

How to write a Ph.D. thesis

by Prof. Dr. Reinhard Jahn

Preliminary remarks

Theses are scientific papers that need to adhere to the standards and to fulfil the criteria of a scientific publication. Keep in mind that the rules of Göttingen University require that your thesis will be publicly accessible, either by obtaining one of the copies, or (if you opt for electronic submission) through the internet. Although exceptions apply, theses are usually divided into the following sections:

- Title and author(s)
- Introduction
- Materials and Methods
- Results (incl. Figures, tables etc. and legends)
- Discussion
- Summary (abstract)
- References
- Acknowledgements
- Supplemental Material, Appendices, list of abbreviations (where applicable)

Furthermore, a thesis needs to meet certain formal requirements that are outlined in the Doctoral Degree Regulations (RerNat-O). Please also note that special regulations apply for chapter-based dissertations.

In the following paragraphs, you will find some comments as well as Do's and Don'ts. When you write your thesis, I strongly recommend referring to this text for guidance.

1. Introduction

The purpose of the Introduction is to guide the reader from the knowledge level of an advanced textbook toward the specific question you are addressing. Ideally, the Introduction is constructed like a "funnel" in which, logically and systematically, the focus is becoming narrower until the question you are addressing becomes almost inevitable.

Literature review

It is mandatory that the relevant literature of the field be quoted. You may resort to reviews in the first part of the Introduction, but it is necessary to discuss individual publications relevant to your topic. If in doubt, you should discuss with your supervisor which papers you need to deal with.

Appropriate citations

Every statement that you are making and that goes beyond basic textbook knowledge needs to be supported by an appropriate reference. However, unlike in the humanities verbal quotations from other papers are highly unusual in our field. In fact, they are strongly discouraged – it is your job to rethink the state of the art and then phrase it in your own words. Verbal copying of paragraphs or even single sentences (“this review is great – this is exactly what I need to say in my Intro....”) is not allowed if not clearly indicated as a verbal quotation with the source given. In fact, such behaviour is branded as plagiarism and comprises a serious violation of the standards of Good Scientific Practice. If someone copies texts from reviews or webpages without giving the source and if this becomes known after the thesis is completed and the degree is conferred, it will have serious consequences. In severe cases the degree may be formally de-recognised.

This also applies for cases of ‘self-plagiarism’: you have to correctly cite your own work (text, figures, data, ...) from published or as yet unpublished manuscripts, Master’s theses, etc.

Another issue is which paper to quote for the discussion of a finding to which many groups have contributed. To some extent, this is a matter of taste, but you should make an effort to identify key papers instead of quoting someone who has only contributed to in a minor way. Finally, I guess we do not need to discuss factually wrong references that do not contain data relevant to the statement you are making. Do not quote papers you have not read or at least scanned. This is highly non-scholarly and cannot be excused by tight deadlines. In fact, such behaviour may backfire sooner than you think...

Extent and scope

There are different opinions about this topic, and you should discuss with your supervisor how extensive the literature review should be. All aspects that lead to the work must be integrated without losing sight of the focus. It may be useful to incorporate one or two schemes and diagrams that can be taken (and, if required, adapted) from publications (appropriately referenced). Keep in mind, though, that an Introduction is not a textbook chapter. The level of knowledge you should expect from the reader is that of a non-specialist biologist, chemist or physicist with a MS/diploma degree – no more, no less. Thus avoid expansion into topics that are not directly related to your work, and do not reproduce textbook pictures, proofs, derivations etc. Rather, these should be covered by references to the appropriate literature (including textbooks). Everything beyond that level of knowledge needs to be explained. In physics-oriented theses, it is particularly important to explain concepts which are essential for a proper understanding and appreciation of the questions addressed, the techniques applied, as well as of the results obtained. If needed, a special “Theory” section may be included after the general Introduction. Here, explanations in terms of physics should be preferred over formalistic-mathematical language wherever possible. The latter might be more precise but is often not the best means to efficiently communicate the ‘physics behind it’.

It is a good practice and helps the reader if the Introduction ends with a paragraph such as “Scope of the thesis” in which you succinctly (no more than one page) summarize the main goals of your work.

Timeliness

An Introduction needs to consider the latest available papers. Often, relevant knowledge becomes available during the time of your thesis (sometimes from directly competing groups) that was not

yet published when you started your work. Such facts need to be mentioned. For instance, you can state that at the time when you began your work the problem x was not addressed but that during the course of your thesis new information became available etc. If such papers, however, affect details of your results but not the scope of your work, the arguments relating to them may fit better into the section "Discussion".

