

PhD in Development Studies, University of Lisbon

Research Seminars I and II (RS I and II)

**Notes about the Preparation of the Research Project**

Carlos Nuno Castel-Branco

(2022)

**Challenges**

By the end of May, first year PhD students will have to submit a research project (which we may also call your research argument), which will broadly guide your research activity until the end of your PhD. The approval of this project is a pre-condition for the continuation of your research work at PhD level.

The standard practice adopted in previous years involves two steps: the submission of the research project to your tutor, to the RS II lecturer and to whoever else is going to discuss and examine your proposal; and the *viva,* or presentation and evaluation of your proposal in a workshop.

The goal of these notes is to help you understand the different stages and tasks you need to go through to produce a good research proposal.

**What is expected of you?**

You must produce a research project (or research proposal, or research argument), of a maximum length of 10,000 words (between 20-25 pages, depending on the format and graphic content). The most important goal of this document, a long essay, is to make plain clear what the project is about (the problem), the field of academic work and literature your project is part of (literature research), why you intend to do it (its intellectual relevance), and how you intend to do it (the theoretical approach, the road map and the methodology). Think of your research proposal as the research argument you are presenting, this is, your argument for doing this piece of research work through your PhD.

The research project needs to be built around a research problem and hypotheses (or research road map), a clear analytical framework and a detailed research methodology. All these four elements are developed, at the initial stage, from the literature research that is undertaken. Therefore, your research argument must include: a research problem (which also explains why this is relevant to study – for example, because of a gap in the literature or an unexplained paradox or problem); research hypotheses, which are the road map for your research; a solid analytical framework (the theoretical field of work you are going to utilise); a methodology, which explains how the research is going to be conducted; and the required literature research that supports the problem, the hypotheses, the analytical framework and the methodology. In other words, the researcher needs to explain how the research argument is developed, taking into account the specific historical and social conditions in which the research problem is located and the existing state of the art in the relevant field of study, by extracting from the literature – scientific, journal articles and books, statistic sources and grey literature (media, reports, mimeos, briefings, policy documents, etc.) – what is essential to identify the *“thesis-shaped hole”* and how this “hole” is going to be addressed/filled/unravelled.

***Research problem – asking relevant questions about something interesting that needs to be unravelled***

First, you need a “problem” – what is the problem you want to address with your research? Your problem typically has two levels: (i) a particular social problem, which is socially relevant; and (ii) a problem with the literature, which is intellectually relevant. Both are important. If the social problem does not exist, or if there are no problems (gaps, errors, etc.) with the literature, you may not have anything to investigate and may need to choose another topic.

A research problem is the definition of what the student wants to unravel with his/her research, which also includes locating the problem in its specific socioeconomic and historical context and in the literature. There is no point in researching what we already know, we only research what we do not know. Suppose that you say “my research problem is that agriculture labour productivity is low”. This is not a research problem, but a problem with agriculture labour productivity. Additionally, if you already know that agriculture labour productivity is low there is no point of researching it. However, you may be interested in a particular question related to low agriculture labour productive, which for example, may explain it or enable you to measure it properly (if the problem is measurement), and which is under-researched or not properly researched. You identify a problem for your research related to your original concerned (the problem must be socially and intellectually relevant, researchable, original and properly formulated – this is, with the adequate theoretical and methodological tools.

Let’s take a more detailed example. Suppose that the student is concerned with how large mining companies change the local socioeconomic context and the relationship between the state and citizens, in the context of decentralisation of the state (empowering of municipalities, for example). You may have heard from family or friends, or from media reports, or observed directly that difficulties arise with local political empowerment when there is a very large private corporation operating in the district or municipality. This may enthuse your curiosity and you may want to understand what the problem is. It is not possible to study everything related to this issue in a single piece of research, as otherwise it would never finish, or it would be too superficial and of limited social, intellectual and academic value. Additionally, you may not know what that “everything related to this issue” is – you may have heard of displacement and relocation of people, expropriation of land, deforestation, employment being created or displaced, and so on and so forth, but you still do not have a common thread that connects the main dots and generate a logical and credible narrative. In other words, you do not know what the problem is, although you have access to anecdotal and loose information about possible manifestations of the problem. So, you need to ask a question, your research problem, that can be answered through research and that guides your research.

