Political Science Research Methods in Action

Edited by Michael Bruter and Martin Lodge

ecpr

palgrave

Research Methods Series

General Editors: **Bernhard Kittel,** Professor of Social Science Methodology, Department of Social Sciences, **Carl von Ossietzky,** Universität Oldenburg, Germany, and **Benoît Rihoux,** Professor of Political Science, Université catholique de Louvain (UCL), Belgium

In association with the European Consortium for Political Research (ECPR), Palgrave Macmillan is delighted to announce the launch of a new book series dedicated to producing cutting-edge titles in Research Methods. While political science currently tends to import methods developed in neighbouring disciplines, the series contributes to developing a methodological apparatus focusing on those methods which are appropriate in dealing with the specific research problems of the discipline.

The series provides students and scholars with state-of-the-art scholarship on methodology, methods and techniques. It comprises innovative and intellectually rigorous monographs and edited collections which bridge schools of thought and cross the boundaries of conventional approaches. The series covers both empirical-analytical and interpretive approaches, micro and macro studies, and quantitative and qualitative methods.

Titles include:

Joachim Blatter and Markus Haverland
DESIGNING CASE STUDIES
Explanatory Approaches in Small-N Research

Alexander Bogner, Beate Littig and Wolfgang Menz (*editors*)
INTERVIEWING EXPERTS

Michael Bruter and Martin Lodge (*editors*)
POLITICAL SCIENCE RESEARCH METHODS IN ACTION

Bernhard Kittel, Wolfgang J. Luhan and Rebecca B. Morton (*editors*)
EXPERIMENTAL POLITICAL SCIENCE
Principles and Practices

Audie Klotz and Deepa Prakash (*editors*)
QUALITATIVE METHODS IN INTERNATIONAL RELATIONS
A Pluralist Guide

Lane Kenworthy and Alexander Hicks (*editors*)
METHOD AND SUBSTANCE IN MACROCOMPARATIVE ANALYSIS

Ingo Rohlfing
CASE STUDIES AND CAUSAL INFERENCE
An Integrative Framework

Political Science Research Methods in Action

Edited by

Michael Bruter

Reader in European Political Science, Department of Government, London School of Economics and Political Science, UK

and

Martin Lodge

Professor of Political Science and Public Policy, Department of Government, London School of Economics and Political Science, UK

Contents

Tables and Figures

Contributors

Jessica Bain is Lecturer in Media and Communication in the Department of Media and Communication, University of Leicester, UK.

Michael Bruter is Reader in European Political Science in the Department of Government, London School of Economics and Political Science, UK.

Natalia Chaban is Associate Professor and Jean Monnet Chair in European Identity and Culture at the National Centre for Research on Europe, University of Canterbury, Christchurch, New Zealand.

Robert Erikson is Professor of Political Science at Columbia University, US.

Mark N. Franklin is Director of the European Union Democracy Observatory at the European University Institute, Italy, where he previously held the Stein Rokkan Chair of Comparative Politics.

Julie Gervais is Lecturer in Political Science at University Paris I – Panthéon-Sorbonne, France.

Sarah Harrison is Research Officer in Electoral Psychology at the London School of Economics, UK.

Will Jennings is Reader in Politics and International Relations at the University of Southampton, UK.

Martin Lodge is Professor of Political Science and Public Policy in the Department of Government and the Centre for Analysis of Risk and Regulation, London School of Economics and Political Science, UK.

Milton Lodge is Distinguished University Professor in the Department of Political Science, SUNY at Stony Brook, US.

Maja Renko is a researcher at the European University Institute, Italy.

Aaron Strauss is a PhD candidate at Princeton University, US.

Brad Verhulst is postdoctoral fellow at the Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University.

Ben Wellings is Convenor of European Studies at the School of Politics and International Relations, Australian National University.

1
Introduction: Writer's Block

Michael Bruter and Martin Lodge

A social scientist's writer's block is not the same thing as a novelist's writer's block. 'I just can't find an interesting enough idea for a book'; that is a novelist's writer's block. 'I've got a hell of a good idea but how on earth am I going to transform it into half as good a research project?'; that, by contrast, is a social scientist's writer's block.

It is true that the work of a social scientist is highly creative, which is one of the reasons why we all find it so stimulating. It is also true that you cannot do research without, first and foremost, a good idea. Nevertheless, the truth of the matter is that most of our time is not spent in the luminous glory of the drawing room but back in the darkness of the engine room. There, good ideas produce research questions, and the existing literature is dissected with about as much consideration as that of the biologist for the frog, whose corpse he is studying. There, research questions generate hypotheses, and the data that already exist or could be gathered to test these hypotheses are weighted, evaluated, and compared with more precision than fantasy. That is our engine room: a place which may not be glamorous but is certainly essential and where, not unlike craftspeople, we continuously work on the machine, which, in turn, will enable us to produce research. Without good and hard work in the engine room, our great idea will never become good research, and nobody may even realize that there, once laid, a good idea. In fact, it is in the engine room, where we construct, shape, and reshape this ad hoc 'research machine', in fact, that great ideas become social science.

