Ensuring quality PhDs in computing

Ian Sanders[a] ⓘD sandersi@acm.org

[a] School of Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa

1 CONTEXT

Since completing my PhD in 2002, I have supervised PhD students and examined the theses of PhD candidates from other institutions. I have also talked with many colleagues and students about PhD studies and supervision. In recent times, I have become increasingly concerned that many PhD candidates seem not to understand what is required to complete a PhD or do not put in enough effort to do a decent PhD. This results in many "revise and resubmit" and even some "reject" recommendations from external examiners.

In the rest of this article I expand on what I believe constitutes an acceptable PhD (Section 2) and I give some examples of the failings I have noticed in theses I have been asked to examine or in my students' drafts (Section 3)[1]. In Section 4 I describe some of the things I do to ensure that my students understand what is required of a PhD thesis and don't fall into any of the "traps" described in Section 3. New supervisors could adopt these ideas in dealing with their own students.

2 WHAT CONSTITUTES AN ACCEPTABLE PHD?

There are currently numerous forms of PhD that a candidate could complete. Gannon (2006, p. 1061) says "a PhD has as many meanings as there are educational systems. It is not – and has never been – a single, well-defined qualification". In this article I will only talk about the *traditional* PhD – as completed by Gannon and myself – as this is the typical form of PhD expected by most South African universities (see for example WITS (2021)). Such a PhD involves investigating a problem (or sometimes related problems) in the student's area of research. The thesis should describe what was known about the problem previously, what research has been done on the problem by other researchers, what was done towards solving the problem, what

[1]Note that while I focus on computing and, in particular, computer science because that is my "academic home", I believe that some of what I present is more widely applicable.

the results mean and how these results contribute to the field, and where or how further progress in the field can be made (Wolfe, 1996). In such a thesis, candidates are expected to show that they can do good research, which includes making an original contribution and arguing that they have done so. Candidates also have to produce an acceptable document. These requirements are expanded on in various university documents (see for example UKZN (2020), UNISA (2018), UWC (2019) and WITS (2021)).

The requirement of doing good research can be broken up as below.

Finding a good problem to study

Sometimes candidates start with a somewhat focused research problem but often the problem to be studied is not that clear. In both cases, candidates will need to do lots of reading to find a focused problem statement and from that develop research questions or aims. Candidates must clearly identify the problem and argue why it is worthwhile investigating. If it is clear why the problem is worth studying and how studying it will make a contribution to the field, then the candidate will start off with a better chance of doing good research.

Reading and understanding the literature in the area

Clearly this is related to the problem of finding a "research gap" (a problem worth studying, as above), but it is also important for candidates to put their own work into context with what has been done, and not done, by other researchers.

The literature review has to be broad enough to cover all of the related research, but it also has to be at sufficient depth to convince the readers (or examiners) that the candidate really understands the theories or methods in the area. This includes discussing seminal papers in depth and then tracing the progression or development of ideas or concepts over time. If any theory, method or approach is to be used in the candidate's research, then details of these must be presented in sufficient depth to convince the examiner that the candidate understands the background, use and limitations of these.

Finding and applying an appropriate approach for conducting the research

Any PhD candidate should know that there are different approaches (research methodologies) that can be taken in executing research. Sanders et al. (2022) argue that an explicit statement of the candidate's research methodology can make for better research, including research in the more theoretical computing topics.

In any thesis it is incumbent on candidates to explain to the reader or examiner what they did in their research and why they did it that way. Candidates should justify the choices they have made.

Actually doing the research

Clearly this is dependent on the type of research being done, but it is important that the candidate is aware of and can explain the implications of any choices that were made.

Thinking critically about what was done

Once the actual research project has been completed, the results of the work must be evaluated. This should be done in terms of the actual results and how these results could have been affected by the decisions made earlier. The results should also be evaluated and discussed in relation to the research questions. Finally candidates should show that they understand the relationship between their results and the research done by other researchers in the field.

