PHILOSOPHY AND METHOD IN THE SOCIAL SCIENCES

ROUTLEDGE

Evidential Pluralism in the Social Sciences

Yafeng Shan and Jon Williamson

Evidential Pluralism in the Social Sciences

This volume contends that Evidential Pluralism—an account of the epistemology of causation, which maintains that in order to establish a causal claim one needs to establish the existence of a correlation and the existence of a mechanism—can be fruitfully applied to the social sciences. Through case studies in sociology, economics, political science and law, it advances new philosophical foundations for causal enquiry in the social sciences. The book provides an account of how to establish and evaluate causal claims and it offers a new way of thinking about evidence-based policy, basic social science research and mixed methods research. As such, it will appeal to scholars with interests in social science research and methodology, the philosophy of science and evidence-based policy.

Yafeng Shan is Assistant Professor of Philosophy of Science at the Hong Kong University of Science and Technology. He is the author of *Doing Integrated History and Philosophy of Science: A Case Study of the Origin of Genetics* and the editor of *New Philosophical Perspectives on Scientific Progress*.

Jon Williamson is Professor of Reasoning, Inference, and Scientific Method at the University of Kent, UK. His books include *Evaluating Evidence of Mechanisms in Medicine*, *Lectures on Inductive Logic, Probabilistic Logics and Probabilistic Networks*, *In Defence of Objective Bayesianism* and *Bayesian Nets and Causality*.

Philosophy and Method in the Social Sciences
Series Editor: Phil Hutchinson
Manchester Metropolitan University, UK

Engaging with the recent resurgence of interest in methodological and philosophical issues in the human and social sciences, this series provides an outlet for work that demonstrates both the intellectual import of philosophical and methodological debates within the social sciences and their direct relevance to questions of politics, ethics or policy. Philosophy and Method in the Social Sciences welcomes work from sociologists, geographers, philosophers, anthropologists, criminologists and political scientists with broad interest across academic disciplines, that scrutinises contemporary perspectives within the human and social sciences and explores their import for today's social questions.

Titles in this series:

The Constitution of Social Practices
Kevin McMillan

Action at a Distance
Studies in the Practicalities of Executive Management
R.J. Anderson and W.W. Sharrock

The Problem of the State
Michael Mair

Dialogical Networks
Using the Past in Contemporary Research
Ivan Leudar and Jiří Nekvapil

Evidential Pluralism in the Social Sciences
Yafeng Shan and Jon Williamson

For more information about this series, please visit: https://www.routledge.com/Philosophy-and-Method-in-the-Social-Sciences/book-series/ASHSER1373

Evidential Pluralism in the Social Sciences

Yafeng Shan and Jon Williamson

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2023
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 9780367697228 (hbk)
ISBN: 9780367697266 (pbk)
ISBN: 9781003143000 (ebk)

DOI: 10.4324/9781003143000

Typeset in Times New Roman
by ApexCoVantage, LLC

YS: To Zifei with love.

JW: To Kika, Charlie and Anouk with love.

Contents

Figures

Tables

Preface

Evidential Pluralism is a new philosophical account of causal enquiry. It espouses two kinds of pluralism. The first, *object pluralism*, holds that in order to establish a causal claim, one normally needs to establish two things: that the putative cause and effect are appropriately correlated, and that there is an appropriate mechanism complex linking the putative cause and effect which can account for the extent of the observed correlation. The second, *study pluralism*, holds that in order to assess a causal claim, one normally needs to assess relevant association studies and mechanistic studies where available. Association studies are studies that test whether the putative cause and effect are correlated, while mechanistic studies test for hypothesised features of a mechanism complex linking the putative cause and effect.

In this book, we argue that Evidential Pluralism can be fruitfully applied to the social sciences. This application is fruitful in three respects: Evidential Pluralism can explain and validate successful examples of causal enquiry; it can help us understand the general structure of causal enquiry; and it can help to inform practice in the social sciences. With regard to informing practice, Evidential Pluralism can help researchers ascertain how best to establish a causal claim of interest, and it can help evaluators assess causal claims that have been put forward by other researchers.

These arguments appeal to a mixture of philosophical theory building and close analysis of practice in the social sciences. We begin with the theory building. Evidential Pluralism was originally proposed in the context of the biomedical sciences. In Chapter 1, we develop a general account of Evidential Pluralism that is not tied to any particular domain of application. We devote some space to carefully stating Evidential Pluralism, explaining the relevant terminology and trying to avoid potential misconceptions. We also motivate the theory and sketch its application to the biomedical sciences. While Evidential Pluralism is an epistemological theory—concerned with how to establish and assess causal claims—we outline one particular philosophical theory of the nature of causality that coheres well with Evidential Pluralism.

In Chapter 2, we explore some views that might be thought of as historical precursors of Evidential Pluralism, including the approaches of Claude Bernard, W.F.R. Weldon and John Goldthorpe. We also note some differences between

Evidential Pluralism and analytic sociology, and some differences between Evidential Pluralism and Roy Bhaskar's critical realism.

The book then turns to methodological developments in the social sciences that are motivated by Evidential Pluralism. In Chapter 3, we argue that Evidential Pluralism leads to a new account of evidence-based policy assessment, which we call EBP+. EBP+ provides the capability to assess mechanistic studies alongside the association studies that are the bread and butter of present-day evidence-based policy assessment. This can lead to better-informed judgements of the effectiveness of social interventions, and thus to better social policy. We compare this new account to related approaches, such as realist evaluation.

In Chapter 4, we argue that Evidential Pluralism provides new philosophical foundations for mixed methods research in the social sciences. Mixed methods research mixes quantitative and qualitative research methods and data. After providing an account of the context within which mixed methods research emerged, we explore the question of its philosophical foundations. We sketch standard approaches to the foundations of mixed methods research and note certain limitations to these standard approaches. We then argue that the account of causal enquiry at the heart of Evidential Pluralism requires a thorough consideration of both quantitative and qualitative methods and so provides strong motivation for mixed methods research in causal enquiry. Evidential Pluralism also provides practical guidance on how to integrate quantitative and qualitative studies.

Chapter 5 responds to four potential objections to Evidential Pluralism. The first two are objections to object pluralism: an objection that establishing correlation and mechanism is not sufficient for establishing causation, and an objection that establishing correlation and mechanism is not necessary for establishing causation. Next, we consider an objection based on causal pluralism—the view that there are multiple concepts of cause in use in the social sciences. Finally, we respond to concerns about how to define mechanisms in the social sciences.

While Parts I and II of the book (Chapters 1 to 5) are pitched at a very general level so as to apply right across the social sciences, Part III (Chapters 6 to 10) seeks to address considerations that are specific to particular social sciences.

Chapter 6 illustrates the advantages of applying Evidential Pluralism to sociology. We show that Evidential Pluralism can shed light on the use of evidence in causal enquiry in sociology by means of two examples: one involving the connection between socioeconomic status and health, and one concerning the link between family background and educational attainment. We also show that Evidential Pluralism generalises and motivates certain approaches to the methodology of causal enquiry in sociology, including that of Morgan and Winship.

In Chapter 7, we turn to economics. Again, we begin by considering two examples: the link between the legalisation of abortion in the USA in the 1970s and the subsequent decline in the crime rates in the 1990s, and the link between unemployment and crime. These two examples highlight the roles of association studies and mechanistic studies in causal enquiry in economics. We argue that Evidential Pluralism can help us to understand the structure of causal enquiry in

economics, and discuss in more detail the role of mechanisms and of theory in economics. Finally, we address concerns about causal pluralism in economics.

Chapter 8 discusses Evidential Pluralism in political science. We see that Evidential Pluralism can account for the need for a diversity of methods in political science, and we examine a case study concerning the role of resources in shaping strategies of violence in rebellions. We discuss causal pluralism in political science and show how Evidential Pluralism coheres well with process tracing, which is a well entrenched method for causal enquiry in political science, as well as multi-method large-*N* qualitative analysis, which is a newly emerging approach.

Chapter 9 considers law. Law is often taken to presuppose a concept of cause that is autonomous from that used in philosophy: this is partly because causal claims in the law are constrained by legal rules about liability, while causal claims elsewhere are not. We argue that while Evidential Pluralism applies directly to our usual notion of cause, a slightly modified version of Evidential Pluralism can shed light on causation in the law. This modification requires considering what we call 'liability-tracing mechanisms' in place of regular mechanisms.

Chapter 10 suggests that Evidential Pluralism also applies to other social sciences, including anthropology, psychology, and demography, for example. Thus Evidential Pluralism has broad scope in the social sciences. We close the book by making some general comments about the overall approach and by suggesting some potential avenues for further research.

We hope that this book provides a good example of what might be called *epistemology-driven* philosophy of science. Much work in the philosophy of social science is driven by metaphysics or conceptual analysis—by views about the nature of social reality or about general concepts employed in the social sciences. That is not the only way to proceed. The approach that underpins this book is one which focuses on epistemological questions surrounding how to establish, confirm and assess causal claims in the social sciences, in order to develop a theory of causal enquiry. This account of causal enquiry, if successful, can then be viewed as a constraint on the metaphysics of causality or on an analysis of the concept of cause. If a theory of the nature or concept of cause can validate and explain a successful account of causal enquiry, so much the better for that theory. On the other hand, if the theory is incompatible with the successful account of enquiry, then it can be viewed as challenged, or undermined, by scientific practice. Exactly which theories are compatible with Evidential Pluralism is an interesting question for further research.

Acknowledgements

We are grateful to the Leverhulme Trust for supporting the research project *Evidential pluralism in the social sciences*, from which this book grew. The book has benefited enormously from discussions with the two other project members, Samuel D. Taylor and Alexandra Trofimov. We are also indebted to Graeme A. Forbes, Brian Kett, Hugh Robertson-Ritchie and Michael Wilde for very helpful comments on an earlier draft. We are grateful to Svein Anders Noer Lie for his kind hospitality in hosting a workshop on Evidential Pluralism in Tromsø, and dedicate the cover image to him and his colleagues there.

We would like to thank Neil Jordan, Gemma Rogers, Alice Salt, Phil Hutchinson and all the team at Routledge for their work in bringing this book to fruition.

Some text previously appeared in the following published works:

Shan, Yafeng, and Jon Williamson. 2021. "Applying Evidential Pluralism to the Social Sciences." *European Journal for Philosophy of Science* 11(4): 96.

Shan, Yafeng. 2022. "Philosophical Foundations of Mixed Methods Research." *Philosophy Compass* 17(1): e12804.

Part I

Philosophical Framework

1 Evidential Pluralism

§1 What is Evidential Pluralism?

Evidential Pluralism is an account of causal enquiry. Throughout this book, we take causal enquiry to include both establishing causal claims and assessing causal claims. Evidential Pluralism can be expressed as follows:

> *Evidential Pluralism.* In order to establish a causal claim, one normally needs to establish two propositions: that the putative cause and effect are appropriately correlated, and that there is some mechanism complex involving the putative cause which is responsible for the putative effect and which can account for the extent of the correlation. So, in order to assess a causal claim, one normally needs to assess relevant association studies and mechanistic studies where available.

In this section, we explain in detail what Evidential Pluralism says. In the next section, §2, we explain why Evidential Pluralism is plausible. Russo and Williamson (2007, §§1–4) put forward the thesis that establishing causation in the health sciences requires evidence of mechanisms in addition to evidence of probabilistic dependencies. Evidential Pluralism is a development of this thesis, and we shall sketch its use in the health sciences in §3. The Russo-Williamson thesis was clarified and developed by Gillies (2011); Illari (2011); Clarke et al. (2014); Gillies (2019); Williamson (2019a), amongst others. We discuss some points of clarification in §4. Evidential Pluralism is a theory about the epistemology of causality—about how we can identify causal relationships. However, Russo and Williamson (2007) used this theory to argue for a particular account of the nature of causality, namely *epistemic causality*. We introduce epistemic causality in §5. In this book, however, we do not commit to any particular account of the nature of causality—we leave this question open. This book is primarily about the epistemology of causality in the social sciences.

Let us begin by clarifying what Evidential Pluralism says. This is best done with the help of a diagram, Figure 1.1.

DOI: 10.4324/9781003143000-2

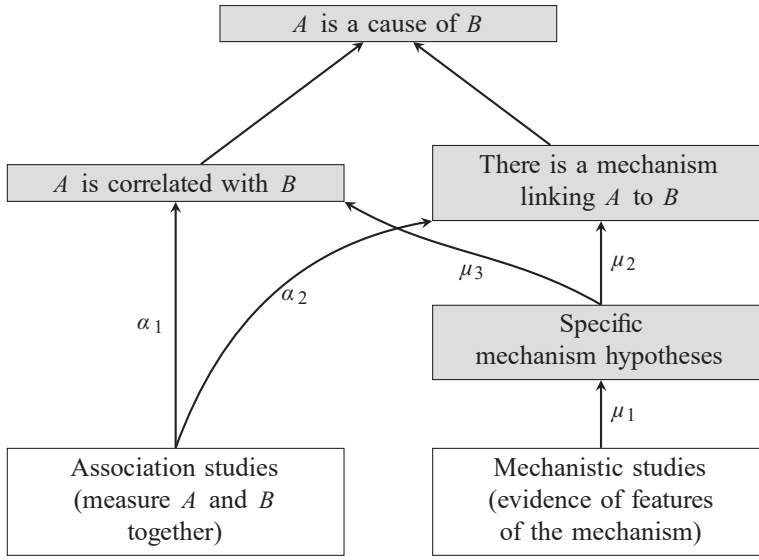

Figure 1.1 Evidential relationships for causal enquiry (Williamson, 2021a).

At the top of this diagram is the causal claim under investigation. Evidence for the claim that A is a cause of B is provided by evidence for two further claims: that A and B are appropriately correlated and that they are linked by a suitable mechanism. This part of the diagram corresponds to the first component of Evidential Pluralism:

> In order to establish a causal claim, one normally needs to establish two propositions: that the putative cause and effect are appropriately correlated, and that there is some mechanism complex involving the putative cause which is responsible for the putative effect and which can account for the extent of the correlation.

We shall call this component of Evidential Pluralism *object pluralism*, because it specifies two objects of evidence—two claims that we need evidence *for* (Illari, 2011). Let us consider in more detail some of the terms that occur in the statement of object pluralism.

'Establish'. A proposition is established just when standards are met for treating the proposition itself as evidence, to be used to help assess further propositions.[1]

1 This leaves open the question of what constitutes evidence. An agent's evidence has variously been analysed as her knowledge, or her full beliefs, or those of her degrees of belief which are set by observation, or her information, or what she rationally grants (Williamson, 2015). Fortunately we do not need to settle this very controversial philosophical question here. For our purposes, all that is required is an account of when a putative item of evidence might be relevant to a causal claim. Evidential Pluralism provides such an account, as we will see.