The research problem cannot be stated as a general concern or statement like, for example, “decentralization is affected by large mining companies” or “large mining companies dominate local politics”. We already know that, and this type of statement would be interesting only if it was not true – if large mining companies did not affect local political, economic and social life then it would be interesting (for research, things get interesting only if we do not know anything about them, or if we cannot explain them, or if they do not happen as expected – if they are paradoxes, real or apparent).

Imagine two friends, A and B, talking, at the end of the day, about how their day was. A says to B *“something interesting happened to me today, I walked in the rain and got wet”*. What do you expect B to reply? If you were B, you would probably say *“so what, what’s interesting about it? You walk in the rain, you get wet”* (it seems that the only interesting thing about A’s story is the motives: why did A walk in the rain if he/she was going to get wet? Or why is A telling B this boring story? So, you would question A’s motives, not the outcome of walking in the rain – the research interest would be on the motives, not on the story *per se*)*.* Suppose that, instead, B says *“how interesting, tell me more about that”.* Then, you would start asking questions about B’s reaction – is he/she kidding? Being ironic/satiric? Trying to flirt with A? Is there more to A’s story than what it seems? Why do you ask these questions, why do you want to understand B’s reaction? Because it was not expected, right? You cannot explain it, you do not know, so it becomes interesting for research.

Now, suppose A says to B *“something interesting happened to me today, I walked in the rain and did not get wet”*, and B replies *“how interesting, tell me more about that”.* Now, B’s answer is the expected one, so it is not interesting *per se* (you may still think that B has no genuine interest in the story, but that he/she’s trying to allow small talk to develop or to flirt with A, but B’s answer is not a “research problem”, in this context, unless you want to unravel B’s real motives to be interested). The fact is that walking in the rain without getting wet is an interesting topic, that raises many questions. Was A under an umbrella or wearing a raincoat? If this is the answer, then the research interest dies there – of course, the chances are that if you are using an umbrella or wearing a raincoat you are not going to get wet in the rain. You still question A’s motivations to tell B a boring story with a very simple explanation.

However, suppose A was not wearing a raincoat or using an umbrella, and walked in the rain and did not get wet – this would make things interesting. B can kill the story, assuming that A had a raincoat and an umbrella. So, B does not know if this is true, but assumed this to be true, so he killed his/her curiosity and never got to the bottom of the story. However, B is a good researcher, so instead of assuming he/she asked questions: *“how interesting, tell me more about that”,* which is an open question that allows A to explain, but which is constrained by what A wants to say, or by what A remembers about the facts, or by A’s interpretation of the facts. So, B can strengthen his/her research inquiry but adding something else to his/her original reply, that helps him/her getting concrete information from A: *“how interesting, tell me more about that. Were you wearing a raincoat and/or using an umbrella? Did the wind pull the rain away from you? How did this happen?”.* By adding small, concrete questions to the open question, B is gong to get concrete information (yes or no) in addition to whatever else A wants to say, and can, thus, decide to pursue or stop the research (notice that the concrete guiding questions, about the raincoat and the umbrella and the wind were informed by assumptions, which in turn were informed by experience and theory. However, the researcher did not stop at assumptions, he/she asked the questions thus obtaining information to stop the research because all had been unravelled – if the answer to raincoat or the umbrella was yes – or to continue the research if the answer to both raincoat and umbrella was no).

To wrap up this conversation, the research problem is a relevant question that arises from an interesting concern of the researcher that cannot be explained without further research. The researcher needs to find the research interest. Walking in the rain and getting wet is not interesting, *per se,* but the motivation to tell that story or to walk in the rain, or the consequences of getting wet by walking the rain are interesting areas of research, because the researcher cannot have an answer to a question that is relevant without further research. Walking in the rain without getting wet is potentially interesting, but before getting too excited about it, the researcher should check if that potential is real – for example, you can solve the problem immediately and kill the research interest by asking the most obvious and simple questions about the raincoat and the umbrella, if the answer is yes to either; or continue your research, if the answer is no to both.

This is what a research problem is, the formulation of a real and relevant issue as a real and relevant question, that can be investigated and needs to be investigated

We may start with a concern (such as “what happens when political, administrative and financial decentralization happen at the same time as the establishment of a large mining or a large plantation company”), but then we need to formulate the problem (why this may be a problem, and how we can research it?). So, the student needs to decide what to focus on, bearing in mind that he/she can only research what is unknown, what is problematic about the socioeconomic issue at stake.