In its full complexity, the construction of that machine – which is, incidentally, also known as a research design – is never fully apparent to those who consume our research. Of course, any rigorous social science book or article provides numerous details as to the choices that were ultimately made – the questionnaires that we created, the way in which we

conducted our interviews, or the manner in which we coded discourses and manifestos. But that is really only what the machine looks like at the end, and it does not say anything about the way it was built, what choices were made, and why. How a research design is put together to tackle a specific research question, how we decide, in the engine room, how it will be powered and how not, where to put the screws, and what it should be able to resist, as well as how to realize when things are not going well and what to do about it, is what this book is precisely concerned with.

The principal aim of this volume is to offer an insight into the engine room of select research projects. As such, therefore, it is in between a grandmother's recipe book and the academic equivalent of Joseph Roth's *Confession of a Murderer* (1936). It is a collection of accounts, by a number of social scientists with different methodological preferences and approaches, of how they set out to do research. In their different ways, all contributors explain how they set up their research design, and how they encountered and dealt with critical choices and dilemmas. Just like any craftsperson, the social scientist has to choose between competing paths to optimally make something 'work'. Each chapter therefore discusses how solutions to particular problems were dealt with, and what the methodological, ethical, and analytical consequences are that should be borne in mind when evaluating the research that results from these choices. The picture that we deliver may neither be that of a high precision laboratory, nor be that of a workshop populated by dodgy cowboy builders, but it includes a series of honest descriptions, by people who are thoughtful, introspective, and passionate about the methodological choices they make as they try to be aware of their limits and potential pitfalls.

Understanding the object of the book: The Bible(s) and the Talmud (or the rules of methods vs the consequences of methods)

Theologists often contrast aspects within religions that emphasize canonical certainty on the one hand and continuous deliberation and debate on the other hand. One typical example distinguishes between the Bible (in any of its numerous Judeo-Christian versions) and the Talmud. The former is said to be defined by its certainty, its absence of room for doubt, and its focus on good and evil. The latter, by contrast, is a permanent, never-ending debate, whereby some religious authorities never reach any agreement on the correct interpretation of one or other aspect of religious rules. In its various forms, this tension between

religious prescription and religious debate has split numerous religions from Judaism to Hinduism and Protestantism. The existing literature already offers a number of excellent 'bibles' on research design. They may be either general or dedicated to specific designs. Some of the most influential books written on research design include Przeworski and Teune (1982), King, Keohane, and Verba (1994), Penning et al. (2005), Coombs (1964), and Shively (2005). To a certain extent, all of these volumes have a prescriptive value, and tell researchers what they should and should not do when preparing to conduct a piece of social science research. In fact, this perceived prescriptive value is such that at least two of these volumes are regularly nicknamed 'the bible' by their respective supporters. And as any text associated with religious status, these books have similarly received their fair share of commentary (e.g., Brady and Collier, 2004).

By contrast, this volume has no claim to be yet another 'bible'. It is, instead, very much a 'Talmud' – a perpetual discussion of the methodological, ethical, and analytical dilemmas that most social scientists come across when engineering their research designs in the face of a messy reality. It is a book which focuses more on the questions than on the answers. These few hundred pages may well disappoint those who seek categorical and universal simple solutions but will hopefully stimulate those who like to think about their own work, have ever experienced doubt about how to tackle an apparently tricky research question, and who chose the social sciences because of their complex human and social object of study and not in spite of it. Of course, the various methods and approaches presented in the following pages are all deemed by the authors to be useful solutions to difficult problems, but none is presented as a panacea or even as a problem-free option. They are not even described as 'compromises', because we are not 'bargaining' ethics, analysis, and methods. Instead, we simply think of them as dilemmas and choices. As such, they are discussed in the context of their positive and negative as well as their intended and unintended consequences, even though each of the scholars who defends them is ready to (and does) explain why they think that the particular positive–negative balance of their approach is overall preferable to its alternatives in specific contexts. This is also why this book has in many ways more to do with how to deal with problems in research design than in trying to come up with ideals of research design. This is a 'plan B', perhaps, but, we believe, an important one, because in a world where the science is in the method, how we think and talk about our methods is of fundamental value, not purely in a pedagogical sense but also in a social sense. The book tells the story of how learning takes place in so-called communities

of practice with their ongoing exchanges about experiences in researching social phenomena, rather than through the repetitive chanting of prescriptive ideals.