Whether a proposition is established depends on prior evidence. Prior evidence must warrant not only sufficiently high confidence in the truth of the proposition, but also high confidence that further evidence will not call the proposition into question, i.e., that confidence in the proposition will remain sufficiently high in the light of new evidence. We can refer to these as the 'threshold' and 'stability' conditions for establishing (Williamson, 2022).

One can distinguish establishing as an evidential relation from establishing as an act. As an evidential relation, an evidence base establishes a proposition when the threshold and stability conditions are met. As an act, an agent establishes a proposition when her evidence base establishes the proposition and she takes the proposition to be evident on those grounds. The act of establishing does not need to be a conscious act. Moreover, several propositions can be established by a single act: for example, causation, correlation and mechanism might all be established at once, if Evidential Pluralism is correct. An agent is 'in a position' to establish a proposition if her evidence base establishes the proposition. One can then construe Evidential Pluralism in terms of either the act or the evidential relation. In the former case, Evidential Pluralism offers practical advice for someone intending to establish causation: for you to establish a causal claim, you need to gather evidence in order to be in a position to establish both correlation and mechanism. Under the evidential relation construal, Evidential Pluralism is impersonal: it says that the evidence base establishes causation just when it establishes correlation and mechanism.

Establishing requires meeting a high epistemological standard. In particular, establishing a causal claim should be distinguished from acting in accord with a causal claim as a precautionary measure. In certain cases in which a proposed intervention has a clear benefit and a relatively low cost, or if failing to act has a high cost, it may be appropriate to initiate the intervention even when its effectiveness has not been established, so that benefits can be reaped in case it turns out to be effective.

Although establishing requires meeting a high epistemological standard, it is fallible. One's prior evidence can be systematically misleading, making a proposition very plausible and making it very plausible that confidence in the proposition will not significantly decrease in the light of new evidence, even though the proposition is in fact false. If the threshold and stability conditions are sufficiently demanding, such cases will be very rare, which reduces the need to revisit previously established propositions. On the other hand, if the conditions are too demanding then they will too rarely be met and enquiry will stall. The need to avoid falsity must be balanced against the need to establish truth.

This balance may differ from discipline to discipline and field to field. A younger field, such as social psychology, may attach less weight to avoiding falsity than an older field, such as particle physics, in the interest of building up a body of established propositions. (This kind of difference between fields is witnessed by their use of different *p*-values when establishing statistical hypotheses.)

As this book has broad scope, covering many different fields, we make no attempt to commit to specific standards for establishing.[2]

A proposition will be said to be 'ruled out' just when its negation is established. Thus the claim that A is a cause of B is ruled out when it is established that A is not a cause of B. In §12 we introduce a range of other possible status levels, intermediate between *established* and *ruled out*, that the available evidence may confer on a proposition.

'Needs'. Evidential Pluralism is primarily a normative thesis, about what one *ought* to do in order to establish, or assess, causation. Whether particular instances of causal enquiry conform to its advice varies: not all practice is good practice. In Part II and Part III, we will examine a range of examples of causal enquiry to investigate the extent to which they can be thought of as conforming to Evidential Pluralism.

Next we turn to 'correlation' and 'mechanism'. Importantly, Evidential Pluralism holds that establishing causation requires establishing the *existence* of a correlation and the *existence* of a mechanism, not necessarily the extent of the correlation, nor the details of the mechanism. In some cases, of course, establishing the extent of a correlation is a means to establishing its existence, and establishing the details of a mechanism is a means to establishing its existence, but these means are not the only means.

'Correlation'. Evidence needs to establish that the putative cause and effect are 'appropriately correlated'. Here, 'appropriately correlated' just means *probabilistically dependent conditional on all potential confounders*,[3] where the probability distribution in question is relative to a specified population or reference class of individuals. Thus, if A is the putative cause variable, B the putative effect variable, and C is the set of potential confounder variables, one needs to establish that A and B are probabilistically dependent conditional on C, often written $A \not\perp\!\!\!\perp B \mid C$.

Note that if the causal claim is single-case, i.e., if the reference class only contains a single individual or a single situation, then the probability function in question needs to be interpreted as single-case probability, such as single-case chance or rational degree of belief. See §4 and §30.2 for more discussion of single-case causal claims. On the other hand, if the causal claim is generic, i.e., repeatedly

2 There is one philosophical approach, *knowledge-first epistemology*, which deems it futile to attempt to provide specific standards for establishing a proposition. See Wilde (2021) for a discussion in relation to establishing mechanistic hypotheses. Note that the knowledge-first approach also takes establishing to be infallible, however, and holds that one cannot in general determine what one's evidence is. This is quite a departure from the common usage of the terms 'establishing' and 'evidence' in scientific practice. In this book, we stick more closely to the standard usage of these terms.

3 'Correlated' is often used in other senses, e.g., as meaning unconditionally probabilistically dependent, or unconditionally linearly dependent. These are not the senses in use here. On the other hand, much of what we say in the book also goes through when 'correlated' is used in senses other than probabilistic dependence conditional on all potential confounders.

instantiable, then the probability function can be interpreted as generic probability, such has limiting relative frequency or generic propensity.[4]

A confounder is a variable associated with both *A* and *B*, e.g., a common cause of *A* and *B* (Figure 1.2). One needs to establish probabilistic dependence conditional on confounders because an observed association between *A* and *B* might be attributable to their association with *C*, rather than attributable to *A* being a cause of *B*. The set of 'potential' confounders should include any variable that plausibly might be a confounder, given the available evidence of the area in question. The set of potential confounders includes variables that are established by current evidence to be confounders (e.g., the known common causes of *A* and *B*), as well as variables that are deemed confounders by more speculative mechanistic hypotheses. As with standards of establishing, standards of what counts as a potential confounder may vary from field to field. If too many variables are taken to be potential confounders, it will be very difficult to establish correlation. On the other hand, if too few variables are considered as potential confounders, establishing correlation will only weakly confirm causation. Thus, a balance needs to be struck.

Establishing correlation is non-trivial for two reasons. First, because it requires establishing a probabilistic dependence in the data-generating distribution, rather than simply in the distribution of a sample of observed outcomes. The method of sampling and size of sample can conspire to render an observed sample association a poor estimate of a correlation in the population at large. Second, the correlation claim invokes all potential confounders, and there can be very many of these.

On the other hand, establishing correlation, i.e., establishing probabilistic dependence on all potential confounders, is not as difficult as establishing probabilistic dependence conditional on all *possible* confounders. The set of all *possible* confounders includes not only the potential confounders but also unforeseeable confounders: variables that, on the basis of current evidence, cannot be anticipated to be confounders. Establishing probabilistic dependence conditional on all possible confounders is an extremely demanding task. Evidential Pluralism does not require establishing this latter form of probabilistic dependence, although some other accounts of causal enquiry do. For example, probabilistic theories of causality tend to identify causality with correlation conditional on all confounders—not

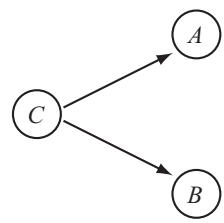

Figure 1.2 Common cause *C* is a potential confounder, with respect to the relationship between *A* and *B*.

4 See Gillies (2000) for an introduction to interpretations of probability.

merely on those variables that might reasonably be anticipated to be confounders. It is hard to see how many standard methods can be used to establish causation under such an account. Indeed, a probabilistic account of causality can lead to scepticism about any causal claim that is supported by anything but the largest randomised controlled trials (RCTs). Evidential Pluralism is less demanding. One merit of Evidential Pluralism is that it shows how study designs other than RCTs can help to establish a causal claim. We will encounter many examples of the confirmatory value of other kinds of study design in this book.

To be clear about terminology, we shall use 'observed correlation' to refer to a correlation found in the sample data, 'genuine correlation' to refer to a correlation in the population from which the data are drawn and 'established correlation' to refer to a claimed genuine correlation that has met the threshold and stability conditions for establishing. Since establishing is fallible, the fact that a correlation is established does not guarantee that there is a genuine correlation, though it makes a genuine correlation very likely. We will use 'association' in a more general sense than 'correlation', to refer to any kind of probabilistic dependence of two variables—either unconditional probabilistic dependence or probabilistic dependence conditional on some further set of variables (e.g., a subset of potential confounders). Thus a correlation is an association between A and B conditional on the set of all potential confounders.

'Mechanism'. We take a mechanism to consist of a complex-systems mechanism, a mechanistic process or some combination of the two. A *complex-systems mechanism* consists of entities and activities organised in such a way that they are responsible for some phenomenon to be explained (Machamer et al., 2000; Illari and Williamson, 2012). An example is the mechanism by which a stock exchange trades stocks. A *mechanistic process* is a spatiotemporally contiguous process along which a signal can be propagated (Reichenbach, 1956; Salmon, 1998). An example is an electronic signal being transmitted via air and cable from a smartphone to a stock exchange. A mechanism might also be composed of both these sorts of mechanisms: for example, a trading mechanism might include the complex-systems mechanism of the smartphone, the complex-systems mechanism by which the stock exchange trades stocks and the mechanistic processes linking the two.

With regard to Evidential Pluralism, it is not enough to establish the existence of some mechanism of action: it is also essential to establish that the influence of this mechanism is not negated by that of counteracting mechanisms. Thus one needs to consider the whole complex of mechanisms that link A and B and ask whether there is a mechanism of action, which when taken together with any counteracting and reinforcing mechanisms, can account for the extent of the observed correlation between A and B. Moreover, when establishing a generic causal claim it is not enough to show the purported mechanism merely exists in some individuals—it needs to be present in sufficiently many individuals to account for the extent of the observed correlation.

'Normally'. Evidential Pluralism says that one 'normally' needs to establish both correlation and mechanism. This is because there are certain cases in which causality is apparently not accompanied by a correlation and there are also cases in which causality is apparently not accompanied by an underlying mechanism. If this is so, one cannot expect to establish both correlation and mechanism in these pathological cases.

In cases of overdetermination, where the cause does not raise the probability of the effect because the effect will happen anyway, the cause and effect need not be genuinely correlated. In many such cases, one can at least expect a counterfactual correlation: if things had been different in such a way that the effect would not have happened anyway—e.g., had a second, overdetermining cause been eliminated—then the cause and effect would indeed be correlated. One might think, then, that one ought to be able to establish a counterfactual correlation for any causal claim, if not an actual correlation. However, there are cases in which the cause of interest and a second, overdetermining cause are mutually exclusive, so that it is not possible both to eliminate the second cause and allow the first cause to vary so as to establish a correlation. For example, an unstable atom may decay to one of two mutually exclusive intermediary states, B and B', on the way to a ground state C. Attaining either one of the intermediary states causes the particle to reach the ground state, even though there may well be no correlation, $P(C|B) = P(C|B') = P(C)$. One cannot simply eliminate B' and vary B here (see Williamson, 2009, §10). Therefore, even the demand for a counterfactual correlation may be too strong.

Let us turn next to causality without mechanisms. Where the cause and/or the effect is the absence of some event, it cannot be connected by an actual mechanism. In many such cases, one can expect a counterfactual mechanism. Suppose cause and effect are both absences: e.g., failing to charge the smartphone causes an absence of a trade at the stock exchange. If things had been different in such a way that what was absent in the cause were present (the phone was charged), then one would expect a mechanism from this presence to a presence corresponding to the effect (the trade). One might think, then, that one ought to be able to establish the existence of a counterfactual mechanism for any causal claim, if not an actual mechanism. However, there are cases where one of the cause and effect is an absence and the other is a presence, and this strategy does not work. For example, suppose that failing to charge the phone causes a trade to go ahead that would otherwise be cancelled. That the cause is an absence precludes a mechanism here, but the effect being a presence precludes any mechanism reaching the effect in the obverse case, namely, charging the phone is a cause of the absence of the trade.

Now, establishing causality in these pathological cases is not particularly problematic in practice. However, it is more subtle than simply establishing both correlation and mechanism, even where counterfactual correlations and mechanisms are admitted. The question as to how Evidential Pluralism can be modified to say something novel in such cases will be not be considered here, because it is not central to the arguments of this book. The use of 'normally' is intended to leave open the possibility that in certain cases, such as overdetermination cases and

causation between absences, one might not need to establish both correlation and mechanism.

Having discussed object pluralism, let us turn to the second component of Evidential Pluralism, represented by the bottom part of Figure 1.1:

> So, in order to assess a causal claim, one normally needs to assess relevant association studies and mechanistic studies where available.

We will refer to this second component of Evidential Pluralism as *study pluralism*, because it says that we should pay attention to two sorts of roles that studies play in relation to a particular causal claim: i.e., we should pay attention to studies that act as association studies and to those that act as mechanistic studies. Again, let us examine the terms that appear here in more detail.

'Association studies'. An association study for *A* and *B* is a study that samples instances of *A* and *B*, often together with a subset *C* of potential confounders, in order to test whether *A* and *B* are probabilistically dependent conditional on *C*. Association studies sometimes attempt to estimate the extent of any probabilistic dependence, and sometimes look for a mathematical function that characterises the value that *B* takes in terms of those that *A* and *C* take. Association studies include experimental studies, such as RCTs, as well as observational studies, such as cohort studies, case control studies, case series and *n*-of-1 studies.

'Mechanistic studies'. A mechanistic study for *A* and *B* is a study that seeks to shed light on features of mechanisms linking *A* and *B*. Features of mechanisms include intermediary variables, entities, activities and the spatiotemporal organisation of these entities and activities, as well as processes and their interactions. A mechanistic study may be devised to test a *specific mechanism hypothesis*, which hypothesises one or more features of the mechanism complex linking *A* and *B*. Mechanistic studies can be very diverse, methodologically. For example, one sort of mechanistic study for *A* and *B* tests whether a variable *D* might be an intermediary between *A* and *B* by testing for an association between *A* and *D*, and/or between *D* and *B*, or between *D* and other variables that have been previously established to mediate between *A* and *B*. Another sort of mechanistic study might seek to directly observe some proposed feature of a mechanism: e.g., by means of video surveillance, a survey, medical imaging or an autopsy. Another sort of mechanistic study might seek to simulate a proposed mechanism in order to isolate key features: e.g., by means of an agent-based model of a social mechanism. Mechanistic studies can be experimental or observational. Confirmed theory and analogical reasoning can play important roles in the design of a mechanistic study or the interpretation of its results.

It is important to emphasise that the concepts of association study and mechanistic study are relativised to the putative cause and effect under consideration, i.e., to *A* and *B*. As we have just seen, a study that tests for a correlation between *A* and

D, where *D* is a variable that is hypothesised to mediate between *A* and *B*, qualifies as a mechanistic study for *A* and *B* but not an association study for *A* and *B*. It does, however, count as an association study for *A* and *D*. For example, consider the question of whether a social intervention *A* causes behaviour *B*. It may be difficult to measure *B*, so a study may examine whether the intervention causes a proxy outcome *D*, namely an *intention* towards behaviour *B*. If the study tests for an association between *A* and *D*, it is an association study for *A* and *D*. However, it is not an association study for *A* and *B*—it is a mechanistic study for *A* and *B* because it tests for a feature *D* that mediates on the mechanism from intervention *A* to behaviour *B*.

