Chapter 1

Introduction

This is a guide to designing and evaluating scientific research in the field of political science, broadly defined to include the study of public administration and international relations in addition to the core domains of national and comparative politics.

Research is about providing answers to questions. Political science deals with some of the most pressing questions related to our shared human experience; for example: *How can we devise just and effective political institutions? How can we avoid war between states and among communities within states? What are the intended and unintended effects of public and social policies?*

Research *design* ensures that the answers we provide are as valid as possible and are discovered as efficiently as possible. By studying this text you will learn to devise effective and efficient research plans that can deliver valid inferences and new insight. Moreover, you will become proficient in assessing the research plans and results of others. Being able to develop an informed opinion about the merits and shortcomings of social and political science research is important for every enlightened citizen and even more so for aspiring policy makers, public officials, diplomats, and political commentators. There is much to gain from mastering research design.

Research design is *not* a settled body of abstract theory but an evolving set of rules, recommendations, some theoretical starting points, and many practical considerations. These inform the main steps of the research process: problem definition, theory development, conceptualization, operationalization, variable selection, and case selection. Some of the material discussed in the chapters to follow, such as the problem of causal inference and the rules about case selection, is simply an extension of general logic. Other parts, like the advice on choosing a research question or selecting the right level of analysis, stem mostly from received wisdom and experience.

Research design is a *craft* as much as it is a science. As with any craft, one learns it better from apprenticeship under a master's supervision than from books alone. Nevertheless, this text can help you get started and serve as a reference along the way.

What Is Research?

Scientific, or academic, research is about providing answers to questions we do not know the answers to. It would not be re*search* if we already knew the answers. This should be obvious, but I can tell you from experience that it is not. Too often, students, and some scholars as well, want to do research only to show what they think they already know – newspapers have undue influence on what people think, states only care about their geopolitical interests, politicians are corrupt, and so on. These might be the right conclusions, but it is the wrong attitude. Research is about seeking answers with an open mind, not about pursuing confirmation for a pre-existing belief, conviction, or prejudice. In rhetoric you search for the best facts and arguments to support your position; in science you search for the position that can be best supported by the available facts and arguments.

Sometimes we can find the answers we are looking for in an encyclopedia, in other published works, or on the internet. But more often than not available research would not provide satisfactory answers and solutions to our problems, questions, and puzzles. Then we need to complement the existing scholarship by bringing in new observations, evidence, and/or theoretical ideas. This text is not about the first type of research, which can be referred to as a literature review, but about the second type, which not only summarizes what is already known but ventures into the uncharted in order to provide new, more valid and precise answers. Naturally, literature review remains an indispensable step to be performed before original research.

Successful research is as much about asking good questions as it is about providing good answers. We review what can be said about the latter in Chapter 2.

The product of research is often in the form of a written report. Especially in the social sciences it is generally assumed that an academic publication, such as a book, a journal article, or a thesis, is the only possible output of a research project. This need not be the case. For example, the product of research can be a predictive model, which might be described in a paper but could also just reside as a software implementation. Alternatively, it could be an interactive web-based visualization that efficiently describes and explores an original dataset. The product of research could also be a policy evaluation and advice – analysis that directly feeds into the selection of public policies or informs a decision about some political choice. In any case, research output needs to be communicated effectively, and this is one of the issues that Chapter 12 deals with.

What Is Science?

Scientific research is most often conducted in the world of academia, but there are other contexts where research is done. A journalist investigating a case has similar objectives – reconstructing a process

or explaining an outcome – to those of a political scientist, although their tools might differ in important respects (see Chapter 10). A political party office that builds a model of electoral behaviour to help increase its vote share would use many of the same techniques that a scientist would use to build a causal model. Political marketing companies rely on the same principles of research design to evaluate the effectiveness of political messages that an academic would use, and pollsters rely on the same theorems that would allow a scientist to estimate public opinion with minimum costs and maximum precision. Policy impact evaluations conducted by a consultancy firm are subject to the same challenges and limitations that an academic study would need to address. A crime detective analysing evidence works in a very similar way to a scientist evaluating competing hypotheses. We can continue adding examples – think about medical doctors, judges, and so on – but the point is clear: the academic world has no monopoly on research.

But if academic scientific research shares so much with other human endeavours, what are the features that distinguish it from journalism, criminal investigations, consultancy, and marketing? It is hard to define science. Scholars have compiled long lists of its essential features (for example, Gerring, 2012b), but the philosophical debates about the nature of science rage on (for a gentle introduction, see Chalmers, 1999). Nevertheless, a minimal definition of science need only highlight two crucial aspects: publicness and adherence to the scientific method.

Publicness

First, academic scientific research is public and open to scrutiny. It is not only that it is almost always, directly or indirectly, funded by the public and that it contributes to the common good rather than to private gain. The requirement of publicness goes to the very heart of what academic science is and how it differs from other applications of the human spirit and intelligence. While a private consultancy would jealously protect the details of its predictive models and algorithms, scientists must always disclose the methods they work with. While political party offices are free to keep any data they use undisclosed, any data employed in an academic project should be made freely accessible. While governments or pharmaceutical companies are not required to publicize any of the prospective evaluations of policies or drugs that they commission (although they should be), scientists have an obligation not to withhold findings based on whether they suit or not their favoured hypotheses or ideological predispositions.

Publicness and transparency of methods, data, and results are indispensable, because science is a community affair. A scientist's work is scrutinized by other academics (his or her peers) and made open to

critique by any valid argument. It is the way science proceeds: by collective scrutiny and criticism that allow for gradual improvements and the correction of mistakes; by replicating, adjusting, and, occasionally, overturning what others have done before.