2. Materials and Methods

An essential (but unfortunately frequently violated) principle is that the experimental procedure given in this section must be described in sufficient detail for allowing a trained student or professional scientist to reproduce the experiment. This means either that the description itself should be comprehensive, or else, that a reference to a publication is provided that contains such a comprehensive description. I know that many students hate this part and omit details that are obvious to them at the time of writing. However, I assure you that many of you will be unpleasantly surprised if after a year or so you go back to one of your old descriptions and try to repeat an experiment that you have not done since! In a PhD thesis you should discuss with your supervisor how detailed the methods description shall be. Scientifically, it is not necessary to provide details if appropriate references are available and quoted. However, these often are used as source for future students and postdocs. Therefore, some supervisors prefer a full description of methods even if they include published standard procedures.

Appropriate referencing

There are few areas where I have seen more carelessness than in the referencing of methods. A full description of a method does **not** obviate the need to provide appropriate references! For instance, if you do not provide a reference for your description of SDS-PAGE or PCR, the impression is created that this is a new method that you have invented during your work. Again, this practice, commonly seen even in our graduate programs, is not in compliance with Good Scientific Practice. A thesis that does not contain references in the Methods section will be given a lower grade.

What is an appropriate reference? Often, a method was originally described by someone and then changed in countless details by others. For standard procedures, it is perfectly legitimate to quote lab manuals (e.g. the "Maniatis"), but if any even seemingly small detail was changed, it needs to be stated ("ZZ was performed according to X, with the following modifications" etc.). It is also acceptable to quote a method, particularly if not a standard procedure, in the format "was performed according to X with the modifications introduced by Y". Is it necessary to quote X? Formally not, but unfortunately nowadays scientists are frequently rated by the number of quotations their paper receives. If X is not mentioned, Y gets all the credit even if his/her contribution was small with respect to that of X.

Check your references! It is a common but deplorable practice that students copy methods references from previous theses without checking. For instance, the famous paper by Laemmli in 1970 is frequently quoted as the sole reference for SDS-PAGE. I recommend looking up the paper - decide by yourself whether this paper is sufficient for a fellow student to run gels the way you have done it. When I wrote my Diploma thesis in the 70ies, purification of mitochondria was a standard method in the lab and everyone quoted an old paper from the forties. When I checked, I

found out that every aspect of the method (starting from the buffers) was very different. I found a paper that was much more appropriate (almost identical) to our procedure that I then quoted, only to learn years later that every subsequent student now copied “my” reference although the method was changed again...

3. Results

The result part needs to be divided in logical chapters, and it contains the data you have obtained during your research. There are differences in style as to how to write this part and how to prepare and label figures. You need to check carefully with your supervisor to make sure you agree on the structure of this section and the display items you want to present.

Outline

I strongly recommend that you first draft an outline of this section in which you delineate each chapter and subheading, clarify the logical flow of arguments, and decide which figures should go into which paragraph. Discussing this outline with your supervisor will probably save you a lot of work and frustration later!

Figures and other display items

Before you start to write even an outline, you should have the figures, tables and other display items either finished or at least drafted. When preparing display items, keep in mind that they should be clearly labelled, and ideally be understandable without having to refer to a long-winded and complicated figure legend. For instance, if multiple panels or lanes are shown, label them to indicate what they show rather than coding them with numbers or letters. This way, the reader does not need to go back and forth between legend and image multiple times to understand what a single panel is supposed to show. This does not mean that one should clutter a figure with text. Also, try to avoid using different lettering sizes and fonts – keep it as simple as possible. Also, adjust lettering in such a way that in the final print size the letters and symbols are readable and of identical or similar size.

Every figure that you show must correspond to original data in your lab notebook, and these data need to be accessible and intelligible to an outsider. If data from multiple experiments are used for a single figure, or if extensive calculations of the original data are involved, I strongly recommend to keep a folder in which copies of all original data (with reference to the corresponding page in your lab protocol) are kept, and in which each step of the calculations is recorded so that the figure can be traced to the original data by an outsider. The rules of Good Scientific Practice require that all original data are stored for 10 years and that all data used for figures or tables in theses and publications must be traceable to original protocols, with all calculations (e.g. for statistics) being on record! Inquire in your institution to which extent such data can be stored electronically, and how such data need to be indexed.