In the light of this relativity to causal claim, it is helpful to think of the distinction between an association study and a mechanistic study as a distinction between two possible *roles* that a study can play in the context of the particular causal claim. Note that these roles are not mutually exclusive. It is in principle possible to design a study that is both an association study for *A* and *B* and a mechanistic study for *A* and *B*. Such studies are not the norm, however: association studies for *A* and *B* are usually not also mechanistic studies for *A* and *B*.

Moreover, the distinction between association studies and mechanistic studies does not correspond to a distinction between the kinds of methods employed in a study. As we will see, each type of study can involve quantitative or qualitative methods, large-*n* or small-*n* methods, and so on. Study pluralism (the second component of the Evidential Pluralism thesis) should not be confused with method pluralism. There is, however, a connection between the two: as we will see in Chapter 4, Evidential Pluralism can be used to motivate mixed methods research.

'Assess'. Here we are concerned solely with assessing the truth of a causal claim, not its utility or its complexity, for example. When assessing studies, we need to explicitly and systematically evaluate their quality, their results and the status they confer on claims of interest. For Evidential Pluralism, the claims of interest are the causal claim, the correlation claim, the general mechanistic claim and the specific mechanism hypotheses.

'Normally'. Evidential Pluralism urges the evaluation of both association studies and mechanistic studies as a means to make better assessments of causality. There are exceptions, however. There is no need to explicitly and systematically evaluate an association study if the correlation that it is studying has already been established or ruled out, and there is no new evidence that challenges this verdict. Similarly, there is no need to explicitly and systematically evaluate a mechanistic study if it is studying some feature of a mechanism that has already been established or ruled out, unless there is new evidence that motivates revisiting that feature. If association studies are assessed first and it is deemed that they would on their own establish (or rule out) causation, then there is no need to go on to assess mechanistic studies unless there are specific mechanism hypotheses that could undermine the verdict of the association studies. A similar point applies to the situation in which mechanistic studies are evaluated first, although it is perhaps less often the case that mechanistic studies suffice to establish causation on their own.

Table 1.1 Possible explanations of an observed correlation between A and B (Williamson, 2019a).

Causation	A is a cause of B.
Reverse causation	B is a cause of A.
Confounding (selection bias)	There is some confounder C that has not been adequately controlled for by the study.
Performance bias	Those in the A-group are identified and treated differently to those in the $\neg A$-group.
Detection bias	B is measured differently in the A-group in comparison to the $\neg A$-group.
Chance	Sheer coincidence, attributable to the size of the sample.
Fishing	Measuring so many outcomes that there is likely to be a chance correlation between A and some such B.
Temporal trends	A and B both increase over time for independent reasons. E.g., British bread prices & Venetian sea levels.
Semantic relationships	Overlapping meaning. E.g., bachelor & unmarried.
Constitutive relationships	One variable is a part or component of the other.
Logical relationships	Measurable variables A and B are logically complex and logically overlapping. E.g., A is $C \wedge D$ and B is $D \vee E$.
Nomological relationships	E.g., a conservation law can induce a correlation between the conserved quantities.
Mathematical relationships	E.g., between profit & loss.

'Where available'. There is no need to evaluate both association studies and mechanistic studies if there simply are no relevant association studies available, or if there are no relevant mechanistic studies available. One can only assess the evidence that can be made available.

We shall explore how this assessment should proceed in Chapter 3 and subsequent chapters.

§2 Why is Evidential Pluralism Plausible?

It is a truism that correlation is insufficient for causation. This is because a correlation between A and B admits a large number of possible explanations, only one of which is that A is a cause of B. Other potential explanations include: that B is a cause of A; that A and B are effects of a common cause; that the correlation is attributable to bias or chance; or that it is attributable to a non-causal connection between A and B. Table 1.1 catalogues a range of such explanations.

So what do we need to add to correlation to be able to infer causation? When A is a cause of B, there is some mechanism complex underlying the causal relationship that explains why instances of A are at least partly responsible for instances of B. If the correlation is not attributable to causation, it is not a mechanism from A to B that explains the correlation, but some other feature—e.g., a common cause C, or a semantic relationship between A and B. So, having established correlation, establishing the existence of a suitable mechanism is precisely what is needed to establish causation and rule out the other possible explanations of the correlation.

(There is one caveat here. If the mechanism in question could only be responsible for a small fraction of the observed correlation, then the bulk of the correlation must be due to confounding, or a non-causal relationship, etc. That this alternative explanation is responsible for the bulk of the correlation makes it more plausible that it is responsible for all of it, undermining the claim that even a small fraction of the correlation is attributable to causation. Thus the mechanism needs to be able to explain much of the extent of the correlation in order to establish causation.)

On the other hand, establishing the existence of a mechanism is also not enough on its own to establish causation. One cannot directly infer from there being a mechanism from A to B that A makes a difference to B: the mechanism may be long and complex, or there may be counteracting components, making it hard to predict whether A has any net effect on B. This is precisely why it is essential to establish a correlation: in order to show that A makes an overall difference to B.[5]

Hence, establishing causation requires establishing both correlation and mechanism. This gives the upper part of Figure 1.1 and object pluralism, the first component of Evidential Pluralism:

> In order to establish a causal claim, one normally needs to establish two propositions: that the putative cause and effect are appropriately correlated, and that there is some mechanism complex involving the putative cause which is responsible for the putative effect and which can account for the extent of the correlation.

Next, let us turn to the lower part of Figure 1.1 and the second component of Evidential Pluralism, study pluralism:

> So, in order to assess a causal claim, one normally needs to assess relevant association studies and mechanistic studies where available.

Association studies test for an association between A and B, so they directly confirm correlation along channel α_1 in Figure 1.1. The presence of a correlation can be confirmed by RCTs that estimate the extent of the association conditional on all possible confounders, or by observational studies that observe associations conditional on some or all potential confounders. Hence, to assess correlation one normally needs to assess available association studies.

In certain situations, association studies can also provide indirect evidence of the existence of a mechanism (channel α_2). For example, RCTs are valued because, when well devised and well conducted, they can reduce the probability of confounding by unforeseeable confounders. Thus, if a high-quality RCT finds a substantial correlation, and if non-causal connections are ruled out, it becomes more plausible that the correlation is attributable to some underlying mechanism—i.e.,

5 That there is a chain of causal relationships linking A to B does not imply that A is itself a cause of B. This is known as the *intransitivity* of causation. See Hitchcock (2001) for one account of this intransitivity.

Table 1.2 Examples of mechanistic studies in the social sciences.

Specific mechanism hypothesis	*Study (and study design)*
More stringent search requirements (as an intermediary variable between strictness of unemployment benefits and unemployment rates) lower the chance of being reemployed by the same employer.	The 'work search' experiment of Johnson and Klepinger (1994) on the effects of monitoring and sanctions on unemployment. (RCT)
Legalised abortion has a disproportionate effect on the birth of those who are most at risk of engaging in criminal behaviour, which in turn decreases crime rates.	The longitudinal analysis in the research of Donohue and Levitt (2001) on legalised abortion and crime rates. See §24 for more detail. (Longitudinal study)
The initial endowments to which rebel leaders have access constrain their tactics of recruitment, which shape the membership profile of a rebel group. The membership profile affects its internal organisation and thereby its strategies of violence in war.	The ethnographic interviews in the research of Weinstein (2007) on causes of the patterns of violence in rebels. See §29. (Interview)
Family socioeconomic status and abilities affect the influence of parents, teachers and friends on a youth and her own observations of her ability, which affect her levels of educational and occupational aspiration, that ultimately influence subsequent levels of educational attainment.	The cross-sectional study in the research of Sewell et al. (1969) on family background and educational attainment. (Cross-sectional study)
Economic interdependence between states influences the expectation that a state has of future trade with other states, which eventually influences the chances of military conflict.	The large-N qualitative study in the research of Copeland (2015) on economic interdependence and military conflict between states. (Large-N qualitative analysis)

confirmation proceeds along channel α_2. Thus, in situations in which explanations other than causation and confounding have antecedently been ruled out, RCTs can suffice to establish causation, via channels α_1 and α_2 of Figure 1.1.

Alternatively, one can directly confirm the existence of a mechanism by positing features of the mechanism complex and confirming the presence of those features. This is why one should normally assess mechanistic studies: it is by means of mechanistic studies that one can confirm specific mechanism hypotheses (channel μ_1) and thus confirm the existence of a suitable mechanism (μ_2). For some examples of mechanism hypotheses and mechanistic studies in social science research, see Table 1.2.

In certain situations, specific mechanism hypotheses can also provide indirect evidence of the existence of a correlation (channel μ_3). For example, having established the key features of the parachute mechanism, one can be more confident that the use of parachutes will reduce the risk of serious injury when falling from an aeroplane. Indeed, the features of the mechanism, together with unsystematic observations of the use of parachutes, are enough to establish causation—there is no need for randomised trials to test the effects of parachutes (Williamson, 2019a, §2.2).

In sum, the reason it is important to assess both association studies and mechanistic studies is that they are what provide evidence of correlation and evidence of mechanisms. Moreover, they reinforce one another. Association studies directly test for a correlation, but they are prone to various biases which can make it hard to infer the existence of a mechanism. On the other hand, mechanistic studies directly test for a mechanism, but it can be hard to tell whether there is a net correlation when a mechanism is long or complex or has counteracting pathways. In isolation, each kind of study is usually an unreliable indicator of causation, but each compensates for the deficiencies of the other, so they are best taken together when assessing causation (Auker-Howlett and Wilde, 2020).

One key virtue of Evidential Pluralism is that it shows how, in favourable circumstances, RCTs can be sufficient to establish causation. As noted, another is that it explains why RCTs are not required to establish causation—Evidential Pluralism shows how it is possible to establish causation by means of a judicious combination of observational association studies and mechanistic studies. We will see throughout this book that Evidential Pluralism provides an integrated account of causal enquiry, into which many existing methods fit. For instance, Evidential Pluralism can show how graphical causal modelling methods (§13), mixed methods research (§17) and process-tracing (§30) can help establish causation.

§3 Evidential Pluralism and EBM+

Evidential Pluralism was developed in the context of medicine and the health sciences (Russo and Williamson, 2007). There, it led to the EBM+ programme, which seeks to improve the methods of evidence-based medicine (Parkkinen et al., 2018). Here we briefly introduce this application to medicine.

The aim of evidence-based medicine (EBM) is to make evidence explicit and to systematically assess this evidence using explicit evaluation methods, in order to improve inferences and decision making in medicine. In practice, the way present-day EBM goes about this is by prioritising some kinds of evidence over others. In particular, EBM rates clinical association studies more highly than other kinds of evidence, including mechanistic studies. This has been the case since the early days of EBM:

> Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiologic rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research.
>
> (Guyatt et al., 1992, p. 2420)

The State University of New York (SUNY) Downstate Health Sciences University EBM evidence hierarchy, Table 1.3, provides a more recent example of this approach: in vitro research and animal research, which can provide very strong evidence of mechanisms, are ranked below ideas and opinions in the evidence

Table 1.3 The SUNY evidence hierarchy (SUNY, 2004).

Systematic reviews and meta-analyses
Randomised controlled double blind studies
Cohort studies
Case control studies
Case series
Case reports
Ideas, editorials, opinions
Animal research
In vitro ('test tube') research

Table 1.4 The Oxford Centre for Evidence-Based Medicine evidence hierarchy for treatment benefits (OCEBM Levels of Evidence Working Group, 2011).

Systematic review of randomised trials or *n*-of-1 trials
Randomised trial or observational study with dramatic effect
Non-randomised controlled cohort/follow-up study
Case-series, case-control studies, or historically controlled studies
Mechanism-based reasoning

hierarchy, with association studies higher up the hierarchy, and systematic reviews and meta-analyses of association studies at the very top. The current evidence hierarchy of the Oxford Centre for Evidence-Based Medicine ranks mechanism-based reasoning at the lowest level (Table 1.4). Here 'mechanism-based reasoning' might be interpreted as reasoning involving the μ-channels of Figure 1.1. Many other evidence hierarchies ignore mechanistic studies altogether, i.e., mechanistic studies don't even qualify as evidence. For example, the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) evaluation process, which is now very widely used, considers only association studies as admissible studies: ceteris paribus, randomised trials are ranked above non-randomised experimental trials and observational studies, and mechanistic studies do not feature at all (Schünemann et al., 2013, §5.1.1). Elsewhere, those involved in the development of GRADE do consider mechanistic studies, but they appear towards the bottom of the evidence hierarchy, just above clinical experience (Table 1.5).

Clearly, then, Evidential Pluralism conflicts with the way in which present-day EBM devalues mechanistic studies. From the perspective of Evidential Pluralism, EBM's monistic focus on association studies is only half the story. Evidential Pluralism is a dualist account, with mechanistic studies treated alongside association studies (Figure 1.1). This dualist approach precludes a single, linear evidence hierarchy encompassing both association studies and mechanistic studies, but, if Evidential Pluralism is correct, this would be a price worth paying for improved causal assessment.

On the other hand, there are some points in common between EBM and Evidential Pluralism. Evidential Pluralism is concerned with making evidence and its evaluation explicit, in order to improve causal enquiry. This is very much in the same spirit as EBM. Moreover, Evidential Pluralism agrees with EBM that

Table 1.5 An evidence hierarchy related to the GRADE approach (Guyatt et al., 2015, p.15).

N-of-1 clinical trial
Multiple-patient randomised trials
Observational studies: patient-important outcomes
Basic research: laboratory, animal, human physiology
Clinical experience

Table 1.6 Examples of sources of evidence of mechanisms in medicine (Clarke et al., 2014).

Direct manipulation: e.g., in vitro experiments
Direct observation: e.g., biomedical imaging, autopsy
Statistical studies: e.g., RCTs, cohort studies, case control studies, case series
Confirmed theory: e.g., established immunological theory
Analogy: e.g., animal experiments
Simulation: e.g., agent-based models

association studies are crucial to proper evidence evaluation, and that fully blinded RCTs can be more informative than other kinds of association study. Thus, Evidential Pluralism motivates modifying, rather than abandoning, EBM.

The EBM+ programme is an extension of EBM which seeks to systematically consider mechanistic studies in addition to association studies. Parkkinen et al. (2018) provide a handbook for EBM+: this includes a theoretical framework in line with Evidential Pluralism, as well as tools for assessing mechanistic studies and integrating this assessment with an assessment of association studies provided by, e.g., the GRADE approach. They divide causal assessment into two tasks: assessing efficacy, i.e., whether the causal relationship holds in the study population, and assessing external validity, i.e., whether the causal relationship can be extrapolated from the study population to some other target population of interest. They argue that evidence of mechanisms is key to both these tasks. Because these two tasks are also of central importance in the social sciences, we will consider them in some detail in §12.

