978-1-009-10039-7 — Finding Your Social Science Project
John Gerring , Jason Seawright
Excerpt
More Information

1 Introduction

There is no royal road to science, and only those who do not dread the fatiguing climb of its steep paths have a chance of gaining its luminous summits.

Karl Marx, "Preface to the French Edition," *Capital*

If we knew what we were doing, we wouldn't call it research.

Albert Einstein (attributed)

Finding a topic for research should be easy. Presumably, anyone with an interest in social science has a primal urge to explore the world. One has questions, and one wants better answers than can be found in extant work.

Yet, if the researcher follows their curiosity so many potential topics beckon that it may be difficult to settle on just one. Finding a topic is easy. Finding an *optimal* topic is not so easy.

We have all faced this dilemma at some point in our careers. It is the scientific equivalent of writer's block. How does one go about "finding a topic"? Why are some people better at identifying fruitful topics than others? What should one advise someone who is struggling to find a topic for their thesis, article, or book?

Little assistance will be found in the annals of social science methodology. Consigned to metaphor – bells, brainstorms, dreams, flashes, impregnations, light bulbs, showers, sparks, and whatnot – inspiration falls outside the traditional rubric of methodology. Rarely is exploratory research included in courses or textbooks.

One might imagine that insights could be garnered from published work. Surely, the product of our endeavors sheds light on its origins. However, articles and books are generally unrevealing. Scientific journals do not provide space to expatiate on the origins of ideas or their subsequent development. While books offer more room for reflection, authors are rarely forthcoming.

This is a symptom of a deeper malaise. For professional reasons, authors of scientific studies are forced to engage in an elaborate and stylized game of deception. In order to avoid charges of "curve-fitting" or "fishing" they must adopt the Dogma of Immaculate Conception. Accordingly, they narrate their project as if the theory was hatched in complete isolation from the data used to test it.

Published work follows a standard protocol. Typically, the author begins by outlining a topic or research question. Then, they state a general theory, and thence a specific hypothesis and research design. Finally, the evidence is presented and discussed, and concluding thoughts are offered. Scientific studies thus present an appearance of order and predictability, a step-by-step descent down the ladder of abstraction.

This is nothing at all like the progress of most research – which, in our experience, is circuitous and unpredictable.[1] Unfortunately, we learn little about this process as it is not in the author's interest to divulge deviations from scientific orthodoxy. The early stages of the scientific journey therefore remain mysterious. The *search* part of research is not well understood.[2]

To be clear, there is no right or wrong way to begin. All that matters is where one ends up. And yet, where one ends up has a lot to do with where one starts out. Decisions made at the beginning of a research project structure everything that follows, as changing topics midstream is costly. Once one has developed knowledge and expertise in an area it is difficult to retool. And once one has gathered evidence it is difficult, expensive, and sometimes impossible to revisit research sites, that is, archives, field sites, interviewees, or respondents.

The choice of topic serves as a critical juncture. Once that threshold is crossed the research process is highly path-dependent. It follows that the earliest stage of research, where a topic is identified, is the most crucial stage of all. Nothing of scientific interest is likely to arise from research on a topic that is trivial, redundant, or intractable. No matter how well-executed, little can be expected from it.

To set the stage for this book, we begin by laying out our vision of scientific exploration. Later sections of this chapter delve into specific features of the book that distinguish it from other texts on similar subjects. A brief *coda* reviews the relevant literature.

[1] See Howitt, Wilson (2014), Medawar (1963), Merton (1967: 4).
[2] See Swedberg (2019), White (2013).

An Eclectic, Holistic View of Scientific Exploration

Where should one look for inspiration when searching for a research topic? Where should one begin?

The immense oeuvre devoted to methodology, philosophy of science, and history of science suggests four general answers to this question, which we summarize under the labels: *abduction*, *appraisal*, *relevance*, and *theory*. These labels correspond to four long-standing paradigms of social science research, which we will briefly review.