Arguing that a significant contribution has been made

Here candidates need to argue about the worth of their research. Is it new/novel? How? Why? How does what they have done fill the gap in the field? Will this be of interest to other researchers in the area?

All of the above needs to be "packaged" in the form of an acceptable document. The requirements for this document are as below.

Correct grammar and spelling

Most word processing packages include tools to check grammar and spelling but even if these are used the final thesis should be carefully proofread by the candidate and the supervisor[s] and preferably by at least one other person. Professional language editing should also be considered.

Consistent layout

The "look" of the document should be the same throughout. This means the same font, line spacing, paragraph breaks, and chapter and section headings. All the figures should be produced using the same tool – cutting and pasting from various sources looks sloppy. The same applies for all tables.

Clear figures

All figures must be big enough that all annotations and other text are easily readable. The figures themselves should be crisp and clear – drawn at a high enough resolution. If colour is used, then care should be taken that grey scale printing is still clear.

Clear tables

All tables should be big enough to be clearly read. Tables should not be split across pages. If this is necessary, then they should be explicitly labelled as continuations.

Accurate referencing

Often a referencing style is mandated or suggested by an institution. Sometimes this is not the case and candidates can select a style. Either way, it is the candidates' responsibility to apply the style used accurately. This means including all of the required information for each type of source and making sure that the formatting of each item is correct.

A coherent logical flow

It is important that candidates lead their readers (or examiners) through the document. This means that the flow of the document should be coherent – the title, abstract, introduction, body and conclusion should be about the same thing.

3   CURRENT SITUATION

Common problems that I see in PhD theses in recent times are listed below. Note that I cannot give specific examples as the examination process is confidential.

Lack of alignment

The title of the thesis, the introduction, the problem statement, the research questions, the methodology and the discussion differ in the details of what the research is about.

Superficial literature reviews

The material presented in the literature review does not go into any depth on any topic. The literature review reads like a collection of "cut-and-paste" snippets from abstracts of various papers.

Methods, theories and algorithms that are relied on in the research are not described in any depth. It is very difficult for the reader to determine whether the candidate actually understands these.

There is also very little synthesis of ideas and concepts presented in the reviewed literature.

Lack of knowledge of research methodologies

As mentioned above, and in Sanders et al. (2022), I believe that all PhD candidates should know about research methodologies and be able to argue that they have selected an appropriate methodology for their research. I have seen theses where there is no discussion of possible approaches that could be taken to answer the research questions or achieve the aims of the research.

Unclear description of how the research was actually done

Most theses have chapters that describe what the candidates have done. The problem is that in many cases the candidates do not describe why they have done the research that way – choices are not justified.

Lack of novelty

I have seen a number of theses where the candidates have completed a few small projects. Sometimes these projects are related and described as such by the candidates. Sometimes they are not. The problem here is often that each sub-project is something

that could be done at the Honours (or possibly MSc) level to demonstrate that the candidate understands the process of doing research. Unfortunately, two or three or four Honours projects do not make a PhD.

This problem is sometimes compounded by the fact that the candidate has a paper (or more than one paper) about some aspect of the research published in a little-known journal or conference proceedings. The candidate seems to feel that such a publication means that the research has been shown to make a significant novel contribution. Even if this is the case, the candidate must still argue this in the thesis.

Lack of critical thought
This point relates to the two preceding points – the candidates often do not really think about (or, at least, do not show in their documents that they think about) the choices they make and how such choices add strength to their argument (or detract from their argument).

Poor grammar and spelling
Some theses are never carefully proofread (until the examiner gets to do so).

Poor referencing
It seems that many candidates do the bare minimum in terms of referencing. They use some reference management tool but never check that they input the correct and complete information into the tool or check the output from the tool.