In medicine, high-quality evidence of mechanisms can be obtained by a wide variety of means. Table 1.6 provides some examples. The same is true of the social sciences, and we will encounter many examples of evidence of social mechanisms throughout this book.

Examples of the application of EBM+ methods include the work of Abdin et al. (2019), who assess whether Amoxicillin is a cause of Drug Reaction with Eosinophilia and Systemic Symptoms (DRESS), that of Auker-Howlett and Wilde (2020), who discuss whether a pegylated combination therapy of peginterferon alfa and ribavirin is effective against Hepatitis C, and that of Auker-Howlett (2020, Chapter 3), who assesses a combination therapy of interferons and ribavirin for the treatment of Middle East Respiratory Syndrome (MERS). Aronson et al. (2018) advocate the use of EBM+ methods in the drug approval process, while Aronson et al. (2021) argue that they are well suited to the assessment of coronavirus interventions. Williamson (2019b) argues that the methods of the International Agency for Research

on Cancer (IARC) conform quite closely to EBM+, and Williamson (2021c) cites this as evidence for the practical feasibility of EBM+. Wilde and Parkkinen (2019) discuss the reliability of IARC's use of mechanism-based extrapolation, while Parkkinen and Williamson (2020) argue that the EBM+ approach helps to shed light on the process of extrapolation from model organisms in pharmacology. Tonelli and Williamson (2020) and Pérez-González and Rocca (2022) advocate an approach based on Evidential Pluralism for clinical practice. Levack-Payne (2022a,b) argues that the EBM+ approach can be fruitfully applied to sports science. Other case studies of the application of Evidential Pluralism to medicine can be found in Clarke (2011); Parkkinen (2016) and Gillies (2019).

§4 Controversies and Clarifications

Initially, Evidential Pluralism had a mixed reception in the literature, for two main reasons. Firstly, it conflicts with present-day EBM, and although EBM and its evidence hierarchies have their detractors, EBM is now orthodoxy, central to many health institutions around the world. Second, there were some misunderstandings about what Evidential Pluralism seeks to do and what it says. With respect to the first point, researchers and practitioners are beginning to see the value of Evidential Pluralism, despite the tension with present-day EBM.[6] With respect to the second point, it may be helpful to highlight certain aspects of Evidential Pluralism in order to try to ward off further misunderstandings.

Objects and types of evidence. We noted earlier that Russo and Williamson (2007) put forward the thesis that 'the health sciences make causal claims on the basis of evidence both of physical mechanisms, and of probabilistic dependencies' (p. 157). As Illari (2011) pointed out, it is important to distinguish the types of evidence from the objects of evidence. The objects of evidence are the claims that are being assessed by evidence: e.g., the claim that A and B are correlated, and the claim that they are linked by an appropriate mechanism. According to Evidential Pluralism, establishing causality normally requires evidence of both objects, i.e., evidence of correlation and evidence of mechanism. The types of evidence are the kinds of study that constitute the evidence: e.g., association studies, or mechanistic studies. Evidential Pluralism does not demand both kinds of study in all situations. It is in principle possible to establish causation purely on the basis of association studies, since association studies can, in certain circumstances, establish mechanism as well as correlation. Similarly, it is in principle possible to establish causation purely on the basis of mechanistic studies, since mechanistic studies can, in certain circumstances, establish correlation as well as mechanism. However, these situations are rare and when assessing causation it will typically be important to consider both kinds of study, should both be available. In this

6 The methods of IARC, for example, have been moving even closer to EBM+ (Samet et al., 2020, p. 34), and Greenhalgh et al. (2022) argues that the COVID-19 pandemic made a move to EBM+ more urgent.

book we will avoid the terms 'mechanistic evidence' and 'probabilistic evidence' where possible, in order to avoid ambiguity between the objects of evidence and types of study.[7]

Mechanism details. A second point to emphasise is that Evidential Pluralism does not demand that one establish all the details of a mechanism. Establishing a specific mechanism hypothesis that posits a few key features of a mechanism can provide strong confirmation along channel μ_2 of Figure 1.1.[8] Moreover, as Wieten (2018, Chapter 3) argues, specifying a mechanism does not require specifying all the contextual factors that enable the mechanism to operate, nor all the factors that might interfere with the mechanism to block its operation. What is important is that the mechanism operates often enough to account for the extent of any observed correlation.

Epistemology. It is important to stress that Evidential Pluralism is an account of the epistemology of causality, not an analysis of our concept of cause, nor a metaphysical account of the nature of causality. The thesis is that causal assessment requires evidence of correlation and evidence of mechanisms, not that *A* causes *B* if and only if *A* and *B* are correlated and connected by an appropriate mechanism. As discussed on p. 12, certain counterexamples challenge an analysis of causality of the latter form—hence the 'normally' hedge in the statement of object pluralism. These counterexamples are not counterexamples to Evidential Pluralism. We will discuss what sort of examples would challenge Evidential Pluralism in Chapter 5, where we look at objections to the claim that Evidential Pluralism can be fruitfully applied to the social sciences.

Evidential Pluralism vs evidential diversity. We use 'Evidential Pluralism' to refer to a specific theory that appeals to object pluralism and study pluralism, as formulated in §1.[9] It should not be confused with other claims about diversity of evidence. For example, the Principle of Total Evidence says that one should consider all available evidence when assessing a proposition. Evidential Pluralism has narrower scope than the Principle of Total Evidence, because it is restricted to causal claims, yet it is more informative, because it specifies the objects and kinds of evidence that are relevant to the assessment of a causal claim, and, via Figure 1.1, how this evidence confirms a causal claim. Similarly, Evidential Pluralism is more specific than the claim that it is desirable to use diverse methods, theories

7 Howick (2011a,b) takes the Russo-Williamson Thesis (RWT) to be demanding both types of study and criticises it on those grounds. However, RWT, and Evidential Pluralism as formulated in this book, demand evidence of both objects, not of both types. See Gillies (2019, Chapters 8–10) for a detailed discussion of the defensibility of a version of RWT that requires two types of evidence.

8 Howick (2011a, §10.4.1) notes that partially understood mechanisms can be misleading: interventions motivated purely by partially understood mechanisms have often turned out to be ineffective. This is why it is essential to consider association studies alongside mechanistic studies, and not to take causation to be established unless both correlation and mechanism have been established.

9 Hence the use of capital letters: 'Evidential Pluralism' is the name of a specific theory.

or perspectives in research (aka 'triangulation'): Evidential Pluralism holds that a very particular kind of triangulation is required for causal inference. Mixed methods researchers advocate another kind of diversity of evidence in the social sciences: the use of both qualitative and quantitative methods. Evidential Pluralism offers a very different conceptualisation. However, as we will see in Chapter 4, there are interesting connections between Evidential Pluralism and mixed methods research.

Extrapolation. As noted in the previous section, Evidential Pluralism also handles the extrapolation of causal claims from one population to another. A causal claim that has been established in a source population can be extrapolated to a target population if the associated mechanisms are sufficiently similar in the source and target populations (Steel, 2008), i.e., if the mechanism of action is sufficiently similar and any new counteracting mechanisms do not completely mask the mechanism of action. Evidence that supports an extrapolation includes (i) evidence from mechanistic studies in the source and target populations that key features of the mechanism of action are preserved from source to target population and (ii) evidence from association studies and mechanistic studies in the target population that any new counteracting mechanisms in the target populations do not completely mask this mechanism of action. Figure 1.3 portrays these evidential relations in full, while Figure 1.4 offers an abbreviated representation. In Figure 1.4, 'source studies' include any relevant association studies and mechanistic studies in a source population that differs from the target population of interest. Extrapolation becomes pertinent when the source studies suffice to establish causation in the source population, and the association studies and mechanistic studies in the target population do not establish causation there, but do confirm mechanistic similarity. In such a case, it may be possible to establish causation in the target population by considering the source studies alongside the target association and mechanistic studies.

Generic vs single-case causation. Evidential Pluralism is intended to apply to both generic causation and single-case causal claims. A causal claim of the form *A* is a cause of *B* is generic if at least one of its relata *A* and *B* is repeatedly instantiable; otherwise it is single-case (Russo and Williamson, 2011, p. 48). Thus, the claim that the legalisation of abortion in the USA in the 1970s is a cause of the drop in violent crime rates in the 1990s is single-case, while the claim that the initial finances of a rebel group are a cause of its strategy of violence is generic.[10] Evidential Pluralism maintains that, in each case, one needs to establish both correlation and mechanism to establish causation, so one ought to consider both association and mechanistic studies. Evidence of correlation may differ substantially, however, according to whether the claim is single-case or generic. Evidence of

10 For more discussion on the case of legalised abortion and crimes, see §24. For the case of rebellion and violence, see §29.

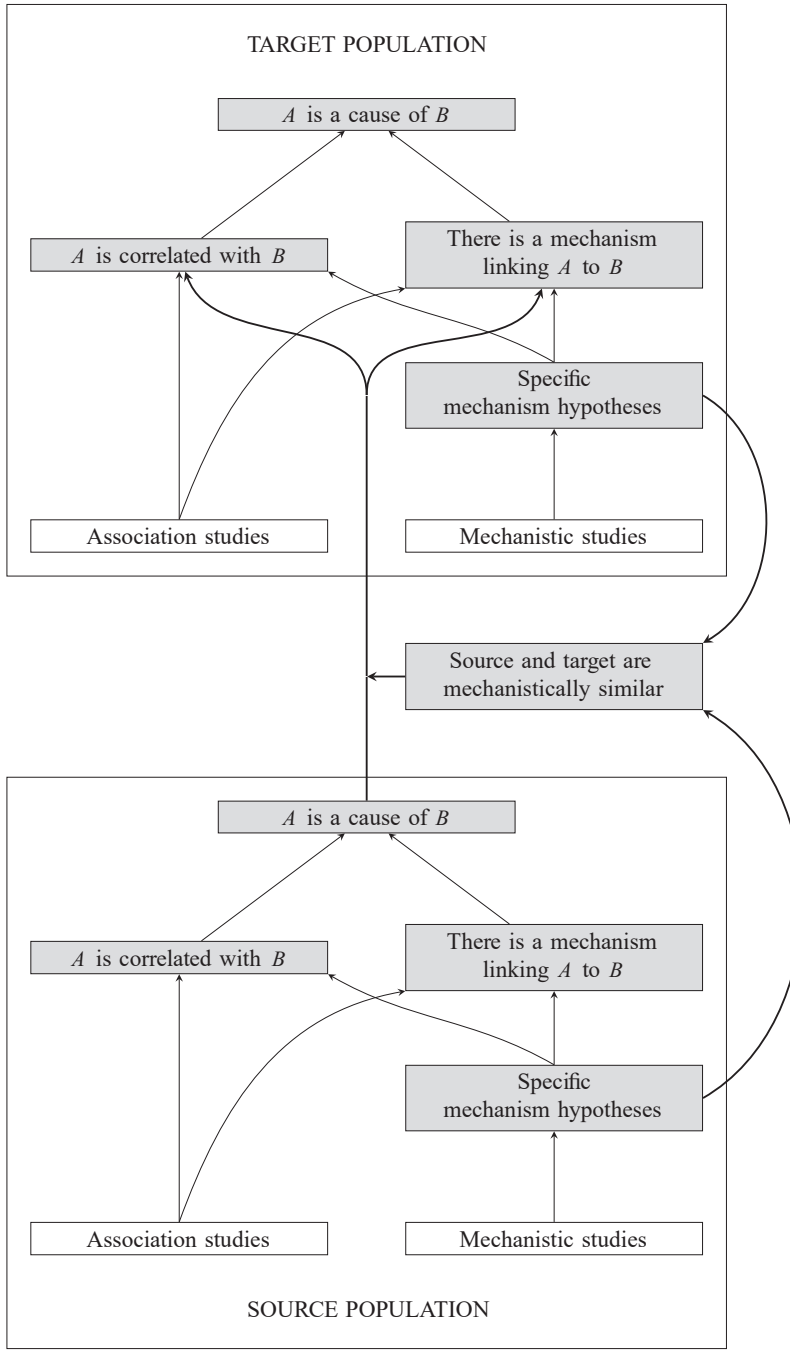

Figure 1.3 Evidential relationships for extrapolating a causal claim (Williamson, 2019a).

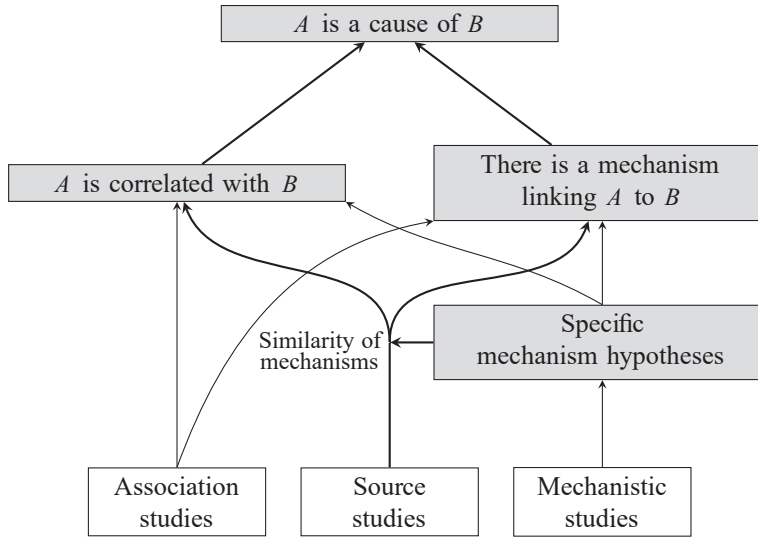

Figure 1.4 Evidential relationships for assessing a causal claim in a target population, including extrapolation from 'source studies', i.e., studies performed in a different source population (Williamson, 2019b). The 'association studies' and 'mechanistic studies' here are studies performed on the target population.

correlation for a generic causal claim is predominantly produced by quantitative association studies that sample a large number of individuals (channel α_1 of Figure 1.1). Evidence of single-case correlation, on the other hand, is often acquired by means of a counterfactual analysis that appeals to specific mechanism hypotheses (channel μ_3), as well as by means of association studies that seek to test a corresponding generic correlation claim (channel α_1). For more detailed discussion of single-case causal claims, see §30.2.

There is a further category of causal claim, namely a universal causal generalisation, which posits that a single-case causal claim holds in each of a set of cases. Runhardt (2022) suggests that multi-method large-*N* qualitative analysis (LNQA) is used to establish this sort of claim, for example. If Evidential Pluralism can be successfully applied to establishing single-case causal claims, then it can also be applied to establishing a universal causal generalisation, which essentially requires establishing multiple single-case causal claims. See §30.3 for further discussion of multi-method LNQA.