Abduction. One group of writers adopts an essentially inductive approach to scientific inquiry, where explanations arise from the researcher's encounter with the world. Accordingly, one's job as a social scientist is to be attentive to signals emanating from the subject under study. To put oneself in the path of discovery is to put oneself in direct contact with some empirical reality. The setting might be ethnographic, archival, or statistical. Regardless of setting, one must "soak and poke" until one figures out that most basic of all scientific questions: *What the devil is going on here?*[3]

Appraisal. A second group of writers follows the precept that criteria for strong appraisal should guide the search for topics; only in this fashion will the area of truth be extended. In practical terms, this means that researchers should look for opportunities to exploit strong research designs, where identification is possible without a lot of potentially problematic assumptions about the data generating process. For causal questions, this entails settings where a treatment is randomly assigned (an experiment) or as-if randomly assigned (a natural experiment). To some, this position may recall the adage of the drunk who looks for their missing key under the lamppost (because that's the only area that is lit). In the drunk's defense, one might point out that efforts to discover a key in pitch-black darkness are likely to be unavailing. Efforts to resolve intractable problems without sufficient empirical evidence may not advance the cause of truth. Accordingly, the central question for this group of scholars focuses on appraisal: *Is it falsifiable (testable)?*[4]

Relevance. A third group of writers believes that important research arises out of important questions or problems, relevant to the real world. Ask a question of social significance and good things will follow. This position is

[3] See Glaser, Strauss (1967), Locke (2007), Peirce (1929, 1934, 1992), Swedberg (2012).
[4] This approach is implicit in methods texts with an emphasis on research design, e.g., Angrist, Pischke (2009), Dunning (2012), Gerber, Green (2012), Shadish, Cook, Campbell (2002).

often embraced by those who see social science as responding to problems in society. While there is no well-developed methodology attached to this position (indeed, there is some hostility to the apparent tyranny of methods in academe), there is a clear point of departure: *Does it matter?*[5]

Theory. A fourth group maintains that good research begins with a well-formed theory, sometimes couched in mathematical language, that is, formal theory. This approach is associated with the logico-deductive method,[6] though many who favor theory would not identify themselves with that tradition. In any case, those in the theory camp are perturbed by the ways in which data and statistical models seem to drive research agendas. They would also be perturbed by the problem-driven approach, which views social science as wedded to normative concerns. From their perspective, the best point of departure for important research is abstract: *What's your theory?*[7]

Putting the Pieces Together

In our view, there is much to be said for all four points of departure – abduction, appraisal, relevance, and theory. They are all so plausible and so indispensable that we find ourselves unable to dismiss any of them. Each seems eminently useful in particular contexts.

We conclude that there is no Archimedean point of entry to social science research. There are, instead, many possible points of entry. Indeed, finding a topic for research is even more wide-open than the four-part typology introduced above suggests. One might begin with a general topic, a research question, a key concept, a general theory, a specific hypothesis, a compelling anomaly, an event, a research venue (e.g., a site, archive, or dataset), a method of analysis, and so forth. Accordingly, research may be problem-driven, question-driven, theory-driven, method-driven, or data-driven. Typically, it is a mixture of them all.

Moreover, the various traditions of research introduced above should be understood as complements rather than rivals. To be successful, a social science project must be successful along multiple dimensions. This includes the key concept(s), the theory, the research design, the empirical terrain, and the findings. While a single project is likely to innovate on only one or two dimensions, the finished research must satisfice along other dimensions. A fascinating theory is not very convincing without a strong research design. A

[5] See Flyvbjerg (2001), Shapiro (2005), Shapiro, Smith, Masoud (2004), Smith (2002).
[6] See Hempel, Oppenheim (1948), Popper (1968[1934]).
[7] See Clarke, Primo (2012), Eidlin (2011), Mearsheimer, Walt (2013).

Cambridge University Press & Assessment
978-1-009-10039-7 — Finding Your Social Science Project
John Gerring , Jason Seawright
Excerpt
More Information

new empirical terrain is not much good without a theory to explain it. A new concept is not much good – indeed, is not even understandable – if it is not connected to other elements of research. And so forth.