Yet another way of describing the specificity of our book is by explaining that it is not so much concerned with the 'rules' of methods as it is interested in the 'consequences' of methods. The contributors are not authoritatively explaining how a method should be used; they are, instead, sharing their reflections on and experience of the consequences of the methodological choices they like to make and to avoid. They provide a first-hand account of the 'secret' considerations that have led them to make their own decisions in the context of complex research projects which, in their combination, represent a rather broad spectrum of the types of research question, approaches, and methods that social scientists focus on across the range of our disciplinary subfields.

Genesis of an engine-room project

Just as the craftsperson with their tools and the materials available to them, the social scientist is faced with numerous challenges in terms of moving from an interesting idea to a working product. In this section we discuss the genesis of a research project from an engine-room perspective. We are dealing with this experience at an abstract level because the following chapters will add their individual flavour to them. We should also stress that the 'genesis' approach discussed here does not suggest that every research project follows linear paths. Far more frequently the progress of research projects requires that we step back and reconsider some previous infelicitous choice only revealed by a later dead end. Nevertheless, the following stages describe what many of us consider to be the typical natural birth process of a research project.

As we have already suggested, a good project always starts with a good idea. This is the first and irreplaceable step in the genesis of a social science research project – genesis of a research project, 'day one'. Good research starts in one's head, and no geeky knowledge of the literature and no engine fiddling will ever replace some hard thinking about what we – social scientists and citizens – need to know. Social scientists face a permanent challenge when it comes to taking enough distance to understand what our discipline has still failed to address, what is the gap that has remained unnoticed between two important advancements of our collective knowledge, what are the apparent contradictions between studies which suggest that a question has still not been taken to its end. Sometimes, finding an idea might seem daunting. At other times one seems to have far too many ideas to ever stop working. But

the coherence of this creative first step in the genesis of social science research comes from something that it is purely introspective. Nobody can find an idea for someone else, and without an idea, and preferably a good one at that, there can be no research.

With regard to this first stage, let us focus on the example of Sarah Harrison's contribution (Chapter 3). She spotted that in between works assessing the respective impacts of institutions, the economy, other parties, voters' demographics, and so on, on the success of extreme right parties, something was missing in the literature – the impact of the discursive and ideological choices of extreme right parties themselves. Looking at whether these extreme right parties' choices matched the preferences of their specific voters, and whether this relative match influences the success of these parties, is something that, paradoxically, had never been touched upon by the literature before. The 'paradoxically' is not unimportant here because the idea makes intuitive sense without being the object of academic exploration before. In that and other senses, it seemed important enough an idea to generate research.

'Day two' of the genesis, however, is the transformation of an idea into a research question. This stage is no less indispensable than the existence of the good idea, which serves as its prerequisite. While the original idea for research may be a theme, or an intuition, the research question is a specifically constructed line of investigation, a workable question, which will be directly, fully, and specifically answered by the research work that will be created for it. A research question has to be set out transparently and explicitly, and any book, article, or thesis will be purely dedicated to answering it, fully and exclusively. It is at the end of this second stage of engine-room proceedings that the social science sketch becomes a recognizable project, and that research is truly on its way as a liveable enterprise. Suddenly, fascination gives way to hard analytical thinking, and creativity to creation.

When it comes to understanding the transformation of a research idea into a research question, let us consider Michael Bruter's contribution (Chapter 2). His idea was to study the emergence of a mass European identity. However, because of the history of identity building and the importance of identities as markers of interaction between citizens and political institutions, this led him to focus on a far more specific research question: Can political institutions, via the generation of political system symbols, and the mass media, via the way they inform us, encourage or impede the emergence of a new mass European identity? The idea was an important theme that needed social science attention, but the research question was a specific line of investigation

which could be answered by a research design and a book. In other words, within the framework of this particular phrasing of the research question, the author could formulate a set of hypotheses that would theoretically answer the question, and create empirical tests that would assess whether these hypotheses were correct.

On 'day three' of the genesis, the social scientist spends a month in the library trying to get a pretty-near-exhaustive knowledge of the relevant literature – in many ways, what architects would call 'exploring the terrain'. One of the difficulties of this particular step is to understand what exactly constitutes the terrain. For instance, if a social scientist decides to study democratization in Russia, the terrain will certainly not be limited to the literature on Russia, as theories used in other contexts and countries are the most likely to help the social scientist to build a novel model adapted to his own case study. Similarly, a study of the motivations of anti-globalization protesters will imply a need to cover the literatures on political protest and political participation more generally, or an enquiry into the determinants of tax policy may require a knowledge and understanding of the literatures on neighbouring public policies as well as political economy and bureaucracy. Unfortunately, there is obviously a fine line between missing out on some essential and relevant literature and comparing 'apples and oranges'. That is why the process of reviewing the existing literature is not just one of digesting existing accounts but also an interactive (if partly 'virtual') discussion between the researcher and his academic environment.