§5 Evidential Pluralism and Epistemic Causality

As we have seen, Evidential Pluralism is a purely epistemological thesis, about how to establish and assess causality. However, Russo and Williamson (2007) used the core idea behind Evidential Pluralism to argue for a particular account

of causality itself, namely the epistemic theory of causality. Although this book focuses on the epistemology of causality, we will provide a brief introduction to epistemic causality in this section, because it helps to shed some light on how Evidential Pluralism might cohere with an account of the nature of causality.

Standard theories of causality. Philosophical theories of causality can be divided into several classes. Some can be categorised as *difference-making theories*: the key idea here is that a cause must make some sort of difference to its effects. One such theory—the probabilistic theory—understands '*A* causes *B*' as claiming that *B* is probabilistically dependent on *A*, conditional on *B*'s other causes (Williamson, 2009). Another, the counterfactual theory, understands '*A* causes *B*' as a subjunctive claim: the probability of *B* were *A* not to obtain would be different to the case in which *A* were to obtain (Lewis, 1973). A third, the agency or interventionist account, construes difference-making in terms of manipulation of the cause: changing *A* changes the probability of *B* (Menzies and Price, 1993; Woodward, 2003). These accounts are all similar, in that they require some kind of probabilistic dependence of *A* and *B*.

Certain other philosophical theories of causality can be classed as *mechanistic theories* (Williamson, 2011). One such theory, the process theory, holds that *A* causes *B* just when there is a physical process mediating *A* and *B* that is capable of transmitting a signal from *A* to *B* (Salmon, 1998, Chapter 12). Another, the complex-systems theory, holds that *A* causes *B* just when *A* is involved in an arrangement of entities and activities, organised in such a way as to be responsible for *B* (Glennan, 2017, Chapter 6).

Difference-making theories and mechanistic theories are *monistic*: these theories hold that there is a single causal relation that is the subject of all our causal claims. *Pluralist* theories, on the other hand, take there to be multiple kinds of causing. For example, Hall's dualist approach takes some causal claims to be talking about a difference-making causal relation and others to be referring to mechanistic causality (Hall, 2004). Anscombe is more pluralist, in that she takes each different causal activity, such as scraping and pushing, to be an instance of a different causal relation (Anscombe, 1971). Reiss (2011, 2012) puts forward an inferentialist account, which is radically pluralist. For Reiss, 'causes' means different things in 'inhaling tobacco smoke is a cause of cancer in mice' and 'inhaling tobacco smoke is a cause of cancer in humans' because the evidence base differs in each case, with randomised trials forming a part of the evidence base of the former claim but not the latter (for ethical reasons).

Epistemic causality. The epistemic theory of causality does not fit neatly into any of these classes of theories. According to the epistemic theory, causal claims are not claims about difference-making, nor about mechanisms, nor even about multiple such relations. Instead, our causal claims are simply a device that enables us to make certain predictions, to construct certain explanations, and to infer how to intervene to achieve our ends (Williamson, 2005, §9.4). There is an analogy here with the Bayesian account of probability, which views probabilistic claims

as claims about a form of belief (degree of belief) that allows us to make characteristic inferences—not as claims about some non-epistemic probability relation. Similarly, for epistemic causality, causal claims are claims about a form of belief (a directed belief) that allows us to draw characteristic prediction, explanation and control inferences ('PEC inferences')—they are not claims about some non-epistemic causal relation.

This theory of causality is called 'epistemic' because causality is understood in terms of a kind of belief, and because it gives primacy to causal epistemology.

A causal epistemology can be thought of as a theory or a set of principles that guides our causal beliefs and helps to determine which causal claims are established by the available evidence and which are ruled out by the evidence. Now, some causal epistemologies are better than others—in the context of medicine, for example, the EBM+ approach may well offer certain improvements in comparison to EBM, and, in turn, EBM has advantages over an unsystematic approach to causal evaluation. The hope is that we are progressing towards some optimal causal epistemology. An optimal causal epistemology would be one that best balances several desiderata. It would need to be reliable, in the sense that the causal claims that it deems to be established would need to underwrite successful PEC inferences. It would need to be strong, in that it would need to establish and rule out sufficiently many claims for science to progress efficiently. It would need to be stable, in that it should be very likely that any causal claim deemed to be established (or ruled out) by current evidence remains so in the light of subsequent evidence. It would need to be complete, in that it should determine the status that evidence confers on any putative causal claim, for any body of evidence. It would need to be simple, in that causal generalisations should not be unnecessarily complicated. And it would need to be feasible, in that one should actually be able to use it to evaluate any given causal claim.

The concept of an optimal causal epistemology can be used to provide an account of causal facts. While difference-making theories hold that causal facts are facts about patterns of correlation, mechanistic theories deem them to be facts about mechanisms, and pluralist theories maintain that different causal facts are facts about different causal relations, the epistemic theory presents an alternative account. According to the epistemic theory, the causal facts are just what would be established by any optimal causal epistemology on the basis of evidence that is not lacking in any respect—i.e., on the basis of evidence that includes all particular matters of fact.

Of course, given our limited evidence we don't know all the causal facts. But we have established many causal claims that have yielded successful PEC inferences and that have withstood the test of time—we are very confident that these are causal facts. Similarly, we are confident that many other causal claims that we have ruled out are causal falsehoods. And we can use these claims, as well as hypothetical cases, to test and compare causal epistemologies. We can criticise a causal epistemology by producing counterexamples (claims that the epistemology deems established but which clearly are not causal facts, or claims that the epistemology deems to be ruled out but which clearly are causal facts). Alternatively,

we can criticise a causal epistemology by showing that it fails to make efficient use of evidence in cases where the causal epistemology fails to establish what is clearly a causal fact, or in which it fails to rule out what is clearly not a causal fact. One can also advance or criticise a causal epistemology on the basis of general principles gleaned from our previous experience of determining what is a causal fact and what is ruled out. This is what we did in §2 when motivating Evidential Pluralism, which can be construed as a causal epistemology.

In sum, then, the epistemic theory of causality focuses on causal beliefs as a means to generate PEC inferences, and causal epistemologies as means to guide our causal beliefs. From the point of view of epistemic causality, it is worth developing Evidential Pluralism because it promises better causal beliefs—i.e., causal beliefs that better generate PEC inferences.

Motivation. There are problems with difference-making, mechanistic and pluralist theories of causality, and the epistemic theory avoids these problems. This provides key motivation for epistemic causality. Here we shall sketch some of these problems.

As we saw on p. 12, there are counterexamples to difference-making theories and mechanistic theories: there are causal relationships in which there is no difference-making (overdetermination cases), and there are causal relationships in which there is no mechanism linking cause to effect (causation between absences). These counterexamples present well-known problems for standard monistic theories of causality. We will see, however, that they are not counterexamples to epistemic causality.

Evidential Pluralism poses another sort of problem for difference-making and mechanistic theories (Russo and Williamson, 2007). If causality were mere difference-making, analysable in terms of some sort of correlation, then it would be sufficient to establish the existence of this correlation in order to establish causality and, having established correlation, to ignore evidence of mechanisms. But this would go against Evidential Pluralism and the arguments of §2. So, if we take these arguments to be broadly correct, difference-making accounts of causality cannot be right. On the other hand, if causality were just some sort of mechanistic connection, then it would be sufficient to establish the existence of this mechanistic connection to establish causality and, having established mechanism, to ignore evidence of correlation. Again, this contradicts Evidential Pluralism. One might try a conjunctive theory which holds that A causes B just when there is both a correlation and a mechanism between A and B. While this new theory would be compatible with Evidential Pluralism, it would be undermined by both kinds of counterexample to the standard monistic theories. Thus, the epistemological problem posed by Evidential Pluralism is mitigated at the cost of worsening the problem of counterexamples.

Pluralist theories of causality also face objections. Dualism about causality seems to go against our usage of causal claims: we do not seek to disambiguate a causal claim by asking 'do you mean difference-making-causality or mechanism-causality?' in the way we might disambiguate a probabilistic claim by asking

'do you mean frequency or degree of belief?' Moreover, any pluralist theory of causality which holds that different disciplines or fields appeal to different concepts of cause—e.g., any pluralist theory which holds that the natural sciences appeal to one concept of cause and the social sciences another—faces the problem that we often enquire about the cause of a phenomenon without knowing whether the cause is from one field of study or another. These causal questions seem to make perfect sense, without the need for disambiguation (MacIntyre, 1976, p. 139).[11] The dualist account also suffers from the epistemological problem posed by Evidential Pluralism, because it holds that some causal claims are to be understood in terms of difference-making but not mechanism, while others invoke mechanism-causality rather than difference-making-causality. To establish a causal claim that is construed as difference-making but not mechanistic, it would be sufficient to establish correlation, contra Evidential Pluralism. To establish a causal claim that is construed as mechanistic but not difference-making, it would be sufficient to establish mechanism, again contra Evidential Pluralism. Finally, the dualist account fails to avoid certain counterexamples, namely those that are cases of causation without either difference-making or a linking mechanism (Longworth, 2006).

The inferentialist theory also faces a problem in accounting for our usage of causal claims: our usage does not suggest that the notion of cause varies with systematic differences in the evidence bases for our causal claims, as is claimed by the inferentialist theory. This theory also struggles to accommodate the idea that moving from one causal epistemology to another can lead to improvements in our ability to establish causal claims. This is because different causal epistemologies exploit evidence in different ways, and thus, according to the inferentialist, concern different concepts of cause. If so, changing the causal epistemology changes the kind of claim that we are inferring—it fails to provide a better way to infer the same sort of claim (Williamson, 2013, §6).

The Anscombian pluralist, on the other hand, struggles to account for the unity of our concept of cause. We classify scraping and pushing as instances of causing, not because they stand in a loose family-resemblance relation, but because they both admit characteristic PEC inferences. Consider: scraping the boot against the boot brush is a cause of the absence of mud from the boot. We can predict that the mud will disappear on scraping the boot, we can diagnose the scraping on observing the absence of mud, we can explain the absence of mud by appeal to the scraping, and we can infer that scraping is a means to achieve the end of removing the mud. Similarly, consider: pushing the car is a cause of its starting. We can predict the starting from the pushing or vice versa, we can explain the starting by appeal to the pushing, and we can infer that pushing is a means to starting the car. These distinctive PEC inferences are enough to characterise a single causal relation, contra Anscombe.

11 See Chapter 9 for discussion of a possible exception: in the law, there is some disambiguation of causal talk.

Epistemic causality avoids all the problems outlined here. Firstly, counterexamples to standard theories are not counterexamples to epistemic causality. This is because the epistemic theory of causality does not provide necessary and sufficient conditions for causality in terms of one or both of difference-making and mechanisms. We can see why this is the case as follows. The epistemic theory characterises causal facts in terms of optimal causal epistemologies. If the overdetermination counterexamples are correct, difference-making is not always necessary for causality. Hence, no optimal causal epistemology would deem difference-making to be necessary for causality: in particular, such a causal epistemology would fail to generate reliable PEC inferences in these overdetermination cases. On the other hand, an account that deems a mechanistic connection to be necessary for causality would also be sub-optimal: it would fail to generate appropriate PEC inferences in cases of causation between absences. Thus, if the counterexamples are correct then the epistemic theory of causality could not deem either difference-making or mechanism to be necessary for causality. The epistemic theory could not fall to the usual counterexamples.

Second, Evidential Pluralism poses no epistemological problem for epistemic causality, of the sort it poses for difference-making and mechanistic analyses of causality. This is because epistemic causality takes neither difference-making nor mechanism to be sufficient for causality. We can see this as follows. If Evidential Pluralism is broadly correct, then it is correct because it can be extended to yield some optimal causal epistemology. Evidential Pluralism does not take difference-making to be sufficient for causality, so neither would the optimal causal epistemology that extends it. Thus epistemic causality could not take difference-making to be sufficient for causality. Similarly, epistemic causality could not take mechanistic connection to be sufficient for causality. So, for the very reason that Evidential Pluralism poses a problem for other theories of causality, it poses no problem for the epistemic theory.

Third, the epistemic theory of causality posits a single causal relation, and so is not susceptible to the objection to pluralist theories that they fail to accord with our usage of causal claims—usage that is apparently monistic. Indeed, its appeal to characteristic PEC inferences allows the epistemic theory to explain the unity to our concept of cause, as we saw when considering Anscombian pluralism.

Finally, the epistemic theory is not prone to the objection to the inferentialist theory that stems from from the fact that it fails to accommodate improvements in our ability to establish causal claims. This is because epistemic causality ties the concept of cause to characteristic inferences *from* our causal claims (PEC inferences), not to inferences *to* a causal claim from evidence. For the inferentialist, changing the way we infer from evidence to 'cause' changes the meaning of 'cause'. For the epistemic theory, in contrast, moving from one causal epistemology to another can be an improvement—as long as the causal epistemology we move to better meets the desiderata for a causal epistemology that are laid out here.

We thus see that epistemic causality avoids several key objections that beset other accounts of causality. This provides some motivation for the epistemic

theory. Of course, attempts have been made to respond on behalf of other accounts of causality to some of these objections. For example, Glynn (2011) argues that one version of probabilistic causality can overcome the counterexamples, while Weber (2009) argues that another version is compatible with Evidential Pluralism. It is worth pointing out that even if one is convinced by some such response and one endorses a non-epistemic theory of causality, one may yet think that there is an important role for epistemic causality. This might be on pluralist grounds: analogously, many proponents of the frequency theory of probability acknowledge that there is also a perfectly viable and interesting epistemic concept of probability. Or it might be because one thinks that epistemic causality yields the same causal facts as one's favourite theory: for example, one might think that some version of probabilistic causality yields causal claims that optimally generate PEC inferences. The epistemic theory might then help to provide an account of the uses, unity and epistemology of causality. So, even if one thinks that the objections sketched here can be overcome, epistemic causality may yet be of interest.

This brief introduction to epistemic causality helps to put Evidential Pluralism in perspective. Firstly, if Evidential Pluralism does indeed provide a causal epistemology that improves upon—for example—present-day evidence-based medicine, it should not be regarded as the end of the story. No doubt other improvements can also be made. Second, although Evidential Pluralism is an epistemological theory that says nothing explicitly about the metaphysics of causality nor about our concept or concepts of cause, it can be viewed as imposing important constraints on metaphysical or conceptual theories of causality—constraints that standard or dualist theories may fail to satisfy. Third, Evidential Pluralism is compatible with at least one account of the nature of causality, namely epistemic causality. Whether it is compatible with other accounts remains an open question.