The scientific method

Second, scientific research is subject to the scientific method. Consultancies and political parties are free to use whatever method they choose, including asking oracles and tossing sheep bones, to come up with their models and predictions; scientists are constrained to follow the rules of the scientific method. That would be fine, if all people who would describe themselves as social or political scientists (or even the subset who get paid by a university or a public research institute) would agree as to what these rules are.

A popular account of the scientific method follows the normative ideal put forward by logical positivists. In short, according to this view, scientists start with a theoretically motivated hypothesis, test the hypothesis with data, and proceed to conclusions rejecting the hypothesis that fail the empirical tests (see Popper, 1959; Hempel, 1965; and Chapter 3 of this text for details). This view is seductively clear and simple, but it has been disqualified both as a descriptive (how science works) and as a prescriptive (how science should work) model (Kuhn, 1962; Lakatos & Musgrave, 1970).

In contrast to the views of logical positivists, in contemporary political science most would agree with three major points. First, empirical puzzles and substantive problems are as common starting points for research work as theories. Second, theory testing is not the only and perhaps not the most important goal of science. Third, ideas are not simply tossed in the garbage bin of history at the first sign of empirical inadequacy. But beyond this consensus, there are a wide variety of ways of doing what goes under the label 'political science' research. This pluralism makes it hard to define a version of the scientific method that would satisfy all political scientists. Yet, we need such a definition to complement publicness as the second criterion delineating science from other human activities.

Instead of aiming for a single, all-encompassing definition, let us explore the main disagreements about how to do science within the field of political science. This should give us a sense of the most important issues involved and should also serve to position the current text in these debates.

Currently, the major dividing lines are three: between subjectivists and positivists (in a broad sense, not in the narrow sense of logical positivism); between empiricists and scientific realists; and between qualitative and quantitative researchers.

Subjectivism and positivism (in a broad sense) The most fundamental faultline is the one between subjectivists and positivists. According to subjectivism as a philosophical position, 'the essential, unique characteristic of human behavior is its subjective meaningfulness' (Diesing, 1966, p. 124). Hence, social science cannot be construed as a value-free pursuit of objective truths about the social and political worlds. In fact, the mere possibility of social *science* becomes highly suspect due to the irreducibly subjective nature of human perception and experience. According to subjectivists, research should be concerned with interpreting the meaning of and reflecting on the reasons for human action. Searching for, and even speaking of, social *mechanisms* and *causal* factors is not only futile, it is misguided and offensive. Under this strong subjectivist view – a view that critical theorists, post-modernist philosophers, and many interpretivists, reflectivists, and feminist scholars tend to espouse – social science can only function as radical *social critique*. Blurring the distinction between scientific research and advocacy/social action is detrimental to the practice of both. For an extensive overview of interpretivist political science, see the four volumes edited by Mark Bevir (2000).

In contemporary political science the opponents of radical subjectivists are not objectivists, as you might have expected, but a diverse group of scholars who would agree with many of the basic tenets of subjectivism but would resist taking them too far. For a lack of a better term, we refer to this group as positivists, although they are very different both from the often naive positivism of nineteenth-century social theorists and from twentieth-century logical positivism. In the aftermath of the seminal contributions of philosophers such as W. V. O. Quine (1951) and Thomas Kuhn (1962), cultural anthropologists such as Clifford Geertz (1973), and sociologists such as Peter Berger and Thomas Luckmann (1966), few if any social scientists would defend a completely objectivist worldview. Contemporary positivists – who include among their ranks the majority of political scientists around the world – would, for the most part, accept that social reality is not set in stone, objectively given, and directly accessible to human perception, but constructed and reconstructed through a variety of social processes. In other words, social reality is to a large degree *inter-subjective*. However, and here is the crucial difference from radical subjectivism, social *science* remains possible. Moreover, social science is conceived as a quest for the discovery and explanation of the causes and mechanisms of social phenomena, including individual events as well as broader regularities and patterns. And it is subject to transparent rules, standards, and procedures that ensure reproducibility, reliability, and validity of the results.

Subjectivists would object that one cannot entirely separate the observation of social facts from values and theoretical notions (see Chapter 3). This might be true, but only in an abstract and purely formal

philosophical sense. Even if there is no completely objective, value-free point of view, we can still do much to acknowledge and limit the influence of our particular values in doing research on politics and governance. While there may be no final truths, some inferences are still less valid than others (given a context), and the task of research is to discover the 'more' valid ones.

The position taken in this text is pragmatic. That is, we concede the philosophical upper hand to subjectivists but endorse a positivist outlook when it comes to the practice of research in political science. It remains important to be aware of the inherent limitations of social science highlighted by subjectivism, but within the territory outlined by these limitations, there is plenty of scope for scientific research subject to rigorous procedures and explicit standards. The rest of this text is devoted to presenting and explaining the logic of these standards and procedures in the context of research design in political science. For the most part, we will not engage with research in the radical subjectivist tradition (with the exception of one section of Chapter 2). But many of the lessons and insights subjectivists have to offer are implicitly integrated in the text, most notably in Chapters 2 and 4, which deal with, respectively, the status of theory and the process of conceptualization.

Empiricism and scientific realism The second divide in contemporary political science is less fundamental. While subjectivists and positivists disagree whether social reality can be studied scientifically in the first place, empiricists and scientific realists (who would both be positioned within the positivist camp) disagree about the ways to conduct the study. At a deeper philosophical level, empiricism and realism imply different ontological views (what is reality), but in practice these are only manifested as differences in epistemology (how to know reality). Empiricists deny reality to unobservable entities such as theoretical concepts and causal structures and usually adopt an instrumentalist view of theoretical assumptions. The latter means that the assumptions of our theories and models do not need to be realistic as long as they work; that is, as long as they prove useful for prediction and manipulation. A famous proponent of the instrumentalist view was the economist Milton Friedman, who argued in a much-cited passage (1953, p. 14) that

> the relevant question to ask about the 'assumptions' of a theory is not whether they are descriptively 'realistic,' for they never are, but whether they are sufficiently good approximations for the purpose in hand. And this question can be answered only by seeing whether the theory works, which means whether it yields sufficiently accurate predictions.