An excellent example is provided by the Mark Franklin and Maja Renko's contribution (Chapter 5). The emergence of the Propensity to Vote (PTV) solution was only conceivable on the basis of a very comprehensive understanding of the electoral behaviour literature, looking at both the item phrasing in existing questionnaires and the apparent substantial limits or paradoxes of existing findings about voters' choice. It is on that basis that Franklin and Renko could provide this new solution to an old but not fully understood problem affecting an entire body of literature.

On 'day four' of the genesis of a research project, the social scientist will normally develop a model, which can answer the proposed research question theoretically and will be enriched by the literature review that was conducted previously. Indeed, the literature review stage of research project engineering is useful not only to understand what we already know and to avoid replicating existing knowledge but also to use others' insights to develop the theoretical and analytical model that we will want to test in a given project. A model is very much a simplified

version of how we believe the world works. It usually takes the form of a series of testable and falsifiable hypotheses, which the research project will confront to determine what we can empirically find out about the real world. 'Testability' means that there exists empirical evidence which could be used to verify whether a hypothesis is indeed upheld in reality (as opposed to something that we could never check in practice); 'falsifiability' means that the hypothesis is phrased in such a way that if it is 'wrong' we should also be able to find that out (as opposed to a phrasing which could never be disproved because it is too vague or too general).

Here, a typical example is contributed by Robert Erikson, Aaron Strauss, and Michael Bruter (Chapter 11) who use pure modelling and simulation to (re)frame the problem of understanding how parties choose their election manifestos. The authors realize that traditional expectations – not least that of convergence under Downsian assumptions – are not actually verified in real life. They also note that existing research had not yet been able to come to terms with these paradoxes. They thus propose to model the choice of their manifestos by two political parties competing in a given party system. They consider the variables that are expected to influence this choice, such as a party's relative preference for vote-seeking or policy-seeking priorities, what we know of an electorate's true policy preference, which party is expected to win and by how much, and how the other party is framing its own manifesto. They then model the way in which these factors may interact for each of the two parties competing in the election. This leads them to a set of mathematical simulations which help us to understand in theoretically meaningful terms under what conditions we expect the two parties to find an equilibrium, and under which circumstances they will not, chasing each other continuously instead. In this case, the model is run purely mathematically without any empirical test, but, ultimately, all social science models hope for an adequate test to be designed and applied.

The design of an appropriate test is precisely what normally constitutes 'day five' of the research enterprise – it is the equivalent of choosing tools for the job in hand. Now that we know what theoretical propositions need to be assessed, how can we choose the most appropriate empirical tests to do so? Can our research question be answered equally well using a quantitative or a qualitative research design? What is the most appropriate empirical test for the model that we propose? What data are available or may realistically be collected, and are they relevant to the way in which the question and hypotheses are framed?

For instance, a research question which has to do with a party's ideology may rely on an analysis (quantitative, qualitative, or both) of party manifestos (Bruter and Harrison, 2011), or on interviews of party leaders, or again on evaluations of the party's policy either by experts (de Lange, 2007) or by the public at large via mass surveys. However, these different types of possible empirical test will have various consequences for what specific questions can be answered and also, of course, the type of bias – or systematic error – that may contaminate the results. For instance, relying on public perceptions might mean that we could measure a population's prejudice as much as (or conceivably more than) the reality of a party's ideology. Conversely, relying on interviews of party leaders may mean that we capture a set of ideological and policy perceptions which may not be shared by the party members, only by their leaders. In this sense the choice of a methodological protocol is a crucial aspect of the craftsperson's work in the engine room. Indeed, it is not only the stage when the social scientist will perform a background census of the methodological options that exist in theory (How could we answer this question?) and in practice (What is the data potentially available to me in real life to answer this question?), but also when he will comparatively 'weigh' the advantages and disadvantages of these various realistic options. In short, this is the main 'engine room' stage of the research project, and the one when the research machine is constructed.

All of the chapters of this book are direct examples of this 'research machine' building. So we may also note the variety of research machines engineered across the breadth of research examples that we provide. They include a whole set of approaches and methods from the analysis of manifestos to that of TV news in Jessica Bain and Natalia Chaban's contribution (Chapter 7), to Julie Gervais' analysis (Chapter 8) 'inside' the secretive world of a French *corps*.