Consider the apparent dichotomy between theory (deduction) and empirics (induction). Barefoot empiricists must – at some point – consider the contribution of their data peregrinations to theory. After all, there is no such thing as a purely empirical contribution. (What is it a contribution to?) Likewise, those who claim to adhere to an a priori, theoretical approach must be conscious of the data they will use to test their theory. After all, nothing of scientific interest will arise from a question that is empirically intractable, however theoretically compelling the question might be. (Does God exist?) We conclude that exploratory research ought to be informed by theory *and* data, which one might envision (loosely) as a dialectical process. An extant theory suggests *A* while the data suggests *B*, leading to a new synthesis.

In the business of constructing social science everything is connected to everything else. Empirical investigation is contingent on preformed concepts and theories as well as our general notions of the world; and yet, further investigation may alter these notions in unpredictable ways. In so doing, the researcher revises their conception of what they are studying. A change in concepts entails a change in theory entails a change in research design entails a change in concepts . . . The knee bone is connected to the thigh bone.

Because everything is connected, everything is contingent in the early stages of research. Consequently, researchers are obliged to consider all elements of social science methodology before settling on a research topic. The sooner these diverse elements are brought into view the more efficient the exploratory process is likely to be. Without this bird's-eye view one may become so enamored of a theory, a research setting, or a problem of special concern that other desiderata are neglected. In this fashion, a great deal of time may be spent – and perhaps wasted – on a project that does not have strong legs.

What we have said so far suggests an *eclectic* and *holistic* approach to exploration. At the same time, it is important to clarify that this approach applies only to exploratory research. Once one moves to a *confirmatory* style of research it is appropriate to separate theory and testing (insofar as possible) so that the truth of a hypothesis can be properly evaluated (see Chapter 11).

A Sandbox of Options

If scientific exploration is eclectic and holistic, one may doubt the prospect of identifying a standard set of rules or procedures to guide researchers in their search for an optimal research topic. This point is reinforced when one considers the great variety of researchers and research topics.

Some researchers have trouble generating new ideas; others generate lots of ideas but have difficulty choosing among them. Some think deeply about a single problem, to the point of exhaustion; they would probably benefit from diversifying their portfolio. Others flit from subject to subject without building expertise; they could probably benefit from greater focus. Some can't see the forest for the trees; others can't see the trees for the forest. People are different. They have different strengths and weaknesses, and no piece of advice is likely to be applicable to everyone.

Likewise, subject areas are wildly different. Some are amenable to data-driven analysis. For others, there is little data available but many opportunities for gathering evidence "in the field." Some topics are organized around well-established theoretical frameworks. Others float loosely without any theoretical mooring. And so forth.

Under the circumstances, it is difficult to imagine constructing standard guidelines. What is good for Sid might not be good for Nancy. What is good for international relations might not be good for behavioral psychology.

Moreover, the process of finding a topic involves a certain amount of stochastic variability ("serendipity") that cannot be explained by features inherent in researchers or topics. A key finding from our own research is that there are few consistent predictors of research success (Chapter 2).

Accordingly, we embrace a *sandbox* model of exploratory research. We stipulate that there are many ways to search for a topic. Which one will work for you in a given instance we cannot say, for we don't know you, your research area, or the resources at your disposal. Under the circumstances, it would be presumptuous to offer specific advice.

What we can do is to lay out some of the options – approaches and techniques that have proven useful to others and might be useful for you. None of these options is very time-consuming or expensive, and only a few demand advanced methodological skills. So, the cost of exploration is low. Readers are encouraged to try them out to see which ones work – generating plausible ideas for future projects.

This open-ended framework should be helpful for those who feel that they do not know where to start or feel stuck in a rut. Even those who feel confident about how to proceed may find it helpful. Note that although one may already have identified a provisional topic, it is possible that an even better topic awaits. To reach this lodestar it may be necessary to think about new things, or to think about old things in new ways. You should be prepared to step outside your comfort zone.