Description vs. interpretation of the data

The Results section is reserved for the description of the experimental data and should not contain extensive interpretation. However, the logical backbone and the flow of arguments need to be evident. For instance, if you formulate a hypothesis, and you show experiments proving it, you need to state this clearly. Also, conclusions you draw from results that are the basis for the next set of experiments need to be stated. For instance, you may start a new section in something like the following manner: “The immunoprecipitation data shown in Section 3.1.3 have shown that protein X and Y form a complex, a view that agrees well with the colocalization of the proteins shown previously by Smith et al. (2003). However, as stated in the Introduction, the data of Wang et al. (2005) are difficult to reconcile.....I therefore asked whether etc.” This way the reader can follow your line of arguments, and if the results are clear, it may not be necessary to return to this aspect in the Discussion (see below). Furthermore, it is often helpful to deal with small accessory arguments in the Results rather than the Discussion, thus allowing you to focus on the major issues in the Discussion rather than dealing with petty details.

The data should be succinctly stated, and you should avoid getting lost in details that do not contribute to the understanding of the experiment and the results you have obtained.

Recognition of contributions by others

In every thesis, you formally certify that the work has been completed by yourself and that no additional assistance has been involved except where stated (see above). However, science is a team effort, and it is perfectly normal and often even essential to collaborate with others to achieve results. While in publications such contributions can be dealt with by including your colleagues as authors, how to deal with this issue in a thesis?

Actually, this is not a difficult matter. Every direct contribution by someone else must be stated at the appropriate place. For instance, if you show an EM figure of a sample you have generated, but the processing for EM and the actual analysis was performed by someone else, state it in the figure legend and in the text. Examples: “Fig. A: Electron microscopy of fractions of early endosomes isolated by the standard procedure (...). Sample preparation and imaging was carried out by Dr. A (Department, Institute).” “All purified proteins to be used in the experiments described in the following section were provided by Dr. Y (Department, Institute)”. If the contribution of another scientist to a data set is important, it is not sufficient to relegate the acknowledgement of the contributions to the “thank you” section – it must be stated in the text and/or the legend to the appropriate figure. Make sure, however, that you keep the use of data contributed by others to the necessary minimum. The same applies if experiments were carried out together with a colleague.

4. Discussion

It is the purpose of the Discussion to evaluate and interpret your results against the background of the relevant literature. For beginners, this is often the most difficult part, and it is also the last one you should write.

In preparing for the Discussion, I recommend collecting ideas and arguments in an unstructured manner while you are working on the other parts of the thesis. The most important next task is then to structure the Discussion and to define a logical order. This usually takes some time, and you

should discuss your ideas with colleagues and your supervisor in order to obtain some feedback before you start writing.

Discussion of data vs. repetition of the results

In the Discussion section it is inevitable to return to your results as a point of departure for your ensuing arguments – after all, this is what the Discussion is all about. However, a common mistake made by many students and even some experienced scientists is to write lengthy repetitions of the data instead of interpreting them. For instance, if you have proven a protein-protein interaction by several complementary methods, and the data are consistent and clear, there is absolutely no need to repeat in detail every experiment supporting your finding. Often it is sufficient to refer to the appropriate figures in the Results section. Rather, you should concentrate on discussing the significance, i.e. why this interaction is important and what you have learned about the biology or physics of the particular process you set out to investigate. It does not hurt to begin the Discussion with small paragraph in which you summarize your main findings and conclusions. However, the reader of your thesis can be expected having just read your Results and thus being aware of the data. Such a reader will be terribly bored by detailed repetitions. Methods discussions may be important in certain cases, but these should be put into a separate section and dealt with at one place. As with all rules, there are exceptions, for instance, if method development constitutes the major part of your work.

Appropriate review of the relevant literature

Essentially, the same criteria apply as in the Introduction. Use the best of your efforts to identify all papers that report data related to your own findings, and to discuss them appropriately. It is mandatory for every scientific text you are writing that you clearly identify what is novel in your own findings vs. what is confirmatory (or contradictory) to published work. With easy access to searchable literature databases there is no excuse anymore for overlooking a paper important to your work. Obviously, no one is free of errors, but deliberate omission of a paper (particularly if it is published in one of the major journals) in order to hide that your “novel” findings are not so novel anymore, is considered as a serious breach of Good Scientific Practice that may have negative consequences (see above). It is your responsibility to make sure such an oversight does not happen. If such an omission is part of a manuscript to be published, this is usually found out rather quickly, often already during the peer review process.

...and in the end...

- Write clearly and succinctly, and avoid long-winded and complicated sentences!
- Keep the use of non-standard abbreviations to an absolute minimum – there is no need to define an abbreviation for a term that is used only a few times in the text!
- Try to reserve a block of time in for continuous writing – it is much more difficult to get back to a partially finished thesis after a lengthy interruption!

And: **GOOD LUCK!**