Empiricists are interested in successful prediction and intervention and care little about *understanding* the underlying structure of the world that generates the outcomes we observe. Even the mere concepts of 'causes' and 'effects' are suspect, as they cannot be directly observed. By contrast, scientific realists strive for a 'deep understanding' that goes beyond the instrumental uses of scientific results (see Chapter 6).

In contemporary political science it is hard to find 'pure' empiricists, but there are two popular styles of research that are closer to its tenets than to those of scientific realism. First, there is a strong line of theoretical work, very much inspired by economics and game theory, that develops models of political processes (legislative decision-making, for example) on the basis of extremely simplified (hence, instrumental) assumptions about human rationality and the nature of social interactions. But such work is often deployed to provide understanding rather than prediction only, in contrast to the goals of pure empiricism. Second, a lot of research in empirical political science, both in its quantitative and qualitative modes (see below), operates at a level very close to empirical reality, making only modest attempts to link explicitly to theory. Such research can be highly rigorous, but it is primarily interested in describing and exploring political phenomena, such as public opinion, electoral campaigns, or policy implementation, rather than building and testing abstract theoretical models. This second form of empiricism common in political science is content with systematic description and avoids even prediction, let alone deep theoretical understanding of reality, as a scientific goal.

In this text, the focus is on explanation and the discovery of causal structures. Causal structures and concepts, although not directly observable, are considered to be the primary targets of scientific inference. The endorsement of scientific realism is reflected heavily in the structure of this text, which deals extensively with the role of theory development (Chapter 3), operationalization of theoretical concepts into observable variables (Chapter 4), and causal explanations (Chapters 6–11).

Quantitative and qualitative The third major dividing line in contemporary political science concerns the mode of research. It runs between those who are more quantitatively oriented (that is, they use numbers and statistics) and those who do qualitative research (that is, they do dense case studies) (Goertz & Mahoney, 2012). This rift is largely artificial and receding. According to the approach taken in this text, quantitative and qualitative research are both subject to the same rules and challenges of inference. In short, there is a place for and value in both. Chapter 11 will also explain how they can be fruitfully complemented.

While there need not be *fundamental* (ontological and epistemological) differences between quantitative research, qualitative comparative

research, and case studies, there are important differences in what each can achieve. Under such a pragmatic understanding of science, there is a place under the scientific sun for a multitude of research approaches, designs, methodologies, and techniques. Some make heavy use of numbers, others rely exclusively on words to advance an argument; some would trade detail for generality, others would rather have precision rather than a broad scope. These are no reasons to conclude, however, that researchers who use numbers are subject to different rules of logic from those who do not. Many things go, but that does not mean that anything goes.

A minimal definition In view of the distinctions made above, is there after all *some* way left to define the scientific method that is logically sound and at the same time fair to the actual diversity of practice of scientific research? Instead of looking for a single definition, we are better off considering the various activities that make up the scientific process separately and putting forward some requirements for each step of the way. For example, theoretical ideas should be internally consistent and have clear and precise observable implications that lay the theories open to refutation. Measurement should be done in a replicable way that achieves valid representations of the underlying concepts. Inference from data should respect the rules of logic. Theoretical and empirical work, discovery and testing, exploration and confirmation, description and explanation might be related in more complex ways than suggested by logical positivism, but there are still some rules that govern each of these activities, that, taken together, make up the method of science.

To some, this discussion might appear unsatisfactory – after all, it seems that we have just kicked the bucket down the road by raising new questions, such as 'What is validity?' and 'What are the rules of inference from data?' These further questions, and more, will be dealt with in much greater detail in the chapters to follow.

In sum, science is characterized by (1) being public and transparent and (2) adherence to certain rules regarding theory development, measurement, inference from data, and other aspects of the research process. So, going back to the examples raised earlier in the chapter, how is the work of a political scientist different from that of a journalist? Well, apart from the fact that a journalist usually works under stricter deadlines and has access to different sources of evidence, as long as he or she derives conclusions in a transparent and rigorous way that respects the rules of inference, there need not be a real difference. Similarly, somebody developing an election prognosis model for a private company is doing scientific research, as long as the details of the model can be made public and open to scrutiny.

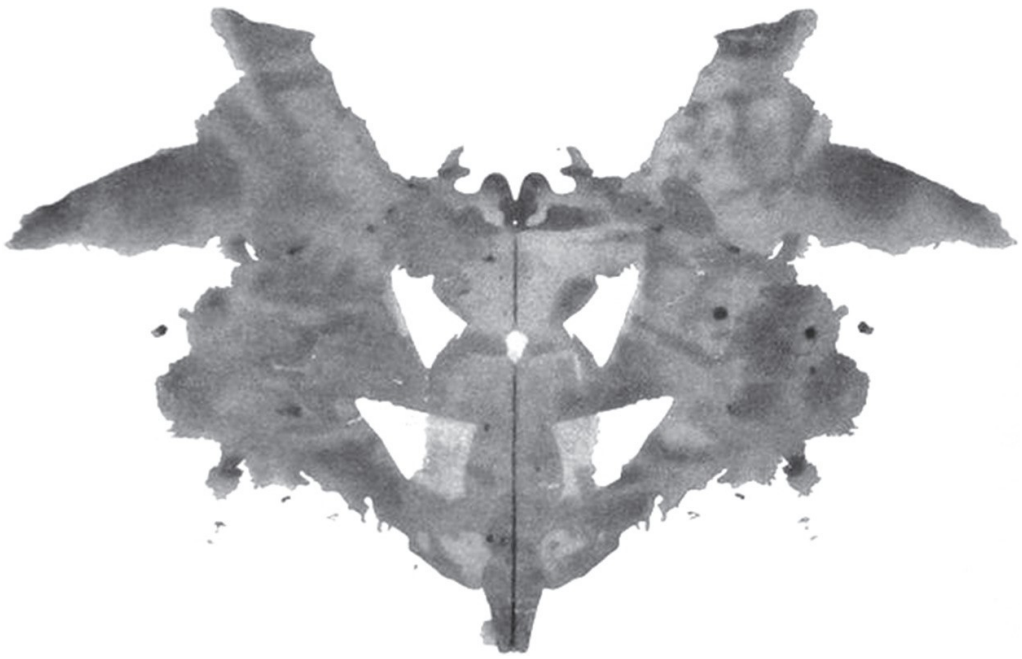
Why Research Design?

