Thesis-writing advice

Or ‘How I learned to stop worrying and love the examiner’

Andy Buckley,
University of Glasgow

LHCb UK seminars
29 June 2023

University of Glasgow

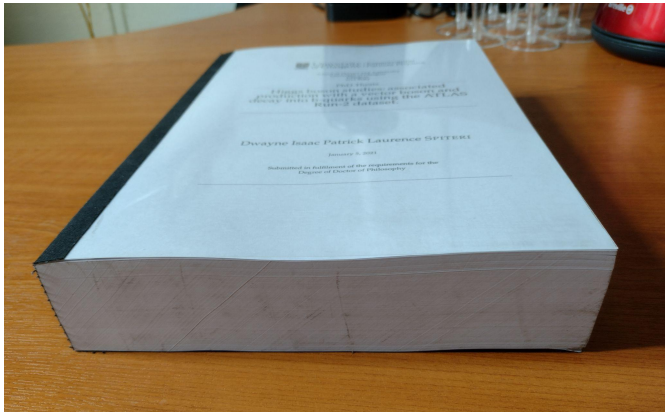
Why does my opinion matter?

- **I'm not going to claim any unique expertise.**
  Only weirdos* write more than one PhD thesis! $\Delta N_{\text{thesis}} = +1$
- But I've examined quite a few… around 16, as both internal and external examiner (according to my vivas/ folder)
- Plus — probably why I'm here — my public thesis gotchas advice document: https://zenodo.org/record/4013748
  - ➢ I actually made a short version of this for the UofG MSc students' report-writing, pasted at the end of these slides
- You might also know me from a variety of LaTeX packages, e.g. hepthesis, hepparticles, hepnames, maybemath…
- Mainly I'm going to talk about the big-picture bits of the gotchas doc. Please ask questions: this is opinions, not theory

* No offence intended to weirdos. You do you

Communication

- **The most essential things every time you need to communicate complex ideas:**
  - Know the audience
  - Tell a story
- Your examiner is the main audience. You need to tell them, clearly, why you deserve a PhD.
- Without a sense of narrative, a thesis will draaaaag. Which is hard work for an examiner →pain for you!
- Can you leave stuff *out*? Physics context beyond your work is needed, but don't set yourself up for awkward viva questions. "Anything you say may be used as evidence…"
- Explain the ideas, as well as the detail. *Why?*



I call this the Mountain of Dwayne. Some others have gone to multiple volumes. Please don't do this! 200+ page theses mean 6+ hour vivas

It’s not a paper!

- **Again, it’s communication: a paper is a polished record of a completed scientific study. Your thesis is actually about *you*.**
- It’s tempting to write “cargo cult” paper-style, especially because there are many bad papers themselves aping “desirable” style. But the best route is to explain your process, including obstacles, some attempts that didn’t work, how you responded. Say what *you* did.
- Many papers are heavy on jargon, assumptions, etc. That’s not great, but can be tolerable. Don’t do that in a thesis: minimise jargon and acronyms, explain the ones that are unavoidable, and again, focus on clearly stating the problem, how you solved it, and the results ⇒ a semi-personal scientific narrative.
- Explaining something complex without jargon is *hugely* more convincing than an impenetrable mass of specialist terms. And examiners will dig into your understanding if given the latter.

South Pacific cargo cults formed after WW2, in communities that benefited from use as airbases and staging posts and the resulting flown-in “cargo”. In attempts to bring back the cargo after demobilisation, without understanding the big picture, many communities built remarkable facsimiles of radars, planes, control towers, etc.

Presentation & nitty gritty

- **This doesn't really seem like the forum to go into fine detail about every typographic gotcha!**
- But I should make the point about why it matters: because — again — it's about efficiently giving the reader what they need.
- If the examiners are fighting past inconsistencies (use macros), unhelpful formatting, bad grammar, they miss your point (and get annoyed)
- A few bees in my bonnet:
  - Spell (and grammar) check! I mean, FFS…
  - Language quality in general. Read. Proof-read, and get friends to help
  - Equations are part of sentences: space, punctuate, etc.
  - Math mode for symbols. Text in equations is text: use \mathrm.
  - TeX is not always right: use manual spacing like \! \, \; \: ~ to improve
  - Capitalise references to Section 3.1.2, Figure 5, Table 1, Ref [x].
    Use non-breaking spaces to keep compound references (& units) together
  - Speaking of compounds: hyphenate compound adjectives, or rephrase!
  - Figures and tables: tbp position, resolution (use PDF), labels, good captions; column alignments

Conclusions

- **First, decide on who you are pitching to**
  - Communication is key: know your audience… the examiners
  - This is a *very* general point… so, a transferrable skill: thesis, papers, talks, discussions, code…
- **A thesis is about *you* as well as the science**
  - Don't worry that your ideas aren't smart enough and need to be hidden in sciencey-looking impenetrable prose. It's not a paper
  - Convey what you personally did, what problems you solved
- **More generally, tell the examiner what they need**
  - All to the key question: did this person do stuff at PhD level?
  - Avoid unnecessary invitations to ask a
- **Do sweat (some of) the small stuff**
  - If the presentation, typography, language, figure labelling, etc. are above threshold, the examiner can focus on the content
- **Start early, with notes in thesis.tex! Enjoy!**

PS. use this!
https://inspirehep.net/bibliography-generator

A long tl;dr: the “top ~10”