At this point in time, in your PhD, most of your research will be done with the literature related to your research problem → the concern leads to literature → literature leads to the development of the problem → the problem leads to more sophisticated, specialised, selective and more focused literature → which leads to the thesis-shaped hole that is going to be the focus of the research → which in turn leads to even more focused literature, a better understanding of the problem, the development of the theoretical foundation of the research, the formulation of the road map/hypotheses, the definition of the methodology, and so on.

The student can build the research problem as follow:

“One of the main debates about the relationship between the State and citizens, in Mozambique, is that of decentralization. This is on everyone's lips. However, **what does “decentralization” mean in a society dominated by the economic, social and political dynamics emerging from mega-projects of an extractive nature?**

*[Note: The sentence in bold is the research problem, broadly defined, as a short question. Research problems usually are questions].*

“The experience of decentralization in Mozambique is very uneven and stumbling, there is a multitude of arguments about decentralization, some more democratic (associated with more or less liberal concepts of citizen participation), others truly anti-democratic (associated with the mere creation of space for sharing of power between political groups with military capacity), but the debate on the subject has been limited to a normative justification of the importance of decentralization (why it is good or should be good) and the “political” conceptualization of the process, regardless of the real socio-economic basis in which the State's power and its broader relationship with citizens, at all levels, is structured. Although the influence on the State that comes from interests, tensions and conflicts around the organization of the economy is widely recognized, this debate about decentralization occurs as if the social and economic structures of reproduction and accumulation do not exist.

*[Note: the above paragraph identifies the specific socio-economic problems with decentralization (uneven, stumbling, supported by democratic and anti-democratic perspectives, occurring within an economy strongly structured by large, multinational corporations); AND with the literature (…the debate occurs as if the social and economic structures of reproduction and accumulation do not exist)].*

“Thus, the starting point for this research is to investigate what, in practice, decentralization means in a context dominated by an extractive multinational, analyzing the experience of the district of Moma and its dominantly organized economic, social and political life around the Kenmare multinational, which exploits its heavy sands.”

*[Note: this paragraph confirms what the research problem is and narrows the scope of the research to one specific district. Of course, there are other options, like, for example, focusing in two different districts, in different regions with different types of extractive projects, such that a comparison can be made about different variables – do different types of extractive industries matter? Do different districts react differently, such that history and local institutional capabilities matter? And so on. In any case, narrowing down the scope of the research matters and needs to be done. So, the student may choose one case study, like Moma, to identify in detail how the process and operation of decentralization under mega project pressure occurs in one case, or he/she may have more than one case study, to control and compare for other variables. It is the student’s choice, but he/she needs to explain what the case study or the comparison of case studies intend to achieve. Of course, the student’s research may not involve case studies, in which case he/she needs not narrow down the scope of the research around cases studies, but around something else –like a period of time, a limited set of issues, and so on.]*

“The study will be better developed within a heterodox and interdisciplinary political economy framework, which involves the description and understanding of the essential characteristics of the agents involved, the structures of accumulation, and the relationships between them. This requires the understanding of the local, national, regional and international historical and social contexts under which the agents, structures and their relationship were built and developed.

*[Note: This paragraph briefly describes the type of social theory and/or approach that the research will adopt. The justification for this, the details of this, and how and why these matter and will be implemented, can be developed in a separate section of the project proposal.]*

In four short paragraphs the research problem has been identified and narrowed to a researchable and realistic set of issues, and tools, and the analytical framework has been introduced.

As mentioned, this research problem has two levels, namely: what is happening with decentralisation and the pressures this process is under, and how the literature discusses it. By addressing one problem, the research also addresses the other problem.

The original contribution of the thesis comes from its critical contribution to the literature (the gaps in the literature having been identified), and how this contribution changes the way the socioeconomic problem (decentralization under multinational influence) is researched and understood. From this point, it is possible to select which particular aspects the thesis is going to focus on (for example, how the presence of one multinational influences and changes the dynamics and focus of municipal planning, how demographics change with tensions related to land expropriation, and so on). The hypotheses are going to be derived from the problem and the set of questions it defined.

***The theoretical framework***

One cannot proceed to do any systematic and rigorous research with a rigorous theoretical framework. A theoretical framework is not a set of opinions and ideas taken seriously only because they are fashionable, of our liking and/or published in a book. The theoretical framework needs to provide rigorous questions, protocols and tools for and be able to guide research and explore the problem that interests the researcher.