You should also be prepared to play. The metaphor of a "sandbox" suggests that good ideas arise out of a spirit of playful provocation. So, throw some sand around and see what shapes arise. Show them to your friends, gauge their responses, and your own. Make revisions accordingly. This is the general idea. Take the job seriously, but not too seriously. Coming up with new ideas should be fun.

This Book

This book aims to enhance scholarly productivity by focusing on the task of topic selection. It should be useful for those who are just setting out and for those who (like the authors) have been kicking around for a while. Several features set this book apart from other work on related subjects (reviewed below).

First, we presume a solid background in social science methodology. This book on beginnings should not be confused with a beginner's guide.[8]

Second, we focus narrowly on exploration, which we understand as the period preceding a researcher's commitment to a specific project.[9] After that, it is possible to conduct "confirmatory" research, where a hypothesis is rigorously tested. We do not have much to say about the latter.

Third, we focus on social science, including the disciplines of economics, political science, sociology, psychology, and their various cognates and offshoots.[10] By contrast, most work on our subject is concerned with natural

[8] For those in search of introductory guides to social science methodology there are many options, e.g., Gerring (2012b), Gray et al. (2007), Kellstedt, Whitten (2018), King, Keohane, Verba (1994), Shively (2016). Specialized texts focus on causal inference and econometrics (Angrist, Pischke 2009; Morgan, Winship 2015; Wooldridge 2016), experimental and quasi-experimental research design (Dunning 2012; Gerber, Green 2012; Shadish, Cook, Campbell 2002), case study research (Gerring 2017), field research (Kapiszewski, McLean, Read 2015), ethnography (Agar 1996), and multi-method research (Seawright 2016).

[9] This is similar to what Swedberg (2014a: 26) calls a "prestudy."

[10] Although examples and citations may betray our disciplinary home in political science, nothing that we say is bound by the arbitrary walls that separate these overlapping fields. We exclude anthropology from our list of social sciences as most work in that field nowadays leans heavily toward the humanities, eschewing the ideal of a generalizing science of human affairs.

science, technical invention, or other realms of creative endeavor (e.g., art, music). We recognize that the creative process has some generic features, and occasionally we draw anecdotes or lessons from other regions. However, many features are particular to the social sciences, justifying our narrow focus.

Fourth, within the rubric of social science we focus on work that is empirical, rather than purely theoretical or methodological. We also give preference to work that is generalizing – aiming to shed light on a population of cases rather than on one or several cases of special interest (though the latter may serve as an empirical focus).

Fifth, in deference to well-established practice in the social sciences, our focus is primarily on work whose goal is causal explanation. Even so, much of what we say should also be relevant for work that is descriptive or predictive.[11]

Finally, while other studies focus on innovation as a product of genetics, core personality, upbringing, or education, this book focuses on aspects of research that individuals have control over. It is intended as a practical guide – not a muse, a philosophical study, or a social-psychological analysis. We are interested in what researchers can do to enhance their own creativity and productivity.

Special Features

There is no established method for studying exploratory research. Typically, writers opine about their own experiences, supplemented by a reading of the literature and stories drawn from the history of science. We shall do the same, with several additional features. These include (a) surveys of practicing social scientists, (b) copious examples drawn from social science research, and (c) causal diagrams.

Surveys

To gain traction on our subject we conducted a series of surveys and open-ended interviews with graduate students (pre-PhD) and researchers

[11] In our view, causal inference is probably overemphasized, and descriptive work de-emphasized, at least within political science (Gerring 2012a). Nonetheless, causality surely deserves a central place in social science and its preeminence seems secure for the time being. Conveniently, causal work lends itself to distinct phases of research, one in which a hypothesis is identified and another in which it is tested – a key premise of this book. These phases are usually more difficult to separate in work of a descriptive nature.

