Leaflet 1: Writing Quantitative Empirical Studies

Merlin Schaeffer
University of Copenhagen

August 26, 2020

> An empirical study is a scientific article with a continuous text flow that makes the following three core contributions: it 1) makes a theoretical argument, 2) proposes how a thorough empirical test of the argument's implications can be performed, and 3) reports the results of this test. The quality of these three contributions define the overall quality of an empirical study. The goal of an empirical study is to convice the reader that the theoretical argument is empirically (in)valid.

To be original, an empirical study must make an innovation to either 1) or 2). That is, an empirical study either makes an innovative argument, or it proposes how an established argument may be put to an innovative empirical test. If you are a MA Student, make sure to meet this criterion of originality. As a BA student, you do not need to come up with own original ideas and tests yet. Your focus should instead lie with a comprehensive 'Literature Review' and a carefully-executed empirical analysis; innovation comes later (at the MA stage). Some empirical studies deviate from this pattern, for instance if they are exploratory or descriptive. As students, you should best not deviate from it, and discuss with me in how far you want to devaite.

Quantitative empirical studies tend to follow a well-established structure that corresponds closely to points 1), 2), and 3): they consist of a 'Theoretical Background' section, a 'Data and Methods' section, and a 'Results' section—always in this order! Published empirical studies oftentimes deviate from this structure in minor ways. But as students, you should try to stick to it closely. You can start to deviate from this well-established structure, once you have mastered it. These three sections are framed by an 'Introduction' and a 'Conclusion'. Thus, any empirical study should consist of five sections:

1. *Introduction* (1 page): The introduction does more than introducing the reader to the topic. The introduction is like a trailer for a movie: it is a preview of the whole study that shall raise interest. Oftentimes it already spoils the movie's best scenes (i.e., paper's most interesting arguments and findings). The first paragraph introduces the topic and explains the societal relevance of it. It needs to draw the attention of the reader and convince her/him that the topic is important. The second paragraph gives a very brief summary of the existing literature on the topic (see 'Literature Review' below). The third paragraph introduces the contribution of your empirical study. It is usually a lucid summary of your argument (see 'Own Argument' below). If you do not make an own argument it is a lucid explanation of your empirical test and why it matters. The fourth paragraph is a summary of the empirical analysis (i.e., which data and which methods were used). The final paragraph is a matter of taste (you can also skip it): it is a brief summary of the core results. As an example, check out the intro of Schaeffer et al. (2016).

2. *Theoretical Background* (2-3 pages): In this section you present your argument. The section should always start with a *literature review*, which gives an overview of the existing literature in the field. That is, give a review of the existing debate on the phenomenon you are interested in. This entails established theories and empirical evidence. Summarize the different theories and empirical findings in a way, so that people who have not read a single text from that field of study can follow the review. Remember, this qualification earns you a job later on. Angela Merkel does not have the time to read into the details herself, she relies on your brief literature reviews exclusively!

   Be advised that a good review of the literature is always much more than an additive summary of theories and empirical findings. A good review sorts/orders them along important dimensions. It highlights similarities and differences and therefore entails many cross-references to its various sections (i.e., 'As explained before', 'similar to Bourdieu' ...). To highlight the sub-dimensions, one may also introduce a third level of sub headings. Note that some discuss theoretical arguments and empirical evidence in separate (sub-)sections; I do not want you to do that, as it interrupts a nice text flow.

   Apart from ordering the debate, the overall purpose of the literature review is to identify a *gap*. That gap may either be:

   (a) *Missing empirical evidence*: Many studies put existing arguments to an original empirical test. In this case, the theory section consists of the literature review, derives hypotheses from existing theories and mentions that empirical evidence for these hypotheses is missing (at least with respect to a specific country, time, or social group). Also, derive the hypotheses directly in the running text and not in a separate hypotheses section. This helps to keep a nice text flow.

   (b) *Theoretical/explanatory gap*: Alternatively, there may also be a theoretical gap. Classical types of theory gaps are: a) The different theoretical arguments and the various empirical findings have not been synthesized; doing so allows us to formulate new hypotheses. b) An aspect of the phenomenon remains unaccounted for, we need a new/and additional explanation. Once you have identified your theory gap, close it in an own sub-section (so as to highlight that you make an own theoretical contribution).