For example, in the above-mentioned problem we noticed that *“Although the influence on the State that comes from interests, tensions and conflicts around the organization of the economy is widely recognized, this debate about decentralization occurs as if the social and economic structures of reproduction and accumulation do not exist.”* Hence, we need a theoretical framework that allows us to take into account the social and economic structures of social reproduction and capital accumulation which may determine and explain the direction of decentralisation. To this this, *“The study will be better developed within a heterodox and interdisciplinary political economy framework, which involves the description and understanding of the essential characteristics of the agents involved, the structures of accumulation, and the relationships between them. This requires the understanding of the local, national, regional and international historical and social contexts under which the agents, structures and their relationship were built and developed.”*

Thus, we define the profile of the theoretical framework we need to understand decentralisation in a given historical and social context, based on the relevance it has to clarifying our problem, and we also defined how we are going to deploy it, in conjunction with, or exploring the understanding of the local, national, regional and international historical and social contexts under which the agents, structures and their relationship were built and developed.

***Research Hypotheses or Research Road Map***

Hypotheses are the research road map – they define what the researcher is going to be looking for and how to interpret what he/she finds. Hypotheses that are too obvious are redundant. For example, for the above-mentioned problem, a hypothesis that says *“Large, extractive projects affect socio, economic, political and cultural life”* has no use whatsoever because we all know this. This is of no interest at all (interesting would be if large projects did not affect life, in which case the research problem would change to *“why don’t large, extractive projects affect the socio, economic, political and cultural life?”*). In other words, such a declaration cannot guide research, thus it is not a hypothesis.

Without a clear definition of the problem, researchers cannot have clear definitions of hypotheses. If the hypotheses are the road map to investigate a problem, of course the problem needs to be clearly defined or there will not be any hypotheses that makes sense (we cannot have a road map for an unknown journey, destination and route). Hence, research needs to start with the problem, which includes the destination, the route and the other keys problems that need to be addressed for the research to reach its destination.

Let us go back to the research problem set above: *what does “decentralization” mean in a society dominated by the economic, social, and political dynamics emerging from mega-projects of an extractive nature?* Remember that we want to research this socioeconomic problem because it has not been researched in the way we want to research it (locating the research within the socioeconomic tensions and contradictions of the decentralization process). Remember, also, that we want to study one specific case for the details of the interaction between decentralization and large, extractive projects; or we want to study different cases to compare, contrast and control for specific variables, like the type of the extractive project, local history (for example, related to labor movements and how they affect local disputes) and local institutional capabilities, etc..

Being a road map, the hypotheses narrow the problem and the field of research and, exploring knowledge we already have, identify exactly what we intend to study, some sort of causal relationships and their effects. For example, taking the decentralization example mentioned above, we can formulate a hypothesis in the following terms:

“The influence of the mega project on the displacement and re-location of people and social infrastructure, in association with the centralization of financial resources and of decision-making power over productive strategies and linkages, limits “decentralized” planning to a defensive role of social reparation and reconstruction of the district in the most favorable format for exploring heavy sands, leaving it vulnerable to the megaproject’s business cycle.”[[1]](#footnote-1)

This hypothesis is formulated for the case of Kenmare, in Moma. What is interesting about it? First, it establishes causal relationships, or mechanism by which large, extractive projects influence the nature of decentralization (displacement and re-location of people and infrastructures; centralization of financial resources and of the power of decision-making over productive strategies and linkages, which leaves the local authorities vulnerable to fundamental decisions that fall upon them; mega projects have volatile busyness cycles, which have a local economic and social impact – for example, on jobs or local taxes, or the mega project commitment to social responsibility project). These need to be informed by data or by other means – we need to be able to prove that these (displacement, re-location, centralization of finance and decision-making, business cycle volatility) exist and operate in a particular manner, and affect the local communities in the way we hypothesize.

Second, it establishes effects, or the results of the influence of large projects on the nature of decentralization via the relationships/mechanism mentioned (limits planning to a defensive role of social reparation and reconstruction of the district in the most favorable format for exploring heavy sands; leaves the district vulnerable to the business cycles of large projects). These effects need to be researched – do they happen and how important are they? Are they related to the mechanisms we mentioned?

Third, it tells us what literature we need to research, what data we need to collect, what relationships we need to be able to establish, how to interpret the results.

By sending us into a route of research with clear issues in mind, the hypothesis guides us and allows us to progress in our research. The hypothesis needs to be well informed by knowledge already acquired (for example, through literature research), or it may be only speculative and not a good guide for research.

