How To Write A Phd Dissertation (aka Thesis)

David J. Greaves (and Simone Teufel)

Computer Laboratory
University of Cambridge, UK

RSP Unit 12th Feb 2025 - Rev A

https://www.cl.cam.ac.uk/ djg11/pubs/how-to-write-a-phd-dissertation-djg.pdf

What is a dissertation/thesis?

The document:

- A document of 100 to 200 pages for the PhD examination.
- Can also serve for other purposes.
- The **thesis** is the new concept, proof or result that you alone are contributing to the world. It is a statement.
- The **dissertation** is the main documentation of this concept.
- It will be discussed at your Viva Voce and likely then revised to some small extent.

The dissertation title can often be turned into a thesis by adding the words '... is a good idea'.

How is good scientific writing created?

- Work — namely, thinking work
- Goals:
  - ► Clarity
  - ► High proportion information $\frac{\text{information}}{\text{ink}}$
- Not goals:
  - ► Beauty
  - ► 'Style' (will follow naturally)
- First drafts are always bad

Teufel recommends 'Writing for Computer Science' J Zobel.

And 'The thinking person's guide to Writing in 21st Century' S Pinker.

Bridging Sentences

Narrative Structure

- The dissertation should tell a story.
- Each chapter should contain opening and closing sentences that narrate.
- It is well worth writing these sentences in advance of anything else.

Think also about the very last sentence of the conclusion (but before any Further Work section).

This should strongly relate to the dissertation title: proving your contribution.

Chapter Structure

Ideally, you have two peer-reviewed papers published and another one nearly ready to submit.

Example ToC — other structures work perfectly well

1. Introduction
2. Previous work
3. Meat 1 - Paper 1 content (largely rephrased but with the same results)
4. Meat 2 - Ditto paper 2, but using a (totally) different approach.
5. Meat 3 - A variation or extension of meat 2.
6. Conclusion.
7. (You can typically include a 'further work' section after the conclusion text, but this should not be in lieu of a proper contribution up to that point.)

Do not blindly paste introduction (or any) text that appeared in each standalone paper.

Introduction Chapter

- Define and explain the problem to be solved
- Canonicalise the possible/academic space of possible solutions
- Explain and justify why you have narrowed the space down for the work presented here
- Finish with a list of personal peer-reviewed publications. Do not worry if you don't have any!

Do not explain standard terms of the art - the reader is a field expert.

Do, here or next chapter, define (concretize) your choice of notation for the standard terms and all new ones. Make sure you stick to these names and symbols throughout.

Any explanation of your own approach and results will be intrinsically redundant here, so keep this to the absolute minimum. Otherwise it can disrupt the narrative.

Prior Work Chapter

Ultimately you have to show clear water between your work and that before:

- Give a summary and criticism of the assumptions and approaches of others to the same problem.
- List any new disruptive events that invalidate or facilitate new approaches.
- Do not even tacitly present your own approach under the prior work heading.

Never here present your own chosen approach as a fait complete with no forward reference to where your decisions will be justified.

It is fine to parenthetically point out differences that, as will be seen in Chapter X, did not accord with my own analysis or work. But better to make that point in the justification for your approach when you properly present it.

Definitions and terminology — recap

- Each technical term you use needs to be explicitly defined
  - You don't have to follow definitions from the literature if you have a reason to go against them
  - Just make sure we understand what your definition is
- The only unambiguous way to define is to say 'X is ... '
- If a sentence uses any other verb, e.g. one that describes X's behaviour, it is not a definition
- The lowest-class citizen in the world of definitions: the definition by example
- Definitions deserve our full attention; be prepared to invest thinking time and space

Active or Passive

Voicing

- Traditional cumbersome passive-voice paper phrasing is well out of fashion
- But remain formal and only use the first person where your personal behaviour or notions are important to the narrative.
- Avoid 'got', 'gonna' and all colloquialisms.

Make sure all comparatives have an explicit second party.

Avoid marketing talk and meaningless superlatives.

Avoid TSJWC (even though the BBC has banned the semicolon).

Evaluation Benchmarks

Benchmark Justification

- A description of benchmark suites and tests can usefully be introduced in Chapter 2, since they are likely to be used over all three central chapters.
- Appropriateness of the benchmark selection should always be **justified** (even if many are standard in the field).
- It is helpful to analyse the benchmarks using various metrics.
- Important to demonstrate they are realistic and give comprehensive coverage of the metric space.

Central Contributions

Define your contribution

- Explain what you did and justify decisions made.
- Explain when and how your results are an improvement
- Sensitivity: what are your partial derivatives?
- Robustness: determine confidence intervals or other appropriate metrics
- Repeatability: is everything archived so your results can be reproduced?

As part of your narrative: explain the societal or other disruptive effect of this result

Defining your contribution

- There are only two types of contribution: **First** or **Best**.
- Use restrictors wisely:
    - Too many and your claim will be weak
    - Too few and you might end up with an unsupported claim
    - There are also 'negative' restrictors such as 'without'.
    - When there is more than one comparison metric, there will be different ways to cut the space with restrictors
- Use *and that was the first time this was done!* as a test
- No need for the word *novel* if you do it right
- The methods you use may or may not be part of the contributions — be clear on this.

Tools

- Use LaTeX and BibTex — start your personal Bibtex folder at the earliest possible date and grow it throughout your academic life
- Use `citep` or `citet`, depending on syntactic context
- Use revision control (eg git).
- Learn how to make nice diagrams.
- Use a flow diagram wherever possible to present project structure.

A presentation slide deck may be invited for the start of the viva:

- If preparation of this raises the slightest difficulty, your dissertation is likely lacking!
- Don't expect to get to the end of the presentation before the end of the viva.

Miscellaneous

IP Rights

- IP and publication restrictions - dissertation must unencumbered, even if viva has restrictions
- You own the copyright of your dissertation and you should insert a relevant notice (the University house style often claims copyright of all content on any web page, which is wrong)
- The relevant animal is Callum's dog Valentine (woof).

Where a group of people have worked on a project, for examination purposes, the dissertation needs to clearly show and delimit the author's personal contribution.

Overall

- Choose your topic wisely.
- Develop a strong thesis statement.
- Develop your narrative and bridging sentences as early as possible
- When running experiments, make sure you collect the results needed for your narrative. False paths are common here!
- Proofreading: When someone suggests a change, even if you disagree with their suggested change, remember some deficiency on your part likely prompted them

The Cambridge regulations do require you to know how your PhD work fits in with the general field and how it is likely to influence it. And also what else is happening in the more general field. You can expect to be asked about that at the viva.

Conclusion

A PhD is

.   1/ a demonstration of a new contribution to knowledge and

.   2/ an explanation of why it is useful.

(C) DJG 2025.