4   THINGS TO DO IN AN ATTEMPT TO ENSURE QUALITY

I believe that it is incumbent on us, as supervisors (and examiners) of PhDs, to maintain standards. I know that some students put pressure on their supervisors to be allowed to submit even when their theses are not ready, but we should try not be swayed by these students. We also should not succumb to pressure from our own (and other institutions) to lower standards so as to increase pass rates.

Some of the things that I, as a supervisor, try to do to maintain standards are listed below.

- I explain to my students what our respective roles of supervisor and student entail.

- I ask students to read at least one good PhD in a related area. To enable this, I keep a list of PhD theses and add to that as I come across new good theses.

- I point my students to available resources (see for example Creswell (2014), Mouton (2001), Oates (2006), O'Leary (2017), Olivier (1999), Pilkington and Pretorius (2015) and Walliman (2001)).

- I provide my students with lots of supplementary material about PhD studies and thesis writing.

- I make sure that my students are aware of what makes an acceptable PhD (see Section 2).

- I read drafts of my students' work and try to give detailed feedback at the conceptual and technical levels.

- I encourage my students to submit papers to conferences and journals. Even if the papers are not accepted, the feedback from reviewers can be used to help to strengthen the student's work. If the papers are accepted, then this provides some validation for the student's work.

  I do, however, point out to my students that publication of papers does not guarantee that they will be awarded a PhD. They must still be able to argue that they have made a *significant novel contribution* with their PhD research.

- I try to have the student's final draft read by a colleague (or two) so that there is an extra level of quality control.

- I also try to get my students to do presentations of their research at postgraduate workshops or other venues (if they do not get to present at conferences).

- I try to resist students' requests to submit until I believe that they are ready to do so.

  If they do eventually browbeat me into allowing them to submit before I feel they are ready, then I explicitly, and in detail, tell them what I believe to be the weaknesses of their work.

- I try not to be swayed by various pressures to recommend a pass for a student if I believe the work is not up to scratch.

We, the supervisors, are the gatekeepers of quality! If I am correct in my perception that many students do not know what is required of them to get a PhD, then it is up to us to help them and to ensure that quality PhDs result from their efforts.

ACKNOWLEDGEMENTS

References

Creswell, J. W. (2014). *Reseach design: Qualitative, quantitative, and mixed methods approaches* (4th). SAGE.

Gannon, F. (2006). What is a PhD? *EMBO reports, 7*(11), 1061–1061. https://doi.org/10.10 38/sj.embor.7400842

Mouton, J. (2001). *How to succeed in your master's and doctoral studies: A South African guide and resource book*. Van Schaik.

Oates, B. J. (2006). *Researching information systems and computing*. SAGE.

O'Leary, Z. (2017). *The essential guide to doing your research project* (3rd). SAGE.

Olivier, M. S. (1999). *Information systems research: A practical guide*. Published by author.

Pilkington, C. & Pretorius, L. (2015). A conceptual model of the research methodology domain. *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015), Volume 2: KEOD*, 96–107. https://doi.org/10.5220/0005613100960107

Sanders, I., Pilkington, C. & Pretorius, L. (2022). Making research methodologies in theoretical computing explicit. *South African Computer Journal, 34*(1), page numbers TBC. DOI-TBC

UKZN. (2020). Information for the guidance of examiners of higher degrees (Form EX 1-8) [University of KwaZulu-Natal].

UNISA. (2018). Procedures for Master's and Doctoral Degrees [University of South Africa].

UWC. (2019). Information and guidelines for examiners of Master's and Doctoral theses [University of the Western Cape].

Walliman, N. (2001). *Your research project: A step-by-step guide for the first-time researcher*. SAGE.

WITS. (2021). Faculty Standing Orders for the Degrees of MSc and PhD [Faculty of Science, University of the Witwatersrand, Johannesburg].

Wolfe, J. (1996). How to write a PhD thesis [Accessed 24 March 2022]. http://www.phys.unsw.edu.au/~jw/thesis.html