Suppose, now, that we do not find exactly what the hypothesis predicts, what shall we do? The answer to this question depends very much on the results and how different they are from the hypothesis. In some cases, the research modifies the hypothesis by requiring some qualifications – for example, the prediction of the hypothesis are confirmed if the process of decentralization is contemporary or posterior to the establishment of the large, extractive project; and, in the opposite case, when decentralization was well consolidated before the establishment of the large project, local planning and empowerment had as much an impact on large projects as these had on decentralization. This is a minor change of significant importance – the hypothesis holds true under specific conditions, and the conditions are well identified. Or, the prediction of the hypothesis are not rejected but there is a another strong causal mechanism at work that was not hypothesized: for example, the so called social corporate responsibility, by which multinationals can directly influence social priorities and resource allocation – so, the predictions of the hypothesis are significantly stronger when we account for the direct impact of social corporate responsibility on resource allocation, which is stronger, still, in districts that have very few other sources of public revenue. In other cases, the research modifies the hypothesis in a much more significant manner – for example, the causal mechanisms are not confirmed, or are not confirmed to be related to the effects, or the effects are not confirmed, or there are other causal mechanisms. In this case, we need to explain why the hypothesis is far from the observations and how it needs to be modified.

Thus, the hypotheses guide the research and work as benchmarks and reference points that help to utilize the research results to challenge, modify or completely change the hypotheses. A well informed and well formulated hypothesis guides research because it tells the researcher what to look for, where to look for it, how to interpret findings and how to reformulate the research hypothesis, if necessary. Without the hypotheses, or road map, we would never be able to decide where to go, to understand where we are, and to modify the route, speed, direction if necessary. We would just be lost in the middle of a world of information and without any analytical tools and structures to help us make sense of the information available.

Depending on the aim of the research topic, on the theoretical framework utilized and on the results and findings, researchers may (or may mot) move into the next gear, namely explaining why what happens does indeed happen, why they are aligned with or different from the hypotheses, and what the historical and social explanations for all of that are – in explaining the essence of the social relationships in their own historical context, we reach the realm and core of political economy (please, remember that we mentioned, in our example of a research problem, that this research was going to be guided by a political economy approach, which may not always be the case).

How is a well informed and well formulated hypothesis built? First, it starts with the research problem, which, as already mentioned, sets the questions we need to research, to which the hypothesis works as a road map. Second, it requires a good literature research that provides (i) a theoretical foundation for the formulation of the hypothesis (for example, that political processes and institutions are routed in socioeconomic structures and dynamics of reproduction and accumulation, that displacement and re-location of people and infrastructures have a strong impact on livelihoods, etc.), and (ii) the specific historical and social information we require to identify the components of the hypothesis (we need to know that there is a multinational corporation called Kenmare, extracting mineral sands in Moma, that the establishment of this corporation implied displacement and relocation of people and infrastructures, that finances and decision making over productive strategies and linkages are centralized, that Kenmare has a corporate responsibility program that affects local resource allocation in public utilities and services, etc., etc.).

The hypothesis is not a vague set of opinions or declarations of no consequence for research. It needs to be sufficiently specific and supported by theory and data to guide research. Hypotheses narrow down the problem of research to researchable road maps, specify the causal mechanisms and channels involved, link them with effects and, above all, tell us what to look for, what to read, which data to collect and how to organize, analyze and interpret them, and how to correct the course of research if, as in most cases, it becomes necessary.

***Research Methodology***

Typically, research methodologies include four elements: the hypothesis, which is the road map for research and defines what needs to be researched; the model, which explains the relationships the research is trying to understand, clarify and describe in order to discuss the problem; the research design or research techniques, which explains what the research will be based upon [what kind of research (for example, case studies), how information will be collected, organized, treated, analyzed and interpreted, if field research is conducted, what type and how, and how to guarantee the highest ethical standards]; and the activities that the researcher will have to conduct, which includes, not exclusively, the research methods [further literature research that needs to be done and for what, where the information is going to come from (for example, interviews, focus groups, surveys, grey literature), which chapters are going to be written and at what stages, if field work is conducted, how to organize the logistics of it, the indicative research plan, etc.).

If the research argument starts with a clear problem and a logical hypothesis, the methodology is easier to develop. Therefore, the methodology is not a neutral, technical field of the research – on the contrary, it is strongly influenced by the problem, the hypothesis and the social theories and social research experiences that inform them.