Few would disagree that science and research are important, but do we really need lessons in research design as such? After all, humans have been able to get by just fine without it for many centuries, and progress in some of the sciences pre-dates any focused concern about method and design. Isn't the innate human problem-solving capacity sufficient to guide substantive research? Don't we all just have the rules of inference encoded in our brains, so that an external guide is redundant?

Humans have been shaped by evolution to be excellent problem solvers, indeed. Faced with new information, we are quick to find similarities and differences and to build implicit causal models of how things are connected to each other (Sloman, 2005). People are very good at finding patterns. In fact, we are too good. This section of the chapter will argue, based on a wide selection of evidence, that there are significant biases to human judgement and problem-solving capacity which require us to pay close attention to how we plan and execute research projects, especially in a domain as emotionally and ideologically charged as politics.

Consider the famous Rorschach psychological tests. They involve showing people inkblots and asking them what they represent. Look at Figure 1.1, which shows one example. What do you see? People have little problem projecting all kinds of images onto the *random* spills of

Figure 1.1 *An example of a Rorschach inkblot (Number 1)*



Source: http://www.test-de-rorschach.com.ar/en/inkblots.htm.

ink on paper – bats, butterflies, moths, humans, and so on. It is *too* easy to find patterns in randomness. But what makes the Rorschach test a particularly good example of this tendency is that not only laymen have been fooled into seeing images in inkblots. Psychologists themselves have been fooled into thinking that the types of images people see in the inkblots are predictive of personality traits, sexual orientation, and emotional functioning. The test has been in use since the 1920s, but at the moment there is very little and much contested evidence that any correlations between what people tend to see in the inkblots and the features of their personalities exist (Chapman & Chapman, 1969). What irony! Not only can regular people see patterns in randomness, but trained professionals and scientists can be just as good (or, rather, just as bad).

For a different example of the human tendency to over-interpret randomness and sustain belief in the utility of their efforts even in the face of recurrent evidence to the contrary, consider the following facts about 'expert' judgements in different domains. On average and over a sufficiently long period, the return to investment of a financial portfolio managed by professional fund managers does not beat a simple index that tracks the average performance of the stock exchange (Fama & French, 2010). And this one: predictions made by political scientists, economists, journalists, and other experts about future political events were only a little better than random guessing (or a dart-throwing chimpanzee, if you prefer the image) (Tetlock, 2006). In general, research has concluded that in

> nearly every study of experts carried out within the judgment and decision-making approach, experience has been shown to be unrelated to the empirical accuracy of expert judgments. (Hammond, 1996, p. 278)

On a sidenote, while an individual expert's judgement is often no better than a random guess, averaging expert opinions seems to provide some predictive leverage. And predictive markets have shown how powerful the wisdom of crowds can be in predicting election winners, even if each individual has very limited and biased information.

The examples above are not mere anecdotes. We actually know quite a lot about how human judgement and inference from data systematically deviate from normative models. Over the past few decades, social psychologists and cognitive scientists have catalogued a large number of biases in and limitations to human decision-making, many of which would directly impair the scientific process. The literature on biases and heuristics is enormous and cannot be summarized here – see Kahneman & Tversky (2000) for an introduction or Kahneman (2011) for a popular account – but we can mention just a few examples to give some illustrations of the problems involved.

The *confirmation bias* is perhaps the most relevant for the context of research design. It relates to the human tendency to seek only information that would confirm a guess or a hypothesis but no information that would contradict it. For example, in a famous experiment conducted first by Peter Wason in 1960, people (students actually) were given a short sequence of numbers, 2, 4, and 6, and asked to discover the rule generating the sequence. They were also given the opportunity to test whether any other sequence of numbers fits the rule or not. Most people would quickly recognize a pattern and form the conjecture that the numbers are generated by the rule 'increasing even numbers', and they would ask whether 8, 10, 12, and so on, fit the rule. Crucially, having formed this conjecture, people would *not* test whether sequences of numbers that *would not pass the test if their initial guess is correct* actually do pass the test. So most people would not even ask whether the sequences 1, 2, 7, or 23, 24, 25, or −2.5, 0, 134 fit the rule. But the sequence *could* as well have been generated by the rule 'Any increasing number', or 'Any integer', or even 'Any real number'. By seeking only information that would confirm their pre-established beliefs, people tend to miss alternative explanations that might just as well account for the patterns.

It is a hallmark of the scientific method in contrast to casual human thinking to search for *disconfirmatory* evidence and properly test intuitions and ideas. This attitude needs some reinforcement, since apparently it does not come to us all that naturally, and our innate tendency to quickly find patterns needs to be disciplined. A good research design fulfils these purposes.