'Day six' and 'day seven' of the genesis of a social science research project are not directly relevant to the engine room. Indeed, by the time they occur, the research machine has been created and is already in full operation. These two stages represent the analysis of the data, and the interpretation and discussion of the findings. While they are not part of the engine-room design (at this stage it is too late to try to change most of the research machine), it is important to remember that the machine that was built earlier will influence what the researcher will see as their findings, and the way in which these will thus be interpreted and discussed. Therefore these are also stages where many high hopes get dashed, and where we often must go back and reconsider our choices at earlier stages. There, with very limited leeway (one simply

cannot redo a new survey instead of the one that was designed and run, or throw 200 interviews in the bin to restart them from scratch), the social scientist has to consider alternative paths to solving the problem. This is why many of us, whenever possible, choose to 'pilot' the fifth stage of the project's genesis to give ourselves a chance to observe and amend the working of the research-machine prototype before making it work at full speed. This sometimes difficult, and never truly straightforward, transition from setting up the engine to making it run smoothly illustrates the purpose of our book. Regardless of whether we find some solutions 'better' than others, analytical and methodological choices have consequences, and almost any one of the approaches that we can pick from will enlighten certain areas of our research question but – even more importantly – leave others in the shade. Indeed, any project entails some implicit and explicit analytical, methodological, and ethical choices with critical implications for the nature and findings of our research.

Three types of consideration: Ethical, methodological, and analytical issues

If the types of problem that a social scientist will face and try to evaluate in the engine room are numerous, they are, broadly speaking, most likely to fall into one of three types of fundamental issue: ethics, methods, and analysis. And, unfortunately, none of them is likely to be simple or straightforward. They arguably occur in different degrees and incarnations across all areas of social science research.

Ethical concerns

Ethical concerns are the cardinal reference of a social scientist's work. They represent the law and morality of the social sciences in a way which should, at least in theory, be non-negotiable. Social scientists do not just work for themselves. They have a number of duties towards their colleagues, their human subjects (i.e., the people who agree to take part in their data collection), and the public and society at large. Ethical concerns broadly define or emphasize these multiple duties and obligations. As a result, ethical references are varied and may address the relationship between the researcher and their work, their human subjects, their colleagues, society as a whole, and many others. For instance, they should not cheat or lie about their findings; they should not expose human subjects to any experimental treatment or interview without their express and informed consent; they should not plagiarize someone else's work; they should protect the anonymous character of the

data collected and take the necessary steps to prevent others to accessing it; and so on.

A certain number of ethical concerns may also seem to vary with fashions or across countries. For instance, virtually all ethics committees in the US will object to a participant's incentive in the form of a drink voucher that could be used in exchange for an alcoholic drink, whilst this same possible incentive is unlikely to raise many eyebrows anywhere in Europe. Finally, a certain number of things which should presumably be conceived as ethical concerns are systematically and happily ignored by our research community. For instance, we know perfectly well that an immense majority of citizens resent being called at home at dinner time by a company wanting to carry out a survey, and that they view it as a gross and unacceptable invasion of their privacy. Yet we continue to rely on survey companies doing exactly that for the purposes of our research. Because we ideally want genuinely random samples, we even prefer the said companies to dial numbers randomly, which implies that we do not even respect national legislation protecting citizens against unwanted calls via various lists ('red list', 'turquoise list', etc., depending on the country and specific scheme of the list).

Thus, altogether, in the context of any research enterprise, the social scientist will firstly have to continuously question their duty towards fellow social scientists, human subjects and participants, the law, and society at large, and to try to minimize their ability to 'disturb', betray, or mistreat them. Secondly, they will also try to understand which remaining aspects of the project still present ethical risks, and to compensate for them in the best possible ways. For instance, the effect of a research treatment will be partly compensated by the use of appropriate debriefing; using other people's ideas will be made acceptable (and even commendable) by properly referencing them; and the use of confidential data will necessitate taking every possible step to protect the said data. In many countries, legislation on research ethics and the protection of private life requires social scientists to immediately destroy confidential data. At the same time, however, other ethical guidelines require researchers to make it possible for people to replicate their work, which implies the retention of the data in question. What is more, whilst our ethical responsibility towards individuals may have been deemed by legislators to entail the destruction of records, is not our ethical duty towards civilizations and communities to protect what may be the best chance of transmission of human oral history? Here, again, we see how complex 'simple' ethical guidelines can in fact be, and how an apparently 'common-sense' measure could, if taken literally,

deprive humanity of, for example, the testimony of the last survivors of Nazi death camps.