***Literature Research***

Literature research is a complex matter, which captures the complexity and dynamics of research. From the previous points, on the research problem, hypotheses and methodology, we can easily conclude that: (i) research of the literature is necessary to set the problem and formulate the hypotheses; (ii) the research problem and the hypotheses are needed to define what to read and what to research in the literature; (iii) the methodology is informed by the literature in an indirect way (through the problem and hypotheses) and a direct way (through the literature on methodology, methodology learning from the previous experience of research on the specific field or topic of research, etc.); and that (iv) as research develops and knowledges accumulates, research problems, hypotheses and methodologies may change (slightly or more fundamentally), and so new literature will need to be researched. There is no doing away with this complexity and dynamics.

Literature research refers to a critical survey and analysis of published, scholarly sources (such as books, journal articles, and theses) on a topic, complemented, if necessary, by different types of grey literature (statistics, media, reports, mimeos, briefings, policy documents, etc.), undertaken not for its own sake, but to explore critically and professionally a field of knowledge and build the tools and path for specific research. It gives an overview of key findings, concepts, and developments in relation to a research problem or question, provides an assessment, classification, comparison and evaluation.

A good literature research does not just summarize sources and information. On the contrary, it needs a research argument and structure and must feed a research argument and structure. A good literature research aims to:

* Show the state of current knowledge in relation to a central research question or hypothesis in a specific field of study, in which research is going to be undertaken. Researchers need to know the state of the art of the literature in their field of study to full professional level, so that they can discuss and contribute to it in ways the professionals will want to listen to what the researcher have to say.
* Analyse, interpret and critically evaluate the literature, highlight patterns, themes, conflicts, and gaps, and learn, such that the literature research helps to define the research problem, the research questions, the research hypotheses and the research design and methodologies. For example, going back to the research problem and hypothesis exemplified above, the researcher needs to be well informed about the process of decentralization (so that he/she can characterize it); needs to know what and where large projects are and, generally, how they operate and what type of local impact they may have (such that, the connection between large extractive projects and the dynamics of decentralization can be hypothesized); needs to be well aware of the limits of decentralization (limited local empowerment) and of the tensions and challenges posed by the relationship between large extractive projects and local planning and action by elected municipal officials and institutions; if the researcher intends to develop case studies, needs to know the literature about research by case studies and the literature on the history of the cases to study. The researcher also needs to have such a broad knowledge of the literature as to be in a position to identify its gaps (in this example, that the literature does away with the analysis of socioeconomic structures and dynamics under which decentralisation takes place). This means that the researcher needs to know, at professional level, the literature that develops the subject from the point of view of social theory and methodology, as well as literature beyond his/her case studies or country focus. Hence, what the researcher needs to know to formulate a problem and a hypothesis depends on good, critical literature research, which, in turn, depends on what the researcher already know and the hypotheses of research he/she is developing.
* The literature research is key to defining the research design and methodologies. The first three components of a research come straight from the literature research on the relevant field of study, namely: the research problem, the research hypothesis, and the research social theory, which then influence and define the research design and its methodologies. From these, the researcher can derive what he/she needs to know to investigate the research problem, guided by the research hypothesis, and to confirm, adjust or change the hypothesis, using a social theory. From the literature research, the researcher derives the research apparatus – this is, the research design, methodologies and activities – the concepts, the models, the approach, the theoretical demonstration that such an apparatus is adequate and sensitive enough to capture and analyse what is deemed important for the research. Literature research helps to differentiate between essential design [which consists of the definition of how the research hypothesis is going to be investigated – the knowledge and data that are required, how they are going to be obtained, how they are going to be related with predicted outcomes (a model, which needs not be mathematical, necessarily), how the analysis is going to be done and the results interpreted] from the research methodologies, techniques and activities (techniques and activities, though important, are analytically subordinated to theory and to the broad design, but also require sound scientific knowledge – such as, for example, how to design and conduct surveys and interviews).
* The literature research also helps to predict and avoid or minimise problems with research (such as data collection and interpretation, for example), and to deal with ethical issues in research (how to conduct interviews, protect sources, and so on).
* The literature research must help the researcher to build a ***“thesis-shaped hole”***, which the research is intending to fill. The thesis-shaped hole bridges the gap in knowledge of what a well-educated, professional scholar in the field knows and what they need to know to read the remainder of the thesis, what books and articles they would need to read in order to do so. This is not a general discussion but one that points forward to the original contribution of the research, that shows/points out to why the research arguments in the thesis are important. It means what the researcher identifies that he/she and his/her fellows in his/her field of work do not know but need to know to answer a specific research question. For example, in our example of a research problem, decentralisation under a dominantly extractive economy, the literature on decentralization is mostly focused on political processes without explicitly analysing the material (socioeconomic) bases under which decentralisation takes place. In this case, this is the *“thesis-shaped hole”* this research is going to focus on, such that we can answer the research question *“what does “decentralization” mean in a society dominated by the economic, social and political dynamics emerging from mega-projects of an extractive nature?”.* The current literature has not researched this question; obviously, that literature cannot answer the questions; so, this is our *“thesis-shaped hole”* in the literature. Bear in mind that there might be important elements in the literature researched that may help us fill this hole. The existence of a hole in the literature does not necessarily mean that the literature is completely useless.