There are further limitations to human decision-making – among others, the *hindsight bias* that makes us believe that events have been more foreseeable than they actually were; *framing effects* that lead people to make different inferences from the same information with only some innocuous-looking change of words; the *availability bias* that makes us take into account only the most salient information that comes first to mind; and so on (Kahneman, Slovic, & Tversky, 1982). People have been discovered to be particularly challenged in mentally manipulating probabilities (especially when they are not expressed as natural frequencies), which is of course an essential skill for drawing correct inferences from observations. For example, people would often judge the event 'dying from a heart attack' as less probable than 'being obese *and* dying from a heart attack', although obesity is clearly just one of the possible causes of heart attack, so the first probability cannot *logically* be smaller than the second. This is the so-called *conjunctional fallacy*, studied by Kahneman and Tversky (as discussed in Sloman, 2005, p. 105).

Moreover, even our direct perceptions – of pain, for example (Quattrone & Tversky, 1984) – and memories (Tversky & Marsh, 2000) can be subject to biases and self-deceptions. But if human cognition can be so easily manipulated, scientific measurement and observation surely

need to be subject to some rules and standards. Research design helps discipline the process of data collection – what kind of evidence should be sought and how – to overcome the limitations of informal human cognition that have been shown to affect laymen and experts alike.

Importantly for students of political science, human reasoning is subject to specific biases related to the political and ideological predispositions people hold. In one study Brendan Nyhan and Jason Reifler (2010) made people read a misleading claim in a news article and a correction afterwards. When the correction contradicted the ideological beliefs of the subjects, they failed to take it into account, and in some cases the level of misperceptions actually increased in response to the provision of the correct information (the so-called *backfire effect*). And just being smart does not always protect against ideological cognitive biases. As Dan Kahan and colleagues showed (2013), people with higher than average numeracy skills were just as likely to fail to draw the correct inferences from factual data, *when* the inferences would conflict with their prior political ideological beliefs. Political scientists cannot be assumed to be exempt from these common fallacies of human reasoning, which is yet another argument for the importance of research design and rigorous data analysis.

While humans can be too quick to find patterns in random data and to seek confirmatory evidence only, the opposite problem also exists. There are many famous examples where people have failed to recognize real connections in the world despite centuries of experience and observation. Just consider that the modern theory of *plate tectonics* (Oreskes, 2003) was only developed during the 1960s, although it has been possible to observe that the shapes of Africa and South America fit together like two pieces from a puzzle ever since the first good maps of the two regions became available (roughly, since the late sixteenth century). The discovery that *scurvy* – a disease that decimated the crews of thousands of ships for many centuries – is a result of vitamin C deficiency and can be prevented simply by eating fruit came only in the twenty-first century despite several occasions of accidental 'near discoveries' and extremely high pay-offs to finding a cure (see Brown, 2003 for a popular historical account). Even though *cholera* has been with humans since ancient times, it was only in 1854 that a careful data analysis by Dr John Snow revealed that the disease is transmitted by infected water (and it took a few more years to persuade the medical establishment of this fact; see Johnson, 2006).

Medical examples are not the only ones demonstrating our limited capacity to uncover 'real' patterns in observational data, but they are often well-documented. The social sciences as well abound with instances of false findings and real associations that remained hidden and, doubtlessly, there are many still remaining that we have not noticed yet. That is why, in addition to helping the production of *valid* inferences rather

than chasing randomness, research design ensures that the process of discovery is as *efficient* as possible, for example by directing focus on theoretically anomalous cases.

As valuable as it is, research design is no substitute for substantive knowledge. It is a framework for the organization of substantive knowledge in the most efficient and effective way to provide novel, valid, and reliable inferences. But no clever design gimmick can salvage a research project that is insufficiently embedded in the substantive problems that animate a research field and that ignores the state of the art of existing scholarship. There are no tips or tricks on offer for choosing a research question, measuring and operationalizing concepts, or developing explanations that can *make up* for a lack of substance. In fact, you can only make full use of the advice offered in this text if you already have enough substantive knowledge to identify, for example, a credible instrumental variable for causal inference (Chapter 8), know which factors to control for (Chapter 9), and know what indicators to use to measure your concepts (Chapter 4). In short, proper research design is a necessary but not a sufficient condition for successful research projects.

But perhaps the importance of research design is a thing of the past, now that we live in an age of big data? Don't the already massive and growing amounts of social, political, and economic data available speak for themselves? Do we still need rules about case selection and sampling now that there are millions of observations at our fingertips? In 2015, the academic journal *PS: Political Science and Politics* ran a symposium addressing precisely these questions (Clark & Golder, 2015). The conclusion was that attention to theory, research design, and the rules of causal inference are even more important in the age of big data. Data are not the same as creative insight, and big data plus computer algorithms cannot substitute for sound thinking, good research design, and careful data analysis. But the *combination* of big data and rigorous social-scientific methods can make a strong contribution to science (Monroe et al., 2015). Current students of political science are well placed to benefit from the synergy between the two.

What Is Political Science?

Little of the discussion so far is specific to the field of *political* science. The limitations to informal human reasoning have general validity, the challenges to define the scientific method concern all social and human sciences, and the value of design to the research process transcends any particular field. So what is it that makes this text one about *political* science, as the title announces?

