e.g., as for Methods: “In Section 2 the research methods are presented” but rather something more insightful and informative, such as “Section 2 motivates the selection of case study as the main research method and the software development environment of ACME.”

Methods

This is the place to describe your study design and details how you collected your data and how you analysed it. Give enough appropriate details, which you must have collected while doing the research, such as numbers of people interviewed, their background, context, etc., and any other sources of information, e.g., log files, memos, presentation material, white papers, etc. When performing literature studies, document your search strings, databases applied in search and snowballing procedures, if any.

Position of Research Questions or Aims

There are two basic models for positioning research questions in an article or a thesis: before literature (previous work) or after it.

The research question can be presented in Introduction before addressing previous work when it is understandable without previous scientific work. Typically this is the case when the research problem comes from elsewhere than scientific literature, e.g., from practice in an organisation, or if it is otherwise a generally known issue.

A natural place for the research questions is after addressing previous work when your work addresses an identified gap in the literature. Here, previous work is discussed in a way that leads to the identification of a gap in the previous knowledge. Research questions are then placed so that you aim to fill the gap. Depending on the details of the structure, the Research questions may be placed (late) in Introduction or as a separate section after Previous work.

Results

Figure 1 highlights the line between Results and Discussion. This is sometimes hard to understand. The main idea is to make it explicitly clear what was found as results and how those results are discussed. Section Results avoids discussion, commenting and any speculation when presenting the results. On the other hand, Section Discussion (discussed below) only discusses results presented in Results – it does and not bring in any new results.

In a thesis document, you may have multiple sections for reporting your results. A general guideline for a Results section(s) is to write the facts, e.g., what you built or what you observed and what you analysed from the data observed.

While reporting the results, you should keep looking back to your research questions to see that you write about answering them. You may need to rephrase your research questions, if

your results do not provide answers to them and a different wording may better describe your results. Just remember to maintain consistency between the different parts of your thesis document.

Discussion

In Discussion, you take a look at what you achieved as reported in Results (not necessarily in the order shown below). You reflect your results to the research questions basically to discuss how well you were able to answer them. Typically, you will discuss the validity and generalisability of the results and compare with related work.

Validity

You need to very honestly bring up any potential problems in terms of the validity of your results. For example, if you interviewed some persons to understand the problems, you need to think carefully whether you had the right person to talk to, they understood your questions correctly, you understood their answers, you have made correct analysis and conclusions of the answers, and so on. If your results are a designed artefact, you need to explain how well the design solves the original problem and how you know that.

Explain all you did to improve the validity of your results, such as justification for the selection of the interviewees, how you tested your questions, the analysis process you used, have you presented the results to the original informants asking if they agree with your interpretation and conclusions, how you tested, measured or otherwise validated your design, etc. Your research approach and study design play an important role here. Nevertheless, remember to be humble with your results, as absolute truths are not that common in software engineering research. But on the other hand, be proud of your results you have worked hard to get. Discussing the validity is not about bashing your work as hard as you can, but to provide the reader justified and intelligent views on your results.

Related work

You also discuss your results in the light of related work. That is, what similar work have others done and how do your results compare with theirs.

The term *Previous work* is used in these instructions to refer to the literature that is presented as a basis for your work, whereas the term *Related work* refers to the comparison of your results with the work of others. Therefore, Related work is placed after the Results section. Depending on your topic, Previous work may be rather brief and does not deserve a section of its own – you can embed it in the Introduction. Similarly, brief Related work may be embedded in Discussion. So, it depends on your topic whether it makes more sense to present the literature mainly as Previous work or as Related work, or to have them both in a prominent role.

Generalisability

Depending on the context of your work, you should also consider generalisability of the results. For example, if you learnt some lessons in a case study, consider if those might be

applicable to other companies as well. Think what was special in your case? What seems typical or common to other similar cases? What lessons or findings from your work could be applicable in different contexts? Give your suggestions and justify them.

Abstract

It is often useful to structure the abstract so that the reader can get a concise overview of the paper. The basic structure can follow the IMRAD format; essentially, the abstract is an extremely condensed version of the paper. Some forums have adopted a so-called structured abstract, which may have the following sections:

Context - the context or background of the study.
Objective - the aim(s) of the study.
Method - the method(s) used to address the aims.
Results - a very concise summary of the most important result(s).
Conclusions - a very concise summary of the implications of the study.

The sections can be written out in the abstract using a colon after each section name (e.g. Conclusions: …). No line breaks are required, in order to keep the abstract compact. Even if not using the section names, using a similar structure can help write a good abstract.