   According to King et al. (1994), a "social science theory is a reasoned and precise speculation about the answer to a research question, including a statement about why the proposed answer is correct" (King et al., 1994, p. 19). This is precisely the task, you are asked for in this sub-section: give reasoned speculation about the answer to a research question. Use the roughly two to three pages of this sub-section to develop a convincing argument why your speculation is reasonable, meaning logically sound. This means, if you choose to tackle a theoretical gap, you cannot give a summary of theories and derive hypotheses from these. Rather, use existing theories and empirical findings as building blocks of your argument in favor of your reasoned and precise speculation (remember the King et al. (1994) quote above). Try to imagine a reader who learns about your speculation, and disagrees. The purpose of this sub-section is to convince your (imagined) opponent that one can make a reasonable (i.e., logically sound) case in favor of your speculation. That does not mean, your speculation is true or that people need to believe it is true after heaving read your theory section. It only means that you need to convince people that it *could be true*. As an example, check out the theory section of Hopkins (2010).

3. *Data and Methods* (2 pages): In this section you explain how you will design a thorough test of the raised hypotheses. Emphasize if your empirical test is innovative. Is it, for instance, the first longitudinal or experimental test of the argument? As an example, check out the research design sections of Dinesen and Sønderskov (2015).

   As to the structure of the 'Data and Methods' section, the first paragraph introduces the data that have been analyzed for the study. The next paragraph details the sample. After that, four sub-sections follow. Sometimes people re-order the sub-sections, Dinesen and Sønderskov (2015) is an example. Anyways, within each sub-section, make sure to discuss potential limitations and whether your procedure is established, falls short from, or improves over other studies.

   (a) *Dependent Variable(s)*: Explain the operationalization of the dependent variable(s) analyzed in the study. For instance, did you use factor analysis? How well did the single items score, what was the wording of the survey questions, etc.

   (b) *Predictor Variable(s)*: Explain the operationalization of the explanatory variable(s) used in the study. Explain why those explanatory variables help us to identify the relations and processes suggested in the theory section.

   (c) *Control Variables*: Explain the operationalization of the control variable(s) used in the study. Also explain why you think these variables should be controlled for in the study.

   (d) *Methods*: Here you explain how you put your dependent variable in relation to the explanatory variables. Why is the method that you have chosen superior to other methods? In some rare and excellent cases, students find the time to also use alternative approaches to cross-validate their results as robustness check. Most of the time you will simply mention that you use some type of regression and briefly explain how it works. How did you handle missing values? Do you estimate separate regressions per group? Will your argument be tested as an interaction effect or as a squared term?

4. *Results* (2-3 pages): A good way to start your results section is by briefly summarizing in how far your results re-confirm other (i.e., beyond your concrete argument) well-established arguments and findings that you have reviewed in your 'Literature Review'. Repeat the arguments briefly and report the evidence. This cross-validates how convincing your empirics are. As an example of this, see the initial paragraphs of the 'Results' section of Schaeffer (2017).

   Then continue to repeat your central claim, and report your results in answer to it. In doing so, do not only mention significance levels. Instead, try to give substantial interpretation of the regression coefficients. Compare the effect size to other predictors in the model. Try to plot results if possible and discuss the figures. But do not discuss results that do not directly concern your claims. For instance, do not discuss R^2 or F-tests, if these do not support of contradict your hypotheses.

   If there are plausible alternative explanations, consider to have a 'Robustness Checks' sub-section in which you try to rule out alternative explanations. In this section you can also discuss further (for instance more rigid) empirical tests or tests among sub-populations. As an example, check out the results section of Hopkins (2010), although be warned that it is very extensive.

5. *Conclusion* (1-2 pages): The first paragraph of the conclusion repeats the central research question and again summarizes your argument in answer to this question.

The next paragraph gives a brief summary of your key results. The conclusion should then also entail one or two paragraphs on limitations of your study and directions for future research. What are the open questions and remaining problems? However, do not end with limitations and problems. The conclusion's final paragraph should re-address introduction's first paragraph to close the circle and thereby also re-emphasize the relevance of your empirical study. As examples check out the 'Conclusion' sections of Dinesen and Sønderskov (2015) and Schaeffer (2017).