§6 Applying Evidential Pluralism to the Social Sciences

The application of Evidential Pluralism to the social sciences was first mooted by Russo and Williamson (2007, p. 169). Weber (2007) also noted the importance of evidence of mechanisms to the social sciences and suggested that Evidential Pluralism is 'correct' in the social sciences (Weber, 2009, p. 278). Reiss (2009) and Claveau (2012) expressed concerns about the prospects of Evidential Pluralism in the social sciences—concerns that we shall address in §18 and §19 respectively. Moneta and Russo (2014) and Maziarz (2021) argued that Evidential Pluralism can be helpful in econometrics, while Beach (2021) and Runhardt (2022) were sceptical of the application of Evidential Pluralism to political science. Ghiara (2019, Chapter 3) argued that the core idea extends to the social sciences more generally and Shan and Williamson (2021) suggested that Evidential Pluralism can inform evidence-based policy, basic social science research and mixed methods research.

In this book, we develop this research programme in detail. We will argue that Evidential Pluralism can be helpful in the social sciences in three respects:

1 Evidential Pluralism can explain and validate examples of good causal enquiry in the social sciences.
2 Evidential Pluralism can help us understand the structure of causal enquiry in the social sciences.
3 By doing so, Evidential Pluralism can inform practice in the social sciences.

In Chapter 2, we compare Evidential Pluralism to historical precursors and to related positions in the social sciences. Then, in Part II, we show how Evidential Pluralism helps to reconceive two important social science methodologies: it motivates a new approach to evidence-based policy that is analogous to the EBM+ approach to evidence-based medicine (Chapter 3) and it provides new, metaphysics-free foundations for mixed methods research (Chapter 4). In Chapter 5, we respond to potential objections to the application of Evidential Pluralism to the social sciences.

In Part III, we show how Evidential Pluralism can be fruitfully applied to a selection of social sciences: sociology (Chapter 6), economics (Chapter 7), political science (Chapter 8) and law (Chapter 9). In each case, we argue for claims 1–3 as set out earlier. In Chapter 10, we argue that the benefits of Evidential Pluralism extend to other social sciences, and we note some questions for further research.

2 Historical Roots

Although Evidential Pluralism arose out of the thesis of Russo and Williamson (2007), the interplay between correlation and mechanisms had previously attracted interest in some quarters. In this chapter, we shall examine some precursors to Evidential Pluralism: the positions of Claude Bernard (§7), W.F.R. Weldon (§8) and John Goldthorpe (§9). In §10, we compare Evidential Pluralism to these positions, as well as to analytic sociology and critical realism.

§7 Bernard

Claude Bernard (1813–1878), a French physiologist, championed what he called 'experimental medicine'. He distinguished between three views of medicine, which can be broadly understood as follows. One view, which Bernard labelled 'expectant medicine', or 'Hippocratism', bases medicine on observation and is largely prognostic:

> Among physicians, there are some who actually believe that medicine should remain a science of observation, i.e., that it should be able to foresee the course and outcome of diseases, but should not directly act on disease. There are others, and I am one of them, who think that medicine can be an experimental science, i.e., that it should delve into the interior of organisms and find ways of altering and, to a certain extent, regulating the hidden springs of living machines. Observing physicians look on a living organism as a little world contained in the great world, like a kind of ephemeral living planet whose motions are ruled by laws which we discover by simple observation, so as to foresee the progress and evolution of vital phenomena in health or disease, but without ever being able to alter their natural course in any way. This doctrine is found in Hippocrates in its purest form. Medicine of simple observation obviously excludes all manner of active medical intervention; for this reason it is also known as *expectant medicine*, that is to say, medicine that observes and foresees the course of diseases without aiming to act directly on their progress.
>
> (Bernard, 1865, p. 197)

DOI: 10.4324/9781003143000-3

A second view, 'empirical medicine', employs both observation and experimentation, and seeks to control phenomena as well as predict them:

> It is rarely that we find a physician purely Hippocratic in this respect, and it would be easy to prove that many physicians, who loudly applaud Hippocratism, do not trust to its precepts in the least when they give themselves up to the most active and disordered flights of empirical medication. Not that I condemn these therapeutic attempts which, most of the time, are only experimentations *to see*; only I say that this is not Hippocratic medicine, but empiricism. Empirical physicians, acting more or less blindly, are, after all, experimenting on vital phenomena, and thus class themselves in the empirical period of experimental medicine.
>
> (Bernard, 1865, p. 197)

While expectant medicine and empirical medicine can be thought of as focusing on associations, the third view, 'experimental medicine', takes mechanisms seriously as well, with a view to prediction, explanation and control:

> Experimental medicine is therefore medicine that claims knowledge of the laws of healthy and diseased organisms, not only so as to foresee phenomena, but also so as to be able to regulate and alter them within certain limits. . . . physiology must be constantly applied to medicine, if we are to understand and explain the mechanism of disease and the action of toxic and medicinal agents. . . . The first requirement, then, in practising experimental medicine, is to be an observing physician and to start from pure and simple observations of patients made as completely as possible; experimental science comes next, analyzing every symptom by trying to connect it with explanations and vital laws that shall include the relation of the pathological state to the normal or physiological condition.
>
> (Bernard, 1865, pp. 196–197)

Note that experimental medicine does not replace evidence of associations with evidence of mechanisms—it combines them:

> empirical medicine and experimental medicine are far from being incompatible, but on the contrary must be intimately united; for both are indispensable in building up experimental medicine.
>
> (Bernard, 1865, p. 218)

One can thus construe Bernard as saying that medicine needs to consider both association and mechanistic studies. However, throughout much of his writing it is apparent that Bernard's interest in mechanisms stems primarily from the desire to explain and understand previously established causal relationships, rather than to help establish them in the first place. This does not as yet constitute Evidential Pluralism, which brings association and mechanistic studies together for the

specific purposes of establishing and assessing a causal claim. But there is some evidence that Bernard does hold that mechanistic studies influence confidence in causation, in accord with Evidential Pluralism:

> Experimenting physicians . . . will make use of all the therapeutic means advised by empiricism; only instead of using them according to authority and with a confidence akin to superstition, they will administer them with that philosophic doubt which is appropriate to true experimenters; they will verify the results on animals, and by comparative observations on man, so as to determine rigorously the relative influence of nature and of medicine in curing disease.
>
> (Bernard, 1865, p. 211)

We can interpret this as the claim that empirical studies can provide sufficient grounds for administering a treatment, but that the effectiveness of the treatment is not established unless confirmed by mechanistic studies. This interpretation is confirmed by the fact that Bernard views empirical medicine as merely conjectural: empirical medicine . . . is conjectural medicine because it is based on statistics which collect and compare cases that are analogous or more or less similar in their outer characteristics, but undefined as to their immediate causes. (Bernard, 1865, p. 214)

In sum, while it would be a stretch to call Claude Bernard an Evidential Pluralist in the precise sense outlined in §1, it is clear that his views are not very far from Evidential Pluralism, and that his desire to move from empirical medicine to experimental medicine is analogous to the Evidential Pluralist's urge to move from EBM to EBM+.

§8 Weldon

W.F.R. Weldon (1860–1906), an English biologist, also emphasised the significance of correlation and mechanisms, but in relation to the study of inheritance. In his unpublished manuscript *Theory of Inheritance*, Weldon was explicit about his methodology:

> The student of heredity has two main objects: the first is to discover what degree of stability is actually exhibited by the various races of animals or of plants, and to determine the extent to which deviation from the average characters of parents or other ancestors is associated into deviation in their descendants; the second object is to acquire such knowledge of the changes which occur during the growth and maturation of the germ-cells, their fusion and subsequent development, as may serve to indicate the process by which the obscure relation between parents and filial characters is brought down. The first object is to make a purely descriptive statement of the actual relation between the visible bodily characters of living things and those of their ancestors or their descendants; the second is to learn the process to which this

> relation is due. These two objects are pursued by different methods, and as it happens they are generally pursued by different men, so that few attempts have been made to consider the learning of what are actually known concerning relation between the visible characters of parents and those of their offspring upon the possible interrelation of structural changes revealed by minute study of the germ-cells and of embryonic processes in germinal.
>
> (Weldon, 1905, p. f.3.r)

For Weldon, the statistical approach was fundamental. The first step in the study of inheritance was to make an accurate statistical description of the pattern of inheritance. This echoed Weldon's early view: the problem of inheritance was 'a problem in statistics' (Weldon et al., 1901, p. 3). Nevertheless, the statistical approach, though important, was not the only approach to the study of inheritance. Unlike Pearson (1898), Weldon did not regard a purely statistical description of the pattern of inheritance as a complete theory of inheritance. For Weldon, another important and indispensable task of the study of inheritance was to look for the mechanism of inheritance. As the title of the series of *Lancet* reports 'Current Theories of the Hereditary Process' suggests, what Weldon was interested in was not only the pattern of inheritance, but also the process (or mechanism) of inheritance ('the hereditary process').[1] For example, one of the central tasks in *Theory of Inheritance* was to study the phenomena of dominance. It should be noted that Weldon's conception of dominance was completely different from the Mendelian notion, which referred to a property of hereditary elements (Shan, 2020, p. 62). For Weldon, dominance referred to a process by which a hereditary determinant gives rise to a visible character. Weldon began with Galton's statistical approach, but he found Galton's statistical law of dominance inadequate:

> Galton's theory . . . leads us to consider a limit to the generality of the statistical law of dominance founded on the particular case examined and on others like it.
>
> (Weldon, 1905, p. f.71.r)

Weldon maintained that identifying the mechanism underlying the phenomenon of dominance was as important as making a statistical description of the patterns of dominance.

> What Galton has so far given us is in the first place a theory of hereditary transmission by means of determinant elements, each capable of assuming a condition in which it is dominant, affecting the visible character of the body, or a condition in which it is latent and is transmitted from the given of one generation to that of the next, without affecting the body through which it passes;

1 'Current Theories of the Hereditary Process' consists of eight brief reports of a series of Weldon's lectures at University College London from 1904 to 1905.

> in the second place he has formulated a statistical law of the dominance of such determinants, expressing the facts of inheritance, whether blended or alternative, for races in which mating occurs at random, so far as the characters studied are concerned. The validity of this law has not been disputed by anyone who has dealt with similar cases of random mating since it was first published. . . . a knowledge of the modifications which occur in particular cases may be expected to throw considerable light on the conditions by which the dominance or the latency of germinal elements is determined.
>
> (Weldon, 1905, p. f.72.r)

Throughout the manuscript, Weldon attempts to look for the mechanism of dominance, which was expected to be compatible with the findings of the statistical studies. Pence (2011) suggests that Weldon seemed to hold a probabilistic conception of causation: for Weldon, 'the only way in which we may accurately claim "causal" knowledge of a system, without destructive simplification, is to point to correlations within the system as a whole' (Pence, 2011, p. 483). It seems that Weldon held a Humean regularity conception of causation in the 1890s. For example, Weldon was explicit on the point that 'when I have spoken of cause and effect, I have always endeavoured to use the words in accordance with the definition given [by Hume]' (Weldon, 1896, p. 294). Nevertheless, Weldon's view on causation changed in the early 1900s. As Shan (2020, pp. 65–66) argues, Weldon's conception of causation, especially in his post-1904 works, might be more in line with Evidential Pluralism than with an approach that focuses on association. In order to investigate the phenomenon of dominance, Weldon identified two tasks: the first is 'to make purely descriptive statement of the actual relation between the visible bodily characters of living things and those of their ancestors on their descendants' (correlation), and the second is 'to learn the process to which this relation is due' (mechanism) (Weldon, 1905, p. f.3.r). Arguably, then, Weldon implicitly suggested that one needs both evidence of correlation and evidence of mechanisms in order to determine what causes a dominant character.

§9 Goldthorpe

The British sociologist John Goldthorpe (1996, 1998, 2001) developed a hybrid approach to causal enquiry in sociology.

From a sociologist's point of view, he critically examined three approaches, developed by statisticians: the dependence approach, the manipulation approach and the mechanistic approach. The dependence approach originated from the economist C.W. Granger's work in the context of the analysis of econometric time-series. The basic idea is that, in order to establish a causal claim that A causes B, one needs to show that there is some robust observed association between A and B, where A is temporarily prior to B. However, Goldthorpe argued that one key problem for the dependence approach was its explanatory inadequacy. If causation is merely about robust dependence or stable association, then establishing a causal claim in sociology amounts to a statistical inference for the purpose

of forecasting or prediction. However, such an approach to establishing causal claims is inadequate: sociology is not only about prediction, but also about explanation. Thus, a causal claim in sociology should provide not only prediction but also explanation. Consider a sociological study in which a correlation between educational attainment and level of income is established. It would be too hasty for sociologists to conclude that education causes income. Much more needs to be studied in order to establish a causal claim about education and income. As Goldthorpe (2001, p. 4) elaborated, 'To establish a causal link between education and occupation or income would then require, in the first instance, situating the variable of "educational attainment" within some generalized narrative of action which would represent one or other such process that is of a "causally adequate" kind. And in the interests of clarity, consistency, and subsequent empirical testing, it would then be further desirable that any narrative thus advanced should be not merely ad hoc but rather one informed by a reasonably well-developed theory of social action'.

The manipulation approach rests on an interventionist account of causation, which characterises causation in a more experimental manner (Rubin, 1974; Cook and Campbell, 1979; Holland, 1986). Accordingly, a causal claim is established if it is experimentally shown that the effect is controllable by manipulating the cause. Goldthorpe identified two main problems with the manipulation approach. One concerns the assumption that the causal variables are experimentally manipulable. This implies sociologists should not make any causal claim involving non-manipulatable variables. As Goldthorpe (2001, p. 6) observed, 'one could discuss the association that exists between sex or race, on the one hand, and say, educational attainment, on the other. But it would be no more meaningful to speak of sex or race as being causes of such attainment than it would be to make statements about what level of education Ms M would have achieved had she been a man or Mr N had he been a woman'. Thus, if the manipulation approach is the only correct approach to causal inference, many causal claims in sociology turn out to be inadmissible. In addition, there is a concern about the distinctive nature of the response of individuals in sociological experiments. The manipulation approach was originally introduced in the context of the applied natural and agricultural sciences (Cook and Campbell, 1979), and causal relationships in the social sciences can behave very differently. In principle, the manipulation approach allows conceptual space for human action only in the roles of experimenter or intervener. For example, in an experiment to test a fertiliser, the experimental set-up is the only source of intervention. Once the experiment is conducted, all else has to follow in the manner of plants responding to the fertiliser. But in sociology, the response of the units in experiments cannot simply be assumed to have the same nature as that of the units in experiments in the applied natural and agricultural sciences. Consider a case of the introduction of some positive discrimination in education, with the aim of reducing class or ethnic differentials in achievement. Members of those classes or ethnic groups whose children would not benefit and who might lose their competitive advantage in schools may respond in order to preserve their advantage. In

this case, one crucial requirement of randomised experimental design would be breached: the response of a unit should not be influenced by which units are treated.