Abstracts should not include references. They should stand on their own without the rest of the article. The reference list is not necessarily available when a reader scans titles and abstracts for interesting articles.

Conclusions

Select and write concluding remarks on your work with the perspective the reader has at this point, that is, after the results and discussion. Summarise briefly what you did, but keep in mind that giving conclusions is not the same as simply summarising all that has been said earlier. In layman's terms, you should answer the question: "So what?" It is important to think over and clarify the central results / main contributions of your work. Consequently, write them crystal clear in Conclusions. It may help to revisit your research question(s) to ensure that your contributions have relevance for the reader who was interested in the answers when starting to read the paper.

Finally, some ideas about new research questions may have emerged during your project. You can now give some suggestions on how you or someone else could carry on doing further research based on your work.

A typical length of Conclusions is half a page or less in articles and 1–2 pages in a thesis.

Variations to IMRAD

Wikipedia:
Usually, the IMRAD article sections use the IMRAD words as headings. A few variations may occur, as follows:

- Many journals have a convention of omitting the "Introduction" heading, based on the idea that the reader who begins reading an article does not need to be told that the beginning of the text is the introduction. This print-era proscription is fading since the advent of the Web era, when having an explicit "Introduction" heading helps with navigation via document maps and collapsible/expandable TOC (Table of Contents) trees. The same considerations are true regarding the presence or proscription of an explicit "Abstract" heading.
- In some journals, the "Methods" heading may vary, being "Methods and materials", "Materials and methods", or similar phrases. Some journals mandate that exactly the same wording for this heading be used for all articles without exception; other journals reasonably accept whatever each submitted manuscript contains, as long as it is one of these sensible variants.
- The "Discussion" section may subsume any "Summary", "Conclusion", or "Conclusions" section, in which case there may or may not be any explicit "Summary", "Conclusion", or "Conclusions" subheading; or the "Summary"/"Conclusion"/"Conclusions" section may be a separate section, using an explicit heading on the same heading hierarchy level as the "Discussion" heading. Which of these variants to use as the default is a matter of each journal's chosen style, as is the question of whether the default style must be forced onto every article or whether sensible inter-article flexibility will be allowed.

For more details, see Wikipedia:
https://en.wikipedia.org/wiki/IMRAD
https://fi.wikipedia.org/wiki/IMRD-rakenne

Referencing

References have a very important role in scientific writing. Fundamental scientific principles rely on building research on others' work and correcting or refining previous work whenever needed. In addition, as novelty is one of the essential characteristics and valued aspects of scientific work, proper acknowledgement should be given to all previous and related work. Appropriate use of references makes all this explicit. Furthermore, based on your selection of references, the reader can determine your maturity in positioning your work in the underlying field of research.

Well-written scientific text uses references in a manner that clearly communicates the necessary information and links it to the work at hand without unnecessarily breaking the text flow. Necessary information may include, for example, the use of the referred article for justifying your claim, quality and reliability of the referred work – all publications are not created equal. Achieving this is hard and requires some practice. We try to give some practical advice below.

- **Cut non-informative clutter.** For example:
  "Möttönen et al. (2012) write in their publication that pink is a challenging colour."
  can be simplified to:

"Pink is a challenging colour (Möttönen et al. 2012)."
The meaning of a reference at the end of the sentence is that this is what Möttönen et al. say in that publication. In any case, be careful not to simplify to the point of misrepresenting the original information.

- **Prefer information-prominence over author-prominence.** Using an author's name in text, e.g., as in "Möttönen (2014) writes...", tends to indicate a reason for doing so, which may be to highlight it is Möttönen who said so. For example, if you tend to disagree with the authors, there is something controversial or unconvincing in the statement made or research conducted, or the point is merely an opinion and you want to highlight that. Generally in scientific writing, it should be of more important what is said than who said it. Sometimes, however, you may need to refer to Möttönen's results later in the text and thus use the authors' names as a means of reference. Even in this case, other options can be considered, e.g., name of the research group, country of the research (e.g., "a Finnish study"), etc.

  In addition, it is often far more important to reveal the context of the research or results. Consider the following examples, as they could contain some much more important piece of information to the reader than Möttönen's name:
  - "The participants of a boxing event found pink challenging as a colour (Möttönen et al. 2012)."
  - "The developers in a Finnish game company found pink positively challenging (Möttönen et al. 2012)."