Methodological concerns

Methodological concerns are of an entirely different nature. They broadly have to do with the need to organize research protocols in such a way as to get observed results that fit reality in the most accurate and least ambiguous way. In the social sciences, to analyse 'reality' we end up selecting it, sampling it, measuring it, and analysing relationships between various elements of it. Each of these stages introduces error into what we do, and this error drags what we observe and describe in research a little bit further away from the 'true' reality of the world that we ideally want to capture.

This error can be of two forms: random or systematic. Random error is like a haze which blurs our vision of the reality we would like to capture. Its random impact makes tendencies look less clear (or more ambiguous) than they really should. This is roughly the equivalent of a writer introducing a number of typographic mistakes into the text that they type, occasionally causing confusion or hesitation as to what they intended to write in the first place and forcing more caution than one would ideally want when trying to understand their meaning. By contrast, systematic error introduces a bias that drags what we observe in a distinct direction away from the 'true' reality we were hoping to capture. This time it is more as though a publisher introduced a chapter by the wrong author into a book, or attributed the book to the wrong author altogether. Importantly, understanding that there is an error is no longer obvious (this 'wrong' book will be internally consistent) and requires a greater critical insight as to what may have gone wrong in the study we have crafted.

Within this context, the two main methodological challenges faced by the social scientist consist of, firstly, trying to minimize this error, and particularly its systematic component, and secondly, understanding as well as possible what error is polluting our empirical analysis, this time again with particular attention paid to systematic error. Even more so than ethical concerns, methodological problems can sometimes be difficult to compensate for after they have occurred, and in many cases the 'error understanding' stage will be the final one.

Of course, in principle, particularly in the context of quantitative models, great numbers of methodological tools exist that 'compensate for error'. Most have to do with a 're-creation' or simulation of a dataset which would be 'errorless'. Most such techniques assume that untrue

cases can be identified 'on paper', and usually, they are simply outliers. The possibility that these outliers may be real cases and simply correspond to a problem with the model used to explain the data, rather than with the data themselves, is thus excluded for better or for worse. This is also the reason why a majority of 'error solving' mechanisms result in improved coefficients because they simply improve the fit of the data to the model (rather than the contrary), if at the cost of robustness. For all of these reasons, and without entering methodological discussions on the specific techniques used by colleagues to mechanically 'solve error', we will simply point out that whether one believes that these tools are a panacea, or that they are a cheapskate, they do not change anything of the fact that we need to understand what is the 'real' error in our data and models if only to get a sense of whether outliers are an anomaly or a reality that we have failed to account for in satisfactory ways. In qualitative contexts, the equivalent of this diagnosis and compensation of error may be more cumbersome. The social scientist can 'spot' outliers and ultimately decide to not include them in their sample if there seems to be good reason to consider them incomparable with the rest (e.g., a researcher conducting interviews on perceptions of politicians amongst citizens may decide to dismiss the interview of a respondent who happens to be a politician themselves or, say, the husband or son of one). But whether or not this exclusion is methodologically justified is open to debate.

What is more, error is not only a problem that can be diagnosed empirically. Possibly the worst form of systematic error of all, precisely because it cannot be diagnosed internally, is the lack of validity of the measurement of a variable. Some authors claim that lack of validity is nothing more than a problem of systematic error which can thus be diagnosed statistically in a quantitative model. This is wrong, at least unless there is also the presence of a known perfectly valid variable in the same dataset, which is rarely the case in the social sciences. In the absence of this 'totally safe' standard of reference, validity must first and foremost be understood as a conceptual, almost philosophical, reality. It represents the match between the concept that one wants to measure and that or those that are indeed captured by the items used by the scientist. Be it in a quantitative or qualitative context, this must first and foremost be evaluated intellectually and not statistically because in the absence of a trustable outside reference, a perfectly consistent measurement of the wrong variable will appear exactly as valid as the perfectly consistent measurement of the right variable. In other words, if I measure perfectly consistently the way in which citizens arbitrate between

support for civil liberties and public order and call it, say, a measure of tolerance, the only foundation of my lack of validity will reside in the fact that arbitration between civil liberties and public order on the one hand, and tolerance on the other hand, are simply not the same thing. If no one is here to make this argument, my measurement will undoubtedly appear as pure and valid as can be.