In fact, not much, apart from the selective emphasis on certain topics (observational designs for causal inference, comparative, and within-case

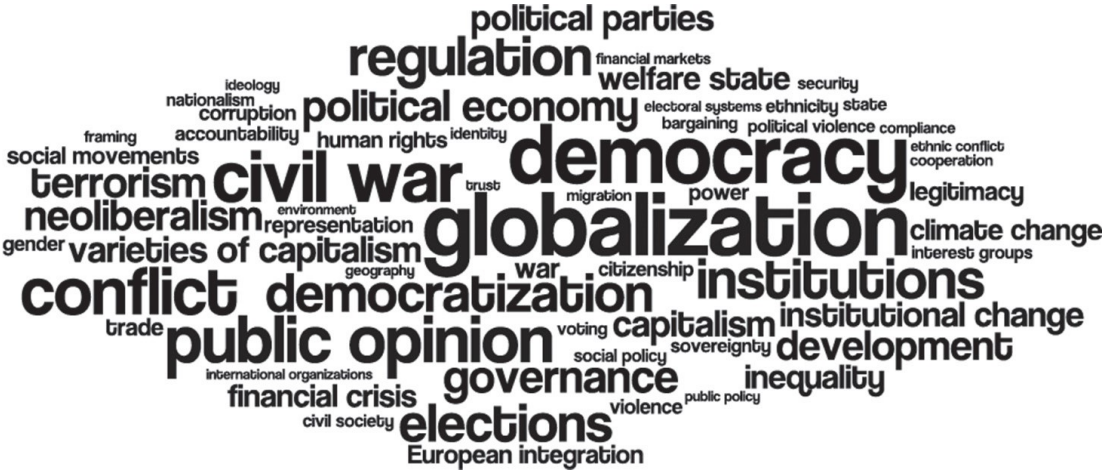
analysis), the relative neglect of others (prediction, machine learning for theory generation), and the choice of most (though not all) examples in the chapters to follow. There is a methodological unity to the social sciences and beyond. Whether one wants to understand the impact of electoral institutions on political polarization, of early childhood education on learning outcomes later in life, of social programs on future employment, or even of exercise on health, the challenges involved and the potential solutions are very similar. This methodological unity is good news, because it means that learning research design in the context of political science gives lessons with a wider field of application (see below). That being said, there are certain peculiarities to research in politics that make some issues more salient than for, say, educational science or economics. But let us first see what *is* political science.

Clearly, political science is the study of politics – voting, elections, and political parties probably first come to mind, especially if you live in a democracy – but it is also more than that. A useful scheme for organizing things is given by the triad *polity*, *politics*, *policy*. Political science is about all three and about their connections as well. The *polity* part of the triad concerns the political organization of nations, the fundamental set-up of the state, the forms of government, and the division of power within societies. It is about themes such as federalism, separation of powers, democracy, and so on. International relations focuses on the relationships *between* polities, and studies conflict and cooperation among the nation states and different forms of international governance. The *politics* part of the triad concerns the political process: the way power is exercised within states and societies. It is about decision-making within political institutions, such as legislatures, and the links between institutions, such as those between the president and the parliament; it is also about political competition, citizen representation, and interest mobilization. Finally, the *policy* part of the triad is focused on the making, implementation, and evaluation of public policies – the products of the political process – in various domains, such as the economy, foreign affairs, or the environment.

To see for yourself what contemporary political scientists actually study, have a look at Figure 1.2. The figure shows the 60 most popular keywords of more than 10,000 academic articles published in the top 20 journals in the field of political science between 2003 and 2013. In this word cloud, the size of each word is proportional to its popularity. Clearly, modern political science is mostly concerned with problems of globalization, democracy, conflict, regulation, development, and more – themes that cut through the division between domestic and international politics, and between politics and administration.

In some countries, the study of public policy and administration is institutionalized as a separate discipline, rather than as a subfield of political science. But it makes a lot of sense to consider the study of

Figure 1.2 *A word cloud of the 60 most popular keywords in political science, based on 10,000 articles published in the top 20 journals over the last ten years*



policy, politics, and polity together, not only because of the important substantive linkages and feedbacks between the parts of the triad and the overlap of issues they focus on, but also to highlight the common methodological ground that these academic (sub)fields share. Therefore, this text adopts a rather encompassing view of political science to include not only the 'core' areas of electoral, legislative, and comparative politics but also international relations and public policy and administration.

Such an encompassing view reflects to heart one of the most influential perspectives on politics conceived a system that translates inputs from society into outputs in the form of policies (Easton, 1965). While a narrow view of political science focuses predominantly on the input side and policy analysis on the output side, a comprehensive view needs to encompass the entire process. Take Easton's own famous definition of politics as 'the authoritative allocation of value' for society, which echoes Lasswell's earlier view that politics is about who gets what, when, and how (1936). The societal allocation of value is not completed when an election is held, or when a government is formed, and not even when a law is passed. To understand the real impact of politics on our lives, an integrated perspective is needed that pays attention to the entire cycle from the formulation of social demands to policy implementation.

To reflect an integrated perspective, the substantive themes and examples used to illustrate the problems of research design in this text are balanced between the different subfields of political science.

The study of politics is a science, but also a vocation. In the words of Aaron Wildavsky (1979), political and policy analysis should be about 'speaking truth to power'. It is important to keep this in mind while learning about research design. Unlike politicians, activists, policy

entrepreneurs, civil servants, diplomats, or lobbyists, political scientists have an obligation only to truth and not to a particular group, cause, or ideology – this is, after all, what makes their work distinctive and valuable. Each of us harbours some political preferences and predispositions of different colours; the scientific method ensures that these are not allowed to play a role in the process of research other than in motivating the choice of problems to study.

Moreover, science has the obligation to report how much uncertainty surrounds what we take for established truths in politics, governance, and international relations. In the public sphere, it is frequently taken for granted that the experts with the sharper predictions, broader generalizations, louder advice, and more self-assured conclusions are the ones who benefit society the most. Not so. Often, the major contribution of a scientific analysis would be to point out how little we know about the likely effects of a certain policy change, how unpredictable the course of a political event is, or how uncertain the future impact of some new institution could be. Research design and the rules of inference help delineate what can be known from what we can only speculate about, and thus provide a much greater service to society than the illusory confidence of clairvoyants, political pundits, and other 'experts'. The concluding chapter will return to these issues, but now it is time to reveal what the chapters in between are all about.