In brief, there is a dialectic and dynamic relationship between knowledge → literature research → knowledge → literature research →… and so on, with the literature, and the awareness of what the researcher seeks from the literature, becoming more focused, sophisticated and specialized in each level. All of this happens for the researcher to build his/her thesis-shaped hole, the field of ignorance the research is supposed to illuminate with new knowledge, the void of information and knowledge, the empty space in the researcher’s matrix of knowledge, which needs to be filled with new information and knowledge.

One of the most frequently asked questions by young researchers about the literature research is how to set its parameters, or, in other words, when should the researcher stop researching the literature and write his/her research conclusions. The good news is that there are answers and tricks that the researcher can benefit from. The bad news is that they are not straightforward and need to be adjusted to different circumstances. Let us look at some *rules of thumb*.[[2]](#footnote-2)

First, to choose and select the literature, there is nothing more practical than a good social theory, which tells the researcher how to define the problem, how to research it and how to interpret the results. ***Social sciences have not yet found a replacement for good social theory and serious, detailed, historical and social research informed by social theory*** in order to organize and undertake good research, including good literature research.

Bibliometrics, which can be useful to identify what has been published about a certain subject and where, depending on the data-bases utilised, are not a good tool to select what to read and what to look for in what is read. On the one hand, bibliometrics tends to do away with theoretical clarity, as research is no longer driven by an explicit theorical approach but only by the metrics of publications. On the other hand, it tends to be skewed towards mainstream ideas, as the literature is identified by frequency of citation, or the raking of the journals in which it is published, and so on, which, by design, tends to mainstream approaches. Finally, it requires assumptions about what is needed in order to study a certain subject, and such assumptions may come from the bibliometrics (reinforcing mainstream tendencies) or from mainstream ideas outside of the realm of research and not questioned by the research process. These assumptions are often not made explicit.

On the other hand, the lists of references in each relevant article and book that is read, and the good literature surveys that are occasionally produced in each field of study as scientific articles on their own merit – because they have an argument around which the literature is surveyed, and such an argument comes from social theory and practical research – are great sources of hints for literature research and of knowledge about the literature debate in the field.

Second, it is necessary to develop a broad and professional knowledge of the literature in the relevant scientific field of research and for the relevant question. We have already mentioned that without this knowledge it is not possible to set a research problem and research hypothesis, because the researcher could not know what the situation is, what the *“thesis-shaped holes”* could be, the knowledge that has already been produced, the research-shaped holes that have already been unravelled. Without this knowledge, the researcher would be going into the void without a single torch or any other source of light that could help him/her find the way around. We know that the research problem and hypotheses are based on what is already known, which provides the theoretical and methodological tools to explore and unravel the unknown, the research questions and its road map. Without such broad knowledge, none of these could be done.

Third, there is no escaping from the tension and conundrum associated with literature research. On the one hand, the researcher needs to be familiar, at professional level, with the literature in order to identify a research problem and hypothesis; on the other hand, the researcher needs a research problem and hypothesis in order to know what should be read and what to look for in the literature; finally, as research progresses, accumulated knowledge may change the formulation of the problem, of the hypothesis and of the literature research deemed necessary.