General suggestions

- *Imitate*: Read this leaflet in reflection of concrete empirical studies that you read in preparation of your seminars. This will help you to understand the abstract points I am trying to make. More importantly, do not try to structure and write your empirical study from scratch! Instead, take a close look at published empirical studies and try to imitate how the experts do it. If you are unsure what belongs into a 'Results' section, imitate the style of two or three results sections that you really enjoyed reading. Also, take a look at the work of your teacher to get an impression how s/he likes to do things.

- *Yes or No?*: It really eases things, if you approach a topic/research question that can be answered with "yes" or "no", or could at least be formulated as a yes/no question.[1] This is important, because it establishes a standard against which you can compare the quality of your empirical study, that is, whether it gives a convincing answer to the yes/no question.

- *Text flow*: Although the structure is predefined and the matter seems very technical, try two write an enjoyable text with nice text flow. Despite the fact that you write an empirical study, try to take the reader on a journey that shall convince her/him of your contribution. Try to captivate your reader in the very first paragraph and then hold the grip by guiding her/him from paragraph to paragraph.

- *Frequent summaries and announcements*: You should start (sub-)sections by announcing/summarizing what follows and how the section is structured. This helps readers not to get lost and to know what is ahead. Also try to give brief summaries at the end of sections, so as to wrap up and ensure that readers have gotten the most important take home messages from each (sub-)section.

- *Levels*: Do not use more than three levels of (sub-)headings. There should only be sections (e.g., 'Theoretical Background'), sub-sections (e.g., 'Dependent Variables'), and the final level of sub-sub-sections.

- *Paragraphs*: Apart from ((sub-)sub-)sections, paragraphs are your most important tool to structure your text! You should thus use them in a very explicit and guided way. In general, the following holds: one idea, one paragraph. Every time you start a new idea, make a paragraph. At best, start your paragraph with that idea. Then use the rest of the paragraph to explain and justify that idea.

- *Finish early*: Be done *at least* two days before deadline. Print your paper. Let it rest for at least a day. After that, Proof read it and be astonished how much there is to revise. Iterate until you are confident with the result.

[1] Take for instance Schaeffer et al. (2016): "Do School Dropouts of Turkish Origin Earn More in Germany?", Dinesen and Sønderskov (2015): "Does Ethnic Diversity Erode Generalized Trust?".

- *Reading club*: Have a circle of fellow students and circulate your papers before you hand in.
- *Zotero*: Check out `www.zotero.org` and use it together with the word plug-inn, unless you are already using some other program that manages your literature.
- *Spelling*: Continuously bad spelling really puts off any reader, irrespective of whether your ideas are great.

Necessary precondition There is a minimum level of coherent, logical and consecutive writing, which any reader of your work expects. Also be aware of the difference between academic claims making versus normative/political claims making. *If your text falls short of these minimum requirements, I will not give you feedback and might even rule the term paper as 'failed'!* If this happens to you, but the text represents your best effort, feel strongly urged to visit a class on academic writing. If you could have done better, you really need to show more effort.

Final consideration If you do not find the time, to write a paper that lives up to your own expectations about how a good paper should look like, why would you expect me to give you detailed feedback?

References

Dinesen, Peter Thisted and Kim Mannemar Sønderskov, "Ethnic Diversity and Social Trust Evidence from the Micro-Context," *American Sociological Review*, 2015, *80* (3), 550–573.

Hopkins, Daniel J, "Politicized Places: Explaining Where and When Immigrants Provoke Local Opposition," *American Political Science Review*, 2010, *104* (01), 40–60.

King, Gary, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton: Princeton University Press, 1994.

Schaeffer, Merlin, "Unfulfilled Status Aspirations and Perceived Discrimination," *SocArXiv/Open Science Framework*, 2017, p. https://osf.io/vykkz/.

_ , Jutta Höhne, and Celine Teney, "Income Advantages of Poorly Qualified Immigrant Minorities: Why School Dropouts of Turkish Origin Earn More in Germany," *European Sociological Review*, 2016, *32* (1), 93–107.