The Purposes of This Text

The purposes and motivation of this text were already touched upon in the Preface and at the beginning of this chapter, but they require some more explication.

The main purpose is to provide students of political science with the knowledge and tools to develop ideas for original research, and to translate these ideas into potent and workable research designs. By studying this text, you can learn how the research process works, from the initial idea to the final communication of the results. You can also get a comprehensive understanding of the menu of available research design options, as well as the ability to assess their major strengths and shortcomings. Taken together, detailed knowledge of the research process and the skills to make the right design choices are powerful assets.

The practical utility of these assets depends on the personal and professional interests of the reader. At the very least, it includes the preparation of research proposals for student papers and graduate theses; for academic articles and scientific reports; for grants, tenders, and project applications.

Knowing how to design research implies that you will have the knowledge and tools to assess the research results and plans of others. We are

all much more often consumers than producers of research. The ability to evaluate claims and inferences made in academic publications or the press is an important ability to have. By studying the material in these chapters, you can learn which claims to trust and which to ignore, and how to assess the real rather than the reported uncertainty of scientific and other conclusions. Altogether, this amounts to a significant upgrade of general critical thinking skills.

There are three main distinctive features of this book. First, the book provides a balanced coverage of descriptive and explanatory, experimental and observational, large-N (statistical) and small-N (comparative), cross-case and within-case (single case) research in a single text.

Second, the book delivers an integrated and consistent picture that not only catalogues different ways of doing research but puts the various designs into a coherent framework. Many edited volumes on research design have excellent individual chapters, but which, however, do not always easily fit well together.

Third, the book makes the lessons of research design as accessible as possible without sacrificing too much of the depth and subtlety of the arguments. To this end, I have avoided any specialized notation, formalizations, equations, and even footnotes and endnotes. I have also used plenty of examples throughout the chapters. The examples are of two types. The first consists of short, 'toy' examples about very general problems that most people can relate to without having a specialized background, such as the links between wealth and voting. The purpose of these is to quickly illustrate a point or add some substantive 'meat' to an abstract discussion. The second type consists of examples from actual published research, introduced in more depth and at a greater resolution. These serve to show how research design problems are tackled in a real-life setting and to give a taste of how various modes of research in political science work. The book is also complemented by an extensive index that can help orientation in the material and that showcases the cross-links among the various topics and issues throughout the chapters.

Of course, no single text on research design and methodology can present a detailed and comprehensive picture of the variety of research in contemporary political sciences. The major limitations of the current one are related to the level of operation, the relative neglect of philosophical issues in favour of more applied ones, and the lack of coverage of techniques for data analysis.

It is customary in military theory to divide the planning of a military campaign into strategic, operational, and tactical issues. Using this analogy, this text is positioned at the level of grand research design *strategy* and at the intermediate *operational* level that connects strategy and tactics. But the minute *tactical* choices involved in implementing any single research strategy and operation are outside our scope. There is a simple reason for that – the details of *any* particular design could easily fill up not a chapter but a book of their own.

The focus on strategy is appropriate because strategy has primacy over tactical issues. The choice of research question, goal, approach, and design (which are all covered in the following chapters) are all made before the choice of data collection and analysis technique, which is also more easy to adapt in the course of a project.

Often, asking a sharp question and linking it with an appropriate research design will make data analysis a breeze. When data analysis is difficult, it is frequently because the design and its implementation have been weak (resulting in noncompliance, missing observations, unbalanced comparisons, wrong level of analysis, noisy measures, selection bias, and so on). While you can always master a specific data-analytic technique once you have collected the data, learning the lessons of research design at that stage would often mean having to start all over again.

Who is this text for?

The primary audience for this text is *students* of political science. As noted, political science is broadly conceived to include the study of public policy and administration, as well as international relations.

The second target group is *professionals* who need to use or commission new research in their daily line of work. For example, policy makers might need an impact assessment of a proposed regulation, government officials might need to know what the public thinks about a social problem, party functionaries might need advice on campaign strategies, and political commentators might need an election forecast. Of course, all these assessments, surveys, advice, and models would be products of research and as such subject to the same requirements, challenges, and limitations as an academic research article or a graduate student paper. Studying this text one can learn why randomized controlled trials would usually provide a better impact assessment than ex post observations of an enacted regulation (Chapter 7), why convenience samples might provide a distorted view of what the public thinks (Chapter 5), how theoretical modelling can inform strategic choices (Chapter 3), and how a forecast is different from an explanation (Chapter 6).

But making decisions informed by scientific research is not a privilege and a requirement for public officials only. I like to think that all citizens, no matter what their educational background or profession, can benefit from some knowledge of the way (social) science works, if only to appreciate its limitations. Those of us lucky enough to live in liberal democratic societies can freely debate and make proposals for the improvement of our political institutions and public policies. A lot of these ideas, debates, and proposals could, and should, be informed by results of scientific research to a much greater degree than is currently the case. The increasing availability of free open political data and the spread of the internet mean that even citizens in non-democratic states

can consult, make use of, and build informed opinions on the basis of facts and scientific evidence.

Political science is no rocket science. Its results, when properly communicated, should be accessible to any intelligent citizen with a minimum of substantive background in the social sciences and some awareness of the design issues involved in the practice of research.