We can think of the solution for this problem as an imperfect one as we may never fully address the tensions between the different stages of work. However, we can think of a (imperfect as it might be) solution based on a stage-hierarchical process of literature research: (i) start with a concern that comes from broad knowledge (familiarity with a real problem that concerns the researcher, from practical experience, or from knowledge that comes from the literature and from the debate, etc.); (ii) move towards the formulation of a research problem based on that concern – this may require that the researcher spends some time checking if his/her concern is a real research problem or only an impression, if others have researched and written about this, what the state of the art of the literature in the specific field of research is, if there are gaps (holes) that need filling (that would be the focus of the research) – for example, going back to our problem of decentralisation, the researcher may know that large mineral projects affect the way local institutions and life are shaped, and he/she wants to research it to find out how this happens, why and whether this is a problem worth researching and how to formulate the problem. So, he/she can start by finding out what has been written about his type of problem; (iii) formulate the research problem and identify what needs to be known to formulate research hypothesis as road maps into the unknown; (iv) formulate the research hypotheses; (v) investigate how best to address the questions raised by the hypothesis and develop the research apparatus. Once the last stage is reached, further literature research is derived from the research problem, hypothesis and methodology, and from the new questions the research raises.

Researchers need to be aware of a few very practical problems. One, use *pegs* for your literature research – without a set of questions the researcher is looking for, and a structure to articulate those questions, literature surveys can become untamed, useless, counterproductive and go out of control. *Pegs* come out from your previous knowledge, your research problem and your research hypothesis, which tell the researcher what he/she is looking for, where, and how to interpret and articulate the questions and the findings.

Two, researchers should write and properly store good summaries of their readings as they go along. These summaries are not the final literature research but are important to build knowledge, identify tensions and gaps, develop methodologies, understand trends, build arguments and counter arguments, record sources and references. They are the raw material of a proper literature research.

Three, at any point in the literature research, the researcher needs to know or, at least, question how and to what extent the readings are related to his/her research problem and hypothesis, are helping to develop a methodology, or show the need to change the problem and hypothesis or reject the literature. There is no scientific way of answering these questions outside of a social theory.

Four, the researcher is trying to answer a set of questions, determined by one research problem and relevant hypotheses, within a limited time frame. Hence, the researcher, with the supervisor’s help, need to shut down the literature survey at some point, or the work will never finish – this can be done in line with the narrow parameters of the research project (a time frame, or a particular context, or something else). Any good piece of research opens the door to further research. This, however, can only be achieved if the research is finished and published. A standard good practice of any good research report (an article, a thesis, a book, etc.) is to conclude by identifying new research avenues and doors that have been opened and need exploring. There is no end to research because new knowledge shows new areas of ignorance and raise new research problems. So, there is no point in trying to find the research to end all research – it will never be found. So, finish, identify new research directions, publish and start again, this is the way forward.

Five, it is important to explore networks of scholars in the relevant field and topic – to know who is doing what, the main schools of thought, key readings. Be aware of the danger of mainstream, orthodoxies and theoretically week/flawed eclecticism. Be aware of buzz wording, which are popular substitutes for proper and sound concepts and theory.

**Evaluation of the research project**

The standard practice for evaluation of the research project is that it takes place at a doctoral workshop in June or July, where all first year students present their projects to examination panels. Independently of your mark, you may always have improvements to make in your proposal. It is important that you consider and make them, but, most importantly, is that your research moves forward (guided by a good enough proposal, of course). At the end of the day, research is going to be the most important process by which your research problems, hypotheses and methodologies are developed or changed. Do not get stuck on continuous and marginal improvements to a document.

Typically, the research projects should not exceed 10,000 words, which is roughly the size of a long, peer reviewed journal article. The project and the presentation should include: the context of your proposal (why you are interested in your topic and why you want to develop it in a PhD programme); the research problem, or the research shaped hole that you need to fill with your research; how the problem is addressed from the point of view of social theory (the theoretical framework); the hypotheses (or road map); the research methodology; how all of these (problem, theory, hypotheses and methodology) are supported by the literature research (or not); any fundamental ethical issues and risks or uncertainties your research project faces and how you intend to address them.

1. There are other ways of defining hypothesis, such as for inferential statistical and econometric research, in which a null and an alternative hypothesis are set. The null hypothesis states that there are no relationships between two measured phenomena or an association among groups (which is equivalent of saying that a person is innocent until proven guilty). The research is undertaken to prove, disprove, accept or reject the null hypothesis to a statistically significant level, such that the findings can be generalized to a defined population. However, this is not the main concern in these notes, because these kinds of studies have already been standardized in the research literature. These notes are concerned with the explanation of the concepts and structures of a research project and not with detailed explanations of research methods. [↑](#footnote-ref-1)
2. The English concept *“rule of thumb”* refers to a principle with broad application that is not intended to be strictly accurate or reliable for every situation. It refers to an easily learned and easily applied procedure or standard, based on practical experience rather than on rigorous theory. [↑](#footnote-ref-2)