Regarding the formatting of the references, we strongly encourage you to use the Author-Date model (e.g., Möttönen et al. 2012), particularly in a thesis. You may use, for example, one of the following guidelines for details:

- A good and clear guide: Harvard Referencing Guide by Monash University (for a quick guide, check the Appendices of that document)
- Often used Guide of APA (American Psychological Association): An Abridged Guide to the APA.

Details and other instructions

Allocate the available space: Page budget

When planning your writing, it is a very good idea to draft a page budget for each part of your article or thesis. First, find out or decide the maximum or target number of total pages in the required format. When allocating this to the individual sections, you may realise that there is not that much room for something. It is of course better to realise this earlier than later. You may use previous work of the same kind as a reference.

Let's consider a six-page, two-column article. Roughly, the typical numbers are:

- Title and author information, etc. 1/3 of the first page,
- Abstract 1/4 of a column,

- Introduction fills the rest of the first page,
- Previous work (if not in Introduction) 0,5–1 page,
- Methods 0,5–1 page,
- Results 2 pages,
- Discussion 1–2 pages,
- Conclusions with References: the final page.

These rough figures already give some basis for considering how to allocate the available space and if something needs more or less emphasis, e.g., more rigorous description of methods is required, how to adjust the whole.

In a thesis, the total number of pages typically has no strict upper limit. It is more important to maintain a good balance. One practical piece of advice is to first figure out how much space the results require. Then allocate adequately for Discussion; it is an important part demonstrating the nature and relevance of your results and your own maturity within and around the thesis topic. The length of theoretical background is constrained by remaining available space. Select what the reader needs from what you have read and what fits in. A common mistake is to write too detailed previous work, e.g. 40 pages, only to find out that there is not enough space. This is important, as summarising existing literature is often relatively easy and produces a lot of text. Do not just summarise: write a critical own analysis of the literature as a background and motivation to the rest of your work.. In other words, do not try to write a textbook on your thesis topic.
For a MSc thesis aiming at 50 pages, you can start, for example, by allocating some 10-15 pages each to Previous work, Results and Discussion. This may leave about 10 pages to the rest, i.e., Introduction (3–5 pages), Conclusions (1–2) and References (5–10). The length of references is not that critical, as the appropriate number of reference is determined by other factors.

Figures and tables

Scientific results are often summarised as figures and tables. Table 1 summarises criteria for selecting between presenting data as a table, a figure or text.

Table 1: How to choose between tables, figures and text to present data (Adapted from Editage Insights, tips on effective use of tables and figures in research papers, retrieved 2016-08-19)

Use a Table	**Use a Figure**	**Use text**
To show many and precise numerical values and other specific data in a small space	To show trends, patterns, and relationships across and between data sets when the general pattern is	When you don't have extensive or complicated data to present

	more important than the exact data values (what to use:graphs and data plots)	
To compare and contrast data values or characteristics among related items or items with several shared characteristics or variables	To summarize research results (what to use: graphs, data plots, maps, and pie charts)	When putting your data into a table would mean creating a table with 2 or fewer columns
To show the presence or absence of specific characteristics	To present a visual explanation of a sequence of events, procedures, geographic features, or physical characteristics (what to use: schematic diagrams, images, photographs, and maps)	When the data that you are planning to present is peripheral to the study or irrelevant to the main study findings

- The difference between own original contribution (own figures) and figures from other sources is done by a reference.
    - Basically: no reference in the caption means the figure is presented as an original contribution!
    - In principle, it is a serious error, even if accidental, to give the impression that a figure or graph taken from elsewhere would be your own contribution.
- A figure and its caption together should form a self-explanatory whole that is understandable even without reading the text.
- Figures should be redrawn and elements not supporting the text removed. Potentially something is reorganised or further annotated with the terminology of your article. The original is referred to, e.g. as:
    - (modified from Möttönen 2012) or (adapted from Möttönen 2012)
- Long lists are not very readable, so include only the necessary and make the text readable, not just a bulleted list.

Language

Use grammatically and idiomatically correct written language. Do not use slang terms, colloquialisms, or informal expressions (unless quoting verbatim). However, do not try to make the text fancier by introducing difficult words or sentence structures, as this will only make the text difficult to read.

As a general rule, use past tense (=imperfect). Examples of usage:

“The study was planned/conducted/analysed in the past.”

> “In this paper, we present the results.” OK, as the presenting is done in the paper.
> “In this paper, we analyse the data from...” NOT OK, as the analysis was done earlier and only presented in the paper.

Sometimes confusing: “In this thesis...”, as one can understand this both as referring to the thesis project or the written thesis. Best to use only to refer to the written document and clarify, e.g., as “thesis project” if that is what is meant.