The process approach assumes that causation is a generative process (or a mechanism) from the cause to the effect (Simon and Iwasaki, 1988; Freedman, 1991; Cox, 1992). Accordingly, it has been argued that a causal claim cannot be established without establishing the existence of a generative process. The process approach was introduced to respond to the problems of the dependence approach and the manipulation approach. It purports to provide greater explanatory power than the two other approaches by means of its appeal to processes or mechanisms. The basic idea is that by identifying the process (or mechanism) that underlies a robust dependence, the process approach yields a better understanding of causation. Thus, Goldthorpe (2001, p. 9) argued that the process approach was usually 'a necessary augmentation' of the dependence approach and the manipulation approach. In other words, the process approach is not a genuine alternative to the dependence approach and the manipulation approach.

In conclusion, Goldthorpe found none of these approaches adequate to analyse and assess causal claims in sociology. Instead, Goldthorpe (2001) proposed an alternative approach. For Goldthorpe, in order to establish a causal claim in sociology, one has to:

> (i) establish the phenomena that form the explananda;
> (ii) hypothesise generative processes at the level of social action;
> (iii) test the hypotheses.
>
> (Goldthorpe, 2001, p. 10)

According to Goldthorpe (2001, p. 10), establishing the phenomena is 'an essentially descriptive exercise' and 'achieved statistically'. In other words, the first step of establishing a causal claim in sociology is to provide a statistical description of some social regularities. Then, a hypothesis 'must be provided', which 'purports to capture the central tendencies' and to show how these central tendencies 'would, if operative, actually give rise, through their intended and unintended consequences, to the regularities' (Goldthorpe, 2001, p. 12). In order to test the adequacy of the hypothesis, one needs to examine whether the hypothesised generative process would be capable of producing the regularity in question. One may further test the hypothesis in an indirect way by examining 'other effects to which the process should give rise apart from those constituting the regularities' (Goldthorpe, 2001, p. 13).

Therefore, it is clear that Goldthorpe highlighted the significance of both correlation and mechanisms for causal enquiry in sociology. Goldthorpe explicitly rejected the view that evidence of correlation alone suffices to establish causality.

> [I]nstead of being regarded as a means of inferring causation directly from data, [the primary use of statistical technology] should rather be seen as

> descriptive, involving the analysis of joint and conditional distributions in order to determine no more than patterns of association (or correlation). Or, at very most, representations of the data might serve to suggest causal accounts, which, however, will need always to be further developed theoretically and then tested as quite separate undertakings.
>
> (Goldthorpe, 2001, p. 11)

Although he did not use these terms, Goldthorpe's view is in line with the thesis that both association studies and mechanistic studies are important for causal enquiry. Association studies play a key role in establishing the phenomena and testing hypotheses, while mechanistic studies shed light on generative processes.

§10 How does Evidential Pluralism differ?

Bernard, Weldon, and Goldthorpe all emphasised the significance of correlation and mechanisms in scientific practice, but their views nevertheless differ from Evidential Pluralism. Bernard's primary concern was to promote the study of mechanisms in medicine in order to explain, rather than establish, causal relationships. Weldon's aim was to develop a new theory of inheritance by studying both the statistical pattern and the mechanism of inheritance. Neither Bernard nor Weldon explicitly talked of causation in terms of correlation and mechanisms.

In comparison, Goldthorpe's approach is closest to Evidential Pluralism. Both Goldthorpe and Evidential Pluralism emphasise establishing a correlation and a mechanism in order to establish a causal claim in sociology. Moreover, both maintain that establishing a causal claim is fallible. As Goldthorpe (2001, p. 15) argues, 'empirical evaluations of [causal mechanistic hypotheses] are not expected to achieve once-and-for-all verification'. That said, Evidential Pluralism does differ from Goldthorpe's approach in several significant respects. Firstly, the rationale behind Goldthorpe's approach is the hypothetico-deductive (H-D) model of confirmation. It begins with the proposal of a causal hypothesis based on empirical data and is followed by the confirmation of this hypothesis. However, the H-D model is not required by Evidential Pluralism. According to Evidential Pluralism, one can perfectly well establish a causal claim by means of methods other than the hypothetico-deductive method. For example, specific mechanism hypotheses might be gleaned inductively from data. Second, the initial step in Goldthorpe's approach is to establish the phenomena that form the explananda, which is 'an essential descriptive exercise' (Goldthorpe, 2001, p. 10). Such a descriptive exercise is implicitly rooted in a distinction between observation and theory. However, the observation-theory distinction is not presupposed by Evidential Pluralism. Third, Goldthorpe's approach imposes a particular order of activities in causal enquiry, namely from (i) to (iii). In contrast, Evidential Pluralism does not insist on any particular ordering of enquiry. What Evidential Pluralism provides is a normative account of the evidence that is needed in order to establish a causal claim. It does not require specific

methodological rules to guide or constrain the practice of evidence-gathering. Whether evidence of correlation or evidence of mechanisms should be obtained first does not really matter, in general.[2]

Evidential Pluralism shares a focus on mechanisms with another well-entrenched approach in the social sciences, namely analytic sociology (Hedström and Swedberg, 1998; Hedström and Bearman, 2011; Hedström and Ylikoski, 2014).

Peter Hedström and Petri Ylikoski elaborate the key idea behind analytic sociology as follows:

> [Analytic sociology] is founded on the idea that social sciences should do more than describe and classify social processes. According to analytical sociologists, the primary epistemic aim of the social sciences should be causal explanation of social phenomena. Sociological theory should aim to develop clear and precise accounts of the social mechanisms by which the intentional activities of social agents bring about social phenomena.
>
> (Hedström and Ylikoski, 2014, p. 386)

One of the central features of analytic sociology is its emphasis on mechanisms. This puts analytic sociology at odds with a quantitatively-oriented approach to sociology that largely shuns the role of theory in social explanation and prefers to stay close to empirical data. It is worth noting that analytical sociology does not oppose the use of quantitative data but it does maintain that sociological research should be more theory-driven and that causal claims should be supported by a theoretical understanding of the mechanisms underlying observed statistical regularities (Hedström and Ylikoski, 2014, p. 386).

There are some important similarities between Evidential Pluralism and analytic sociology. Both are dissatisfied with a purely association-based approach to causation. Moreover, both maintain that mechanisms play an important role in enquiry. That said, Evidential Pluralism differs from analytic sociology in a key respect: mechanisms are crucial to Evidential Pluralism for the purposes of establishing and assessing causation, while their role in analytic sociology is primarily to explain social phenomena. Thus the primary use of mechanisms in analytic sociology is in line with Bernard's primary use of mechanisms: explanation. In contrast, mechanisms are important to Evidential Pluralism insofar as they help to confirm causal claims.

As in the case of Bernard, however, analytic sociologists do sometimes take mechanisms to play an ancillary justificatory role:

2 Where the status quo involves searching for and evaluating association studies, this can indeed provide a natural first step to an evaluation based on the principles of Evidential Pluralism, as we shall see in Chapter 3. Similarly, process tracing sometimes takes its point of departure to be a previously established correlation (§30.2), and multi-method large-*N* qualitative analysis tends to begin by establishing correlation (§30.3). Such orderings are based on historical contingency, however, and are not essential to Evidential Pluralism.

> [T]he information about the causal mechanisms . . . provides justification for causal claims. Causal claims are much easier to accept if one can provide an account of the mechanisms by which the changes in the suggested explanans bring about the changes in the explanandum.
>
> (Hedström and Ylikoski, 2014, p. 390)

Evidential Pluralism can be viewed as providing a more explicit account of this justificatory role of mechanisms in causal enquiry. This account is complementary to the account provided by analytic sociology of the use of mechanisms for explanation.

Evidential Pluralism is also related to Roy Bhaskar's views on causation, which have been influential in social sciences such as economics and international relations. Bhaskar's view on causation is rooted in his transcendental realism, subsequently known as 'critical realism'. Mechanisms are core to this view. The key idea behind transcendental realism is that the central objects of knowledge are the structures and mechanisms that generate phenomena. Accordingly, 'the aim of science is the production of the knowledge of the mechanisms of the production of phenomena in nature that combine to generate the actual flux of phenomena of the world' (Bhaskar, 1975, p. 6). Moreover, for Bhaskar, causal laws are by nature generative mechanisms (Bhaskar, 1975, p. 3).

It is clear that Bhaskar's view of causation is rather different from Evidential Pluralism. First, Bhaskar presupposes a specific metaphysical account of causation, while Evidential Pluralism is an account of the epistemology of causation. Second, Bhaskar maintains that a causal law is established if a generative mechanism is established, while Evidential Pluralism emphasises the significance of both evidence of correlation and evidence of mechanisms to causation. We will encounter some further differences between Evidential Pluralism and approaches based on critical realism in §13.

In sum, the complementary nature of correlation and mechanism has been recognised in the natural and social sciences since at least the nineteenth century. Evidential Pluralism can be viewed as a recent development of this tradition.

Part II

Consequences and Concerns

3 Evidence-Based Policy
EBP+

Our first example of a methodological development that is motivated by Evidential Pluralism relates to evidence-based policy (EBP). In this chapter, we argue that the move from EBM to EBM+ (discussed in §3) warrants an analogous move from present-day EBP to EBP+, a new approach to policy appraisal which takes evidence of mechanisms more seriously. In §11, we introduce EBP and EBP+. In §12, we set out the steps by which an EBP+ evaluation can proceed. Finally, in §13, we discuss the relationship between EBP+ and other approaches to evaluation.

§11 EBM and EBP

In the 1990s, the methods of evidence-based medicine quickly spread to the evaluation of social interventions, leading to what is now known as evidence-based policy. The Cochrane Collaboration, which promotes EBM, was set up in 1993, while the Campbell Collaboration, which promotes EBP along similar lines, was created in 1999.

In the UK, for example, the primary organ of EBP is the government-led 'What Works Network', which includes the National Institute of Health and Care Excellence (NICE) as well as a dozen other centres tasked with evaluating social interventions. The What Works Network is built around the use of association studies (in particular, RCTs) as the evidence on which to base an evaluation (What Works, 2018a, p. 4), and membership of the network is restricted to centres which share the ranking of evidence promulgated by present-day EBM/EBP (Cabinet Office, 2018, p. 3). Allied to the What Works Network is the UK Government Trials Advice Panel, which was set up in 2015 to promote the use of RCTs in public policy decision making (What Works, 2018b). These structures ensure that the monistic methods of present-day EBM/EBP are entrenched at the heart of policy making in the UK.

The situation in the UK is just one instance of a global phenomenon. In the US, the dominant approach to EBP is also modelled on EBM, with a focus on RCTs (Baron, 2018). In addition, the United Nations actively promotes a global vision of EBP based on statistical association studies—see, e.g., United Nations (2013). EBP and EBM continue to develop hand-in-hand: e.g., the Cochrane and Campbell Collaborations share methods at events such as the 'Global Evidence Summit', held in 2017 in Cape Town and 2024 in Prague.

DOI: 10.4324/9781003143000-5

Thus EBP is modelled on EBM, which, as we noted in §3, underestimates the importance of mechanistic studies. Given this, there is arguably a need for EBP+, i.e., an analogue of EBM+ but applied to policy evaluation. As with EBM+, the aim of EBP+ is to provide methods for systematically assessing mechanistic studies and integrating these assessments with those of association studies in order to determine the status of a causal claim.

The need for EBP+ arises because Evidential Pluralism applies equally to medicine and policy making. If Figure 1.1 captures the key evidential relationships when evaluating a causal claim, then it applies as much to the social sciences as it does to medicine. Indeed, one cannot draw a sharp distinction between causal enquiry in medicine and causal enquiry in the social sciences. This is because causal claims often overlap the health and social sciences: health policy interventions are interventions in both medicine and social policy; hence the inclusion of NICE in the What Works Network. Insofar as one can generalise, the main methodological difference between the biomedical sciences and the social sciences is that in the social sciences it can be harder to isolate an experiment from contextual factors that might influence its results and that can thwart replication. In many cases, it can also be harder to properly randomise individuals to social policy interventions, to construct a placebo intervention for a control group, and to ensure adherence to the social policy interventions being tested. All these considerations favour a shift away from the almost exclusive reliance on RCTs exhibited by present-day EBP and towards Evidential Pluralism and EBP+.

EBP+ is also required for successful extrapolation. If Figure 1.3 captures the key evidential relationships when extrapolating a causal claim from a source population to a target population, then it applies as much to the social sciences as it does to medicine. In both medicine and the social sciences, it is only by considering mechanisms of action that one can decide whether a causal relationship discovered in a study context can be extrapolated to a target context of application (see, e.g., Steel, 2008; Wilde and Parkkinen, 2019).[1] In fact, it can be harder to successfully extrapolate a claim about the effectiveness of a social intervention from one setting to another than to extrapolate a claim about the effectiveness of a medical intervention. This is because social settings can vary much more widely in their social mechanisms than do human bodies in their pathophysiological mechanisms. The mechanisms responsible for a particular social phenomenon can vary widely from context to context; consequently, mechanisms responsible for the failure of these social mechanisms can vary widely, necessitating variation in the mechanism of action of an intervention to rectify such failures. In addition, contextual mechanisms (mechanisms that counteract or reinforce the mechanism of action of the intervention) can vary widely in the social realm. Only by learning about these mechanisms can one determine whether extrapolating an intervention

1 Cartwright and Hardie (2012, §1.A.1.1) provide a nice example of the importance of extrapolation to public health policy. They also criticise the way in which EBP focuses almost exclusively on association studies—in particular on RCTs. Evidential Pluralism can be thought of as a principled way of addressing some of these concerns with extrapolation and RCTs.

to a new context is likely to be successful. This consideration favours a shift towards Evidential Pluralism, which takes evidence of mechanisms seriously, and away from the reliance of present-day EBP on association studies.

§12 EBP+ Evaluation Procedures

In this section, we set out the main steps by which an EBP+ evaluation can proceed. The approach taken here develops that of Parkkinen et al. (2018), which was applied to evidence-based medicine.

This section will focus on the procedural recommendations of EBP+, rather than its justification, which has already been sketched in §2.[2] These procedural recommendations will be presented at a very general level here. This is because the uses of evidence-based policy vary enormously and there is a danger that more specific advice will be appropriate to some situations but not others. The advantage of taking this bird's-eye view is that the overall logic of evaluation becomes clearer from such a height.

The first step to any EBP+ evaluation will be to clearly specify the main claims that need to be assessed. This step is discussed in §12.1. The status of each of these claims will depend on the quality of the available evidence. Status and quality are discussed in §12.2.

Recall that Evidential Pluralism draws a distinction between evaluating causation in a study population—evaluating *efficacy*—and evaluating causation in a target population of interest—evaluating *effectiveness*. These terms are usually used to apply to the evaluation of interventions, but Evidential Pluralism extends their use to causal claims in general. It is often the case that there are too few high-quality studies on the target population to establish effectiveness directly. In such a situation, it is natural to try to establish effectiveness indirectly by establishing efficacy in some suitable study population and then extrapolating the causal claim from this source population to the target population. Establishing efficacy in a study population is sometimes referred to as establishing the *internal validity* of the causal claim in the study population. If the causal claim can be successfully extrapolated to a target population then its *external validity* can be said to be established in the target population.