Moreover, as people living in the information society, we are all bombarded with journalistic interpretations of scientific results about the effects of new drugs, cosmetics, health regimes, and exercise routines; of educational interventions, parenting styles, and student evaluations; of macroeconomic policies, gun control laws, and foreign policy interventions; of genes, history, and culture. Often, the journalistic interpretations exaggerate, misrepresent, or fail to point to critical assumptions of the underlying studies. Although this is a text about political science research, one can take many valuable lessons about how to critically evaluate scientific claims coming from a wide range of disciplines. The fundamental problem of causal inference and the more specific issues of random variation, unrepresentative samples, underpowered studies, confounding, self-selection, reversed causality, measurement validity, generalization, and more, are similar no matter whether the context is research on voting, international conflict, and bureaucratic delegation, or epidemiology, social psychology, and economics. The critical thinking skills that this text is designed to sharpen have wider applicability than political science research proper.

There is very little prior background assumed for understanding this text. If you have had a basic course in research methods that introduced concepts such as variable, sample, population, and the like, you will probably find it somewhat easier to think about research design in these terms than if you encounter them for the first time here. And some familiarity with philosophy would certainly ease comprehending a few themes (such as the nature of causality or counterfactual reasoning). But none of this is strictly speaking required, because I try to explain potentially unfamiliar concepts and ideas along the way.

Though the chapters link the foundations of research design to broader philosophical concerns about how we can know anything and what it means to explain, our aim always remains pragmatic: to derive useful lessons and clear some ground for the applied political scientist, rather than to resolve or even comprehensively present the inevitably knotty philosophical debates involved.

This text stops where the collection and analysis of data begin (and picks up again when the results from the analysis are clear and need to be communicated). One cannot learn how to chart a plot, run a logistic regression, perform a qualitative comparative analysis, or conduct an effective elite interview from these chapters. Again, the reason is that each data collection and analysis technique deserves a book of its own.

Nevertheless, it is quite impossible to understand the rationale of different research designs unless one has at least a rough idea of the subsequent step of data analysis. Many research design decisions are taken in view of the requirements, limitations, or assumptions of different data-analytic techniques. You need to be aware that certain problems of design, such as missing data or censored observations, can be addressed during the stage of data analysis and do not have to mean the end of a project. For these reasons and more, I introduce in very broad strokes how analysis of experimental, large-N, comparative, and within-case data proceeds in the respective chapters. In sum, this is not a text about research *methods* as such, but about research *design* proper.

The Way Ahead

There are two different ways to work with this text. One is to first read the general chapters (1–4, 6, and 12) and then, if you know what you need, focus on a specific class of designs (see below). The other is to follow the order of the chapters to get acquainted with all the options before you settle for one, and then go back to study it in detail. In any case, here is a brief summary of what each chapter deals with.

This chapter explained what research design is and argued for its importance. Chapter 2 delves further into the differences between normative and positive; theoretical and empirical; descriptive, predictive, and explanatory; and theory-building, theory-applying, and theory-testing research. These distinctions are fundamental, as all other decisions about design are guided and constrained by the type and objective of the research project. The chapter also looks into the more practical side of choosing and sharpening a research question once the overall topic, type, and objective of the research have been settled.

Whatever the type of research, theory has a role to play. Chapter 3 deals with the place and functions of theory in the research process. It clarifies the structure of explanatory theories and presents a few different ways in which theory can be developed deductively. The chapter discusses the criteria that good theories should meet and offers some practical considerations about choosing, developing, and using theory in political science research.

Chapter 4 explains how you get from theoretical ideas to empirical observations. It introduces the idea of a 'concept' and shows how to translate abstract concepts into observable and measurable factors and variables via the processes of conceptualization and operationalization. Various criteria and challenges for fruitful conceptualization and valid operationalization are introduced and critically discussed.

Chapter 5 focuses on measurement and descriptive research. Description exists in various modes, and the chapter presents in some detail

several quite different ways of doing descriptive research in political science. First, quantitative description via surveys is discussed. Since surveys and other sources of quantitative data play an important role in contemporary political science research, the chapter outlines some common methods and tools for the efficient analysis and presentation of quantitative empirical data. Second, ethnographic thick description via participant observation is presented. Third, the peculiarities of historical description based on archival and other documentary sources are noted.

While description is indispensable, explanation remains a primary goal of scientific research. Therefore, the focus of the remaining chapters is on explanatory theoretically informed empirical research. Chapter 6 prepares the ground for understanding the various research designs for causal inference presented in Chapters 7–11. It introduces the notions of mechanistic causal explanation and counterfactual causality. These allow us to express clearly the fundamental challenge of causal inference and to outline some potential solutions pursued in the remainder of the text.

The first class of designs for causal inference, presented in Chapter 7, are experimental designs, not only because of their growing importance for political science but also because they show very clearly what the problems of making causal claims are. The chapter introduces the logic of experimental research and explains why random assignment is so powerful. The basic design elements of experiments are also covered, including choosing the level of analysis and the sample size. Furthermore, various complications that arise in practice are introduced together with some ways of dealing with them. The chapter presents in detail several real applications of (different types of) experiments in political science research and concludes by discussing the limitations of the experimental approach.

Chapter 8 is the first of several to deal with observational designs and discusses large-N quantitative research for causal inference. Four distinct approaches to identifying causality in large-N research are presented – natural experiments, instrumental variables, mediation, and conditioning. In addition, the chapter deals with the issue of estimating causal effects (and the related problem of statistical significance) and goes on to discuss design questions such as choosing the level of analysis and case selection.

Chapter 9 considers comparative research approaches when only a few cases are available for observation. It argues that comparative designs are a hybrid, mixing cross-case logic and within-case analysis. The most prominent strategies for case selection in comparative designs are discussed and the extension to qualitative comparative analysis and fuzzy sets is made.

Chapter 10 is about single-case studies. Within-case analysis relies on a different rationale to derive proper explanatory accounts, and the