- **A “short” precis of https://zenodo.org/record/4013748**
  - 1. Know your audience: The absolute number-one rule in all communication is to know who you are communicating to. This applies to presentations, papers, and even emails as much as it does to a PhD thesis. What do they need to know? What do they already (not) know? What's most important, and what can be left out? Will they be looking for particular angles that you can anticipate and satisfy? Often the audience is "someone at my level in this general subject, but not specialising in the same area".
  - 2. Explain the ideas: A dump of obscure-sounding jargon found in existing literature will convince no-one that your understanding is high-level. Really good work explains even complex ideas in relatively simple, clear and intuitive terms. Even better if it can have a narrative, i.e. a coherent story is told from start to finish, anticipating readers' questions and confusions and answering them at the natural points. Do not assume that jargon or ideas specific to the area will be known to your reader, including examiners: in particular, define all acronyms on first use.
  - …

A long tl;dr: the “top ~10”

- 3. Check the language quality: In particular, always run a spell-check program (in the target language) over your text just before submitting. Especially if not a native speaker of that language, but also if not entirely confident about flowing and coherent writing, appropriate academic tone, and more advanced punctuation and multi-clause sentences, use a grammar checker (some online ones cover a very wide range of issues now) and ideally get someone to proof-read for you -- offer to do the same for them.
- 4. Consistent presentation and attention to detail: Specific typographic rules for the most part do not matter (except for mathematics, where there are established connections between form and meaning), but once you have chosen a presentation style, stick to it. Changing styles through the document confuses the reader and distracts from the content. If you have a bit of design sense, choose attractive and reasonable fonts and layout, but default styles are usually good. Make sure that your document is appropriately sectioned, with clear and consistent section headings that are distinct from the text. For mathematics, make sure variables are italic (even when used inline in sentences); any text in equations is upright, including in subscripts and superscripts; and that the equation is part of the surrounding text sentences, including punctuation. Do not define the variables in an equation in a new following sentence starting "Where...". Values with units also have rules: put a space between the value and the unit, and the unit should be in the same text style as the surrounding text, not e.g. italic.

A long tl;dr: the “top ~10”

- 5. Figures: Good figures are essential to maintaining reader attention, and can be essential to helping them understand concepts which are convoluted and awkward in text alone. Make sure they interact with the text, i.e. are referred to explicitly by the body text, and are explained in a sufficiently detailed caption. Make sure the image resolution is sufficient and the display size appropriate: there should be no blurring or unreadably small text. For diagrams and plots, as opposed to photographs, vector formats like PDF are far preferable to pixelated formats like PNG or JPG. In plots, make sure the axes are labelled, units are given, and line-styles are distinct and defined. The best positioning of figures is usually "floating" at the top or bottom of a page, or sometimes several floating figures gathered on a dedicated page with no text. Placing figures in-between text paragraphs is sometimes appropriate, but it is a rarity and usually breaks the reader's flow, sometimes stranding a lone sentence at the start or end of a page, or between large images.
- …

A long tl;dr: the “top ~10”

- 6. Tables: First, ask yourself if a table could be better as a graph. A big list of numbers is not easy to look at: keep it short, and make sure the meaning is clear, and the precision available via written-out numbers is essential. Speaking of precision, make sure numbers are quoted to the lowest necessary precision for understanding, and never quote numbers more precisely than their uncertainties permit. Align numbers and text for ease of comprehension: often left-aligned text and right-aligned numbers work best, with centre-alignment rarely the best choice. If feeling fancy, siunitx can align numerical columns around the decimal point. The placement issues described above for figures also apply to tables, and if floating rather than inline, they should have a caption.
- 7. Attribution: Any key insight or factual statement that is not your own and which you don't want to explain in full detail should be attributed to a reliable source. The usual way to do so is with a compact citation marker in the text, matching an entry in a detailed reference list at the end of the document. It's a good idea to use an automatic citation-handling tool, to make sure your references are consistent, ordered correctly (usually numerical markers in increasing order as they are read), and the reference list well-formatted and complete. Detailed rules about boldening journal numbers, italicising titles, etc. are not important, just make sure the formatting is clear, consistent, and unambiguous.

A long tl;dr: the “top ~10”

- 8. Avoid bad breaks: A bad break is where bits of the document that should be kept closely together are split, e.g. over a line-break, a page-break, or even separated by multiple pages. Line-break separation is bad for e.g. values with units, where the value is at the end of one line and the unit at the start of the next, or referring to Chapter 2, but a line-break occurs between Chapter and 2. Most editors, including Word and LaTeX, allow you to use a non-breaking space in these contexts (a ~ in LaTeX, or use the siunitx package; a Ctrl-space in Word). Page breaks are a bad idea between e.g. a section heading and its following content: LaTeX will almost never do this, in Word you may need explicit action, but step one is to use the explicit title vs paragraph styles rather than just making paragraph text bigger and bolder for titles. Multi-page breaks are a potential problem for figures and tables that get separated from "their" text: as things move around while writing, fix this in your final proof-read, to make sure display elements are conveniently visible from the relevant (and referencing) piece of the text.

- - 9. Referring to document elements: capitalise proper nouns (i.e. names) like Section 1.1, Figure 3, etc. when referring directly to them. Citation markers are usually placed "silently" after the sentence they are backing up, without needing an explicit piece of referring text unless you want to discuss that source, e.g. with "as seen in Reference [12]...".

A long tl;dr: the “top ~10”

- 10. Enjoy writing, and start early: Getting to explain and advocate for something you're immersed in is a great experience... once you get started. Enjoy the freedom and challenge of finding the best explanation and phrasing you can, to show the world your own personal picture of how things work. But until you start putting words on the page, it can seem intimidating and hard to know where to begin. Break the task down into small parts, and start with whichever currently appeals most... and be prepared to review and rewrite tomorrow. Scheduling a regular writing slot can also help. Good luck!

NB. The last few points are quite fine-detail rather than general, but I wanted ten!