When it comes to the question of validity, if there is no ultimate salute in statistics, there is no ultimate salute in the literature either. Science, in all its branches and meanings, is full of stories of disciplines spending several centuries looking at questions from the wrong end of the lens until someone suddenly figures out that a change of paradigm is necessary to solve the inconsistencies and inaccuracies others had disregarded till then. The ultimate argument between classic physicists behind Einstein and quantum physicists following Bohr will probably remain one of the most inspiring examples to date of the possibility for a discipline to live parallel lives around both sides of a traumatic schism. There is obviously no reason to assume that, as social scientists, we are not guilty of the kind of disagreements that have regularly divided doctors, physicists, and biologists. Most of them will be due to widely overlooked problems of validity in the measurement of the number of variables we use every day. For instance, from that point of view, the potential consequences of the use of PTVs to measure party preferences as per Mark Franklin and Maja Renko's contribution (Chapter 5) are immense in that they could invalidate entire segments of the literature on which our understanding of electoral behaviour is built. Similarly, the contribution from Robert Erikson, Aaron Strauss, and Michael Bruter (Chapter 11) shows how one can sometimes need to take a model to its ultimate limits to assess the extent to which some of the universally accepted assumptions we make to study the world in which we live may not be tenable in the first place.

Altogether, methodological concerns will thus primarily rest on a need to minimize error and to understand it, as well as on a duty to think about the conceptual and philosophical soundness of the adequacy of our measurement and operationalization (be it quantitative or qualitative) for the needs and foundations of our conceptual model.

Analytical concerns

This naturally leads us to the last type of problem which we must think about when arbitrating between several alternatives of research design, analytical concerns. The question of validity already points to the natural border between methodological and analytical queries, but the latter

have a far-reaching impact when it comes to the logic that the social scientist uses to reduce an un-apprehensible world into a simplified universe that can be studied empirically. Modelling is very much a problem of 'sorting things out'. It consists of an essential preliminary (i.e., technically, treated as 'exogenous') and what needs to be paid attention to (i.e., treated as 'endogenous'). It also means understanding how the relationship between these endogenous variables must be logically organized in the real world and, therefore, in our simplified model.

Here, again, there is no obvious diagnosis for analytical problems even though their theoretical and empirical consequences can be immense. One of the most usual criteria to use when it comes to assessing the soundness of an analytical model is that of 'model specification'. The notion of model specification aims to assess the extent to which a proposed model includes (i) all the variables that should be included in a model; (ii) no variable that would not have its place in the said model; and (iii) all of the causal links that should be expected to be true.

Forgetting some variables which should be included in a model results in the misspecification of the model and will cause bias – for instance, in the form of the observation of some spurious relationships.

A traditional example of the occurrence of such a problem has to do with the study of the impact of race on electoral turnout. When evaluating the impact of race and ethnicity on the likeliness to turn out in elections in the US, political scientists long thought that the 'obvious' way to study this question was to look at the National Election Study (the main survey on national elections in the US), create a dummy variable separating 'white' people from 'non-white' people, create another variable measuring turnout in recent elections, and look at the link between the two variables on the second – for example, by using a simple test of difference of means, such as analysis of variance (ANOVA). When this technique is applied, such a test will invariably suggest that people from ethnic minorities ('non-whites') are less likely to vote than white people, on average. However, when using the frequency of vote in recent elections as the dependent variable in a multivariate regression where, besides including the 'non-white' independent variable, one also controls for a number of other factors, such as the respondent's level of education, the researcher will reach dramatically different conclusions. Indeed, invariably, in such models where education is controlled for, we find that with a comparable level of education, non-white voters are actually more likely (and not less) to participate in elections.

In short, the type of research machine we use in this example will lead us to draw diametrically opposite conclusions – the bivariate machine

will suggest that non-white voters are less likely to vote, while the multivariate machine will suggest that members of ethnic minorities are more likely to vote, everything else being equal (i.e., if we control for as essential a variable as education). Needless to say that the discussion that will follow – for instance, when it comes to policy consequences – will also be radically different. Moreover, the conclusions that would be drawn from the research would also be different if we split the non-white category into more specific groups, and would be different again if one chose to use qualitative methods (such as interviews or a focus group) instead of quantitative ones.

A generalization of this 'race and turnout' example is that a first rule of thumb of model specification is that any variable which is distinct from both one of the independent variables of interest and the dependent variable, and correlated to them both, should be included in the model. Forgetting a variable which fits both of these categories will lead to misspecification and bias, while including a variable which would in fact not be distinct from one of the other independent variables will lead to severe multicollinearity and a lack of stability of the model. Finally, including variables which are entirely exogenous to the model will lead to a lack of parsimony of the model.

The problem of misspecification has very different implications in quantitative and qualitative models, as well as different diagnoses. Quantitatively, forgetting an endogenous term can obviously only be realized when one thinks of reintegrating it into a model (which is tantamount to saying that if one is in doubt about whether a variable should be included in the model or not, this is easily testable, and problems will only occur, once more, when theoretically one failed to consider the possible need to include that variable). Similarly, multicollinearity is easy to test for and to correct. By contrast, in a qualitative context, the consequences of problems of misspecification are very similar, in some ways, to problems of validity. They will mostly result in causal effects being attributed to the 'wrong' independent variables in ways that will only be discovered when someone else makes an analytical argument against the social scientist's (mis)interpretation of his own model.