(post-PhD). In this fashion, we attempt to develop a systematic, replicable methodology for the study of exploratory research. The methodology is laid out in Appendix A and some of the key findings are presented in Chapter 2.

However, it should be emphasized that this is not the sole basis for the observations and recommendations contained in subsequent chapters. It is not possible – at least not yet – to empirically test all propositions relevant to exploratory research.

Examples

Throughout the text we offer examples of social science in use. These examples – typically from published work – serve to make general principles concrete and provide readers with a place to go for more details if they wish to pursue a particular approach.

It should be clarified that we are not endorsing the validity of these studies, most of which fall outside our areas of competence. Some of the cited research is old and has doubtless been superseded. Other studies may be flawed, though perhaps in ways that are methodologically informative.

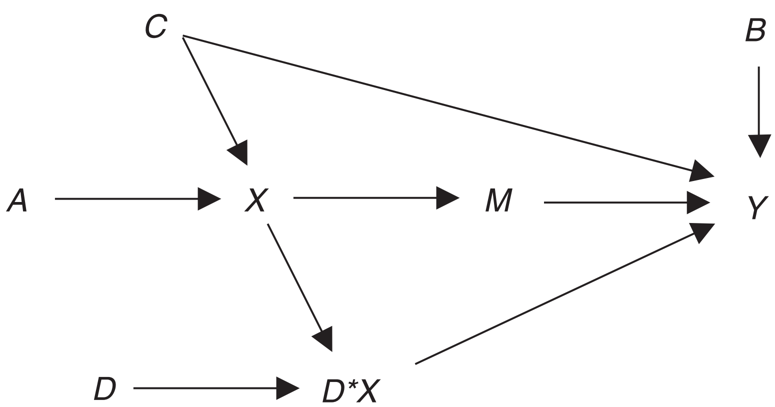
Our purpose in discussing a particular study is to illustrate a particular methodological point, of which the cited study offers a good example. That is all that readers should infer.

Causal Diagrams

In discussing causal relationships, it is important to distinguish the role of various factors. To that end, we make frequent use of causal diagrams. While every diagram is different, the most common elements are illustrated in Figure 1.1.

By convention, X is the causal factor of theoretical interest and Y is the outcome of interest. Additional causes of Y, if uncorrelated with X, are understood as noise (B). A common-cause confounder (C) affects both X and Y. A moderator (D) affects the impact of X on Y. An antecedent cause (A) lies prior to X and may serve as an instrument in an instrumental-variable analysis. Background factors of all sorts, for example, A, B, C, or D, may be referred to generically as Z. A mechanism (M) lies in between X and Y, representing the pathway through which X generates a change in Y.

These terms, and this sort of diagram, will be referenced at various junctures in the book. Several clarifications follow.



Nodes represent a single variable or a vector. (For emphasis, the latter is often printed in bold.)

A	Antecedent cause, which may serve as an instrument in a two-stage analysis
X	Causal factor of theoretical interest ("treatment")
C	Common-cause confounder
M	Mechanism, aka pathway, intermediate variable, mediator
B	Cause of *Y* uncorrelated with *X* ("background noise")
D	Moderator
Z	Any background factor, e.g., *A*, *B*, *C*, or *D* (not shown)

Figure 1.1 A causal diagram

First, each factor in Figure 1.1 and subsequent figures may refer to a single variable or a vector of variables. In the latter case, we print the variable in bold when used in the text of the book.

Second, any causal diagram depicts *assumptions* about the world. They are true only by assertion. Nonetheless, the diagram plays an important role in clarifying those assumptions.

Third, causal diagrams take shape within the context of a specific piece of research. They are not static, ontological features of the world.

Finally, causal diagrams can serve two, quite different purposes.

If the goal is *theory elaboration*, only factors relevant to the theory (e.g., *X*, *M*, and *Y*) are likely to be included. Here, the causal diagram is a tool for the clarification of an idea.

If the goal is *causal inference*, the causal diagram should represent the entire data generating process (DGP), including the strategy of conditioning