The same discrepancy will, to an extent, occur when trying to understand the direction of causal effects. Ultimately, causality is one of the most difficult problems to sort out in terms of the analytical elaboration of our models. Does being insulted by members of an ethnic minority make you more likely to become racist, or does being a racist make you more likely to be insulted by members of an ethnic minority? Both probably, and in many other social science problems, excluding any one

direction of causality will often be hard in principle. Establishing which of two causal links is likely to be dominant will largely depend on the strength of the theoretical arguments that can be made. To an extent, a number of techniques make it possible to assess the dominant causal link between two variables in a quantitative context. Some of these techniques are based on a goodness of fit model, and are subject to all of the other criticisms that can be, and are, addressed to goodness of fit models in general in terms of their inductivity and their relative lack of robustness. Meanwhile, others imply a recourse to time series models, which, in a way, allow us to arbitrate between the prehistoric role of the chicken and the egg in complex social science models. In qualitative models, the same use of dynamic models could be conceived in theory, although, in practice, they are very difficult to implement in the absence of dynamic measurement. Either way, here, again, the prime answer to causality issues has to be theoretical and analytical, and even more so when it comes to deciding whether it is better to 'simplify' causality by relying on a dominant causal path or to respect the natural complexity of a process by drawing a reciprocal causal path between two variables in a model, with all of the problems that may arise in terms of endogeneity (amongst others).

Overall, analytical problems will therefore have something to do with the way in which we are willing to model the world. Social scientists are faced with a permanent dilemma which best summarizes the difficult strategies of model specification and causal inference. Where is the 'just balance' between being too assertive on the one hand and including too many things (variables or causal links) in our analysis on the other hand? Do we prefer the risk of bluntly asserting the wrong thing and generating bias in our analyses, or that of not only losing parsimony but also introducing multicollinearity or preventing ourselves from using the most appropriate techniques of analysis for a given problem?

These modelling dilemmas are at the heart of analytical problems, where they add an empirical dimension to the more general issue of relying on a deficient theoretical framework which would miss out on some crucial links and explanations. Thus analytical problems are perhaps even more difficult to evaluate than ethical issues and methodological ones. Beyond the strict realm of quantitative 'warning signals' of model misspecification, there is indeed no clear or universal way to realize that one is entering a zone of analytical turbulence or failing to provide the right explanations for a puzzling phenomenon. Here, again, the human brain cannot easily be replaced by a machine (let alone happily so), and

it will usually simply take more introspection – or, more worryingly, a mean referee – to reveal the inadequacy of a research project.

Coping with trade-offs?

In an ideal world, a social scientist would be able to sort out his ethical, methodological, and analytical dilemma either all at once or, failing that, individually, as if they were discrete. Ultimately, however, the various ethical, methodological, and analytical problems that we face tend to have a bad habit of not being solvable by the same solution. Worse, in many cases, there is no other clear way to answer one problem than by generating another (of a same or different order) so that social science research often becomes a recurrent choice between pest and cholera.

The examples of the contradictions that we face between various ethical, analytical, and methodological dilemmas (if not polylemmas) are numerous. Some are obvious and others are not, proceeding from contradictions long buried by the social science literature or that as a series of disciplines we have traditionally preferred to overlook as irrelevant, which is not to say that they do not constitute problems.

Let us think of some examples that illustrate the extreme complexity of some of the problems that we face. In theory, factor analysis is a model testing technique, which tells us to what extent variance in individual items is explained by variance in a latent variable which corresponds to a concept assumed to exist but not independently measured. However, progressively, psychometricians have started to recognize it as one of the best ways to approximate the measurement of this latent variable. In this way, if several items (e.g., the extent to which you would mind having a neighbour from another race, the extent to which you would mind your daughter marrying someone from another race, and the extent to which you believe that people from different races have different intellectual capacities) are expected to covary and their shared variance can be hypothesized to represent the latent concept (say racism) which we would ideally want to measure, the factor created by a single factor solution of the factor analysis of all of these items provides us with an excellent estimate of this latent concept. What is more, this measure will have been 'cleaned' of the irrelevant variance (e.g., the fact that you answered 'a lot' to the marriage question because you hope your daughter will become a nun and would therefore resent her marrying altogether) in each of the individual items, whilst other indexing techniques, such as a mean index, would not get rid of this irrelevant variance.