This indirect route to establishing effectiveness can be summarised as:

Effectiveness = efficacy + external validity

The principal steps to an EBP+ assessment of efficacy are laid out in §12.3. §12.4 provides a guide to the assessment of external validity. §12.5 puts these tasks together to provide an account of the assessment of effectiveness.

2 Grüne-Yanoff (2016) and Marchionni and Reijula (2019) provide specific motivation for considering evidence of mechanisms in the context of policy assessment, and Stegenga (2022) argues that the use of evidence of mechanisms to assess the effectiveness of interventions can be motivated on Bayesian grounds.

§12.1. Specifying the key claims

According to Evidential Pluralism (as depicted in Figure 1.1, for example), assessing causality requires assessing four different kinds of claims: (i) the causal claim that *A* is a cause of *B*; (ii) the correlation claim, that *A* and *B* are probabilistically dependent conditional on the potential confounders; (iii) the general mechanistic claim, that there is a complex of mechanisms which invokes *A* as partially responsible for *B* and which can account for the extent of the correlation; and (iv) specific mechanism hypotheses, that posit features of such a mechanism complex. Let us consider each of these claims in turn.

(i) The causal claim. Firstly, we need to be clear about what *A* and *B* are. In the evidence-based policy setting, the putative cause *A* will typically be a policy intervention and the putative effect *B* an outcome of interest, which might be an intended goal of the policy or a potential harm, for instance. When scrutinising an evaluation it is important to check that *B* really is the outcome of interest, rather than a proxy that is more easily measured or that has been more widely studied. Often, short-term outcomes are measured when longer-term outcomes may be of more relevance, just because the short-term outcomes are easier to study. Sometimes, there is a political agenda to the choice of outcome variable: a public body that wants to stop using a costly intervention may choose a less relevant outcome variable for which the intervention is ineffective to argue that the intervention is ineffective simpliciter. Alternatively, an organisation that wants to claim that a particular intervention is 'evidence-based' may choose a less relevant outcome variable with respect to which effectiveness is more easily demonstrated.

A and *B* will normally be repeatedly instantiable, i.e., the policy and the outcome of interest will be of a kind that can be instantiated in a range of different contexts, such as in different geographical areas. *A* might be a binary variable, which takes the value 'true' if the policy is instantiated and the value 'false' if no new policy intervention is made, or alternatively, if some other well-specified intervention is carried out. The outcome variable *B* may be a binary variable, or may take a broader range of possible values. For example, in the case of a skills instruction intervention on members of a workforce, an outcome of interest might be average change in salary over the subsequent 10 years, which could take a wide range of values, positive and negative. While evidence-based policy is usually concerned with evaluating repeatedly instantiable interventions, there are situations in which it is a particular implementation of an intervention that is of primary interest. In such situations, *A* and *B* will be single-case variables and the question is whether that particular intervention was a cause of a particular observed outcome.

It is also essential to specify the population within which the causal claim is supposed to obtain. In the context of evidence-based policy, this is a population within which the policy intervention *A* is supposed to lead to outcome *B*. Recall that if this is the population upon which one is interested in intervening, it is called the *target population*. If it is some other population which has been more widely studied than the target population, then it is usually referred to as a *study population* or *source population*. Note that if the causal claim is single-case, then the population has

only one member. It is of course essential to clearly identify this member, in order to avoid ambiguity in the causal claim.

(ii) The correlation claim. We need to be clear about what the correlation claim says. If *A* and *B* are repeatedly instantiable, the correlation claim is most naturally interpreted as a frequency claim: that certain values of *B* occur more frequently in the presence of the intervention than in its absence, when controlling for potential confounders. Of course, this requires specifying what the potential confounders are. Recall that these are variables which, according to mechanistic and causal background knowledge and previous studies, might reasonably be anticipated to be causes of *B* that are significantly correlated with *A* (e.g., common causes of *A* and *B*). Any particular association study will measure this correlation conditional on some subset of potential confounders, and a range of association studies can be used to estimate the correlation conditional on all potential confounders. RCTs are often thought to be particularly informative in this regard, because they can be used to provide an unbiased estimate of the 'average treatment effect', which is one way of quantifying the association between intervention and outcome.[3]

If *A* and *B* are single-case, on the other hand, then the correlation claim is most naturally interpreted as a claim about chance or Bayesian rational degree of belief. Under a chance interpretation, the correlation claim asserts that the chance of the observed outcome was greater having carried out the intervention than it would have been in the absence of the intervention, when holding fixed the values of potential confounder variables. Under a Bayesian interpretation, it asserts that one should have believed to a greater extent that the observed outcome would occur having carried out the intervention than had it not been carried out, holding fixed potential confounders. Either way, there is an important counterfactual element to a single-case correlation claim. Evidence for such a claim can be obtained from the sorts of association studies that are used to confirm a frequency-based correlation claim. Additionally, well-confirmed mechanism hypotheses will typically play an important role here, because mechanistic information is often required to underwrite counterfactual inferences (channel μ_3 of Figure 1.1).

(iii) The general mechanistic claim. The general mechanistic claim is an existence claim: it asserts that there exists a complex of mechanisms which explain instances of *B* by appeal to instances of *A* and which can account for the extent of the correlation. Confirming this claim does not require establishing details of the relevant mechanisms (channel μ_2 of Figure 1.1)—the claim can also be confirmed indirectly via channel α_2. Recall that a mechanism can be thought of as composed of mechanistic processes, often representable as chains or networks of mediating variables that help to account for the value that *B* takes, and complex-systems mechanisms, i.e., collections of entities and activities organised in such a way as to be partly responsible for the value that *B* takes (Illari and Williamson, 2012).

3 However, as Deaton and Cartwright (2018); Twisk et al. (2018) and Ye et al. (2014) argue, it is important not to read too much into this virtue of the RCT approach.

What we are interested in is the whole complex of mechanisms linking *A* and *B*, including the mechanisms of action of the intervention as well as any counteracting or reinforcing mechanisms. Can this complex of mechanisms as a whole account for the magnitude of the observed correlation?

(iv) Specific mechanism hypotheses. Specific mechanism hypotheses postulate features of a mechanism complex linking *A* and *B*. These features may include mediating variables, entities, activities or organisational features. Mechanistic studies are required to confirm or disconfirm such features (channel μ_1 of Figure 1.1). A specific mechanism hypothesis may include features that have been previously established, however. When all the features posited by the hypothesis are established, then the hypothesis itself is established.

§12.2. Quality of evidence and status of a claim

An EBP+ evaluation seeks to determine the status of the causal claim from the status of the correlation claim and that of the general mechanistic claim. In turn, the status of the general mechanistic claim is determined by the statuses of various specific mechanism hypotheses.

The status of a claim depends on the balance and weight of evidence. The balance of evidence denotes the direction in which the evidence points and the extent to which it points in that direction, while the weight of evidence concerns how conclusive the evidence is with respect to the claim in question. Weight of evidence is often used to measure the quality of the evidence, as per Table 3.1. This approach to quality of evidence was put forward by the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) Working Group (Atkins et al., 2004). The status of the claim can then be determined by the weight of the evidence, understood in terms of quality of evidence, and the balance of evidence understood as the confidence that the evidence inspires, as outlined in Table 3.2. This approach appeals to seven status levels, providing the opportunity for fine-grained assessments of claims.[4]

Table 3.1 Quality levels of evidence (Parkkinen et al., 2018).

Quality level	*Interpretation*
High	Further research is highly unlikely to have a significant impact on our confidence in the claim.
Moderate	Further research is moderately unlikely to have a significant impact on our confidence in the claim.
Low	Further research is moderately likely to have a significant impact on our confidence in the claim.
Very low	Further research is highly likely to have a significant impact on our confidence in the claim.

4 While seven may seem a lot of status levels, more levels allow for finer discrimination. There is some evidence that judgements are improved by offering a greater ability to discriminate (see, e.g., Felig et al., 2022).

§12.3. Assessing efficacy

In this section, we outline some key steps in an EBP+ evaluation of the efficacy of an intervention. The key question here is whether the intervention is a cause of the outcome under scrutiny in a particular study population. The particular ordering of the steps should be thought of as suggestive rather than prescriptive. The ordering provided is motivated by the fact that current EBP evaluations focus on assessing association studies and that an EBP+ evaluation can build upon current practice. But Evidential Pluralism does not insist upon any fixed ordering of the steps of causal enquiry, as noted in §10. See Parkkinen et al. (2018) for more detail on each of these key steps.

§12.3.1. Assessing association studies

An efficacy evaluation can begin by assessing the available association studies. We will not explore this aspect of the evaluation in any detail, as there are already a plethora of existing techniques for systematically searching for and assessing association studies. Systematic review procedures and meta-analysis can be applied here. As in the case of EBM (see §3), these approaches rank association studies in accord with hierarchies of evidence, viewing RCTs as producing higher-quality evidence than cohort studies, which are in turn ranked more highly than case control studies, case series and case reports. Some of these hierarchies of evidence, such as that advocated by GRADE, offer some flexibility as to the ranking, with better examples of studies at one level able to count more than poor studies at a higher level.

It is important to emphasise that approaches for assessing association studies usually tackle three different problems at once. The first is to assess the individual studies themselves: which studies have been carried out well and which badly, and how informative the individual study designs are, in order to get an idea of the quality of the studies. The second is to assess the correlation claim, usually with the aim of estimating the extent of any probabilistic dependence between the putative cause and effect, conditional on potential confounders. The third is to assess the causal claim: i.e., to ascertain whether intervention really is a cause of the outcome in question.

From the point of view of EBP+, the second and third tasks are incomplete. An assessment of association studies can only lead to a preliminary assessment of the correlation and causal claims, since no assessment of mechanistic studies and specific mechanism hypotheses has yet been undertaken (Figure 1.1). Nevertheless, this preliminary assessment can be very helpful and, in certain cases, it can be decisive, as we shall see next.

§12.3.2. Screening the need for a mechanistic evaluation

Depending on the results of the assessment of association studies, there may be no need for a full mechanistic evaluation. There are two scenarios in which this can be the case.

Table 3.2 Status of a claim (Parkkinen et al., 2018).

Status	*Interpretation*
Established	A claim is established when community standards are met for adding the claim to the body of evidence—i.e., for granting the claim and treating it as evidence for other claims. In order to establish a claim, evidence must warrant a high level of confidence in the claim and this evidence must itself be of high quality.
Provisionally established/ provisional	Moderate-quality evidence warrants a high level of confidence in the claim.
Arguably true/arguable	The claim is neither established nor provisionally established, but evidence of at least moderate quality warrants significantly more confidence in the claim than in its negation, or low-quality evidence warrants a high level of confidence in the claim.
Speculative	A claim is speculative if it falls into none of the other categories.
Arguably false	The claim is neither ruled out nor provisionally ruled out, but evidence of at least moderate quality warrants significantly more confidence in the negation of the claim than in claim itself, or low-quality evidence warrants a high level of confidence in the negation of the claim.
Provisionally ruled out	Moderate-quality evidence warrants a high level of confidence in the negation of the claim.
Ruled out	A claim is ruled out when community standards are met for adding the negation of the claim to the body of evidence. In order to rule out a claim, high-quality evidence must warrant a high level of confidence in the negation of the claim.

The first scenario is one in which the evaluation of the association studies reveals that these studies, when taken on their own, would suffice to rule out the existence of a correlation between the putative cause and effect. In the face of such strong evidence against correlation, it is practically impossible for a mechanistic evaluation to lead to much confidence in the existence of a correlation. Granted, evidence of mechanisms might undermine some of the association studies: it might, for example, identify potential confounders that have not been adequately controlled for (Jiménez-Buedo and Squitieri, 2019). The evidence base as a whole, however, will typically fail to establish the existence of a correlation, given the negative evidence from association studies. Hence a full mechanistic evaluation is not needed in order to settle the question of whether causation is established: causation is not established in such a scenario, because correlation is not established.

There is a second scenario in which there may be no need for a full mechanistic evaluation. Suppose the evaluation of the association studies reveals that these studies, when taken on their own, would suffice to establish causation (via confirmation channels α_1 and α_2 of Figure 1.1). If none of the specific

mechanism hypotheses under consideration could significantly undermine the correlation and mechanism claims, then this is another situation in which there is no need to assess mechanistic studies. In this scenario, causation is already shown to be established.

In other situations, however, a full mechanistic evaluation is likely to be more informative. It is then important to proceed to the next step—the search for mechanistic studies.

§12.3.3. Searching for mechanistic studies

The search for mechanistic studies needs to be guided by the specific mechanism hypotheses. In particular, mechanistic studies only need to be scrutinised where: (i) they bear on aspects of specific mechanism hypotheses that have not previously been established or ruled out, or (ii) they significantly undermine previously established claims about such features.

In the latter case, it is likely that domain experts will be aware of studies that subvert established claims. It is thus important that domain experts review the specific mechanism hypotheses to judge which features have already been established or ruled out, and to identify studies that may yet undermine such judgements. If crucial features of a specific mechanism hypothesis have been ruled out, and no studies have been identified that can undermine such judgements, then one may already be in a position to rule out the specific mechanism hypothesis in question.

For those mechanism hypotheses that have not been ruled out in this way, the next task is to formulate review questions to search the literature for studies that bear the features that have not already been established or ruled out.

Next, a systematic literature search needs to be carried out for studies that address the review questions and those that threaten to undermine judgements about features that have been established or ruled out.

The list of studies included in this search then needs to be refined in order to eliminate studies that are not informative in the presence of the remaining studies.

Finally, the review questions themselves may need to be refined and further searches carried out, in the light of the results of the previous searches.

§12.3.4. Assessing mechanistic studies

The quality of each of the resulting mechanistic studies needs to be determined next. There are three main criteria in play here: relevance, methods and implementation. In terms of relevance, the primary concern is how relevant the population of the mechanistic study is to the study population. There is also the question of how relevant the variables of the mechanistic study are to the features of the specific mechanism hypotheses under scrutiny. Next, one needs to ask whether the study uses well understood and reliable methods. Finally, the study needs to be assessed with respect to how well it implements these methods.

Once one has a grasp of the quality of each study, one needs to consider the credibility of the results of the studies. Results are particularly credible if they

are verified by independent methods, consistent from study to study, and robust across varying contexts.

§12.3.5. Assessing the specific mechanism hypotheses

One is then in a position to assess the status of each specific mechanism hypothesis. Recall that a hypothesis is established just when the evidence warrants a high level of confidence in the claim and the quality of evidence is such that further research is unlikely to have a significant impact on this confidence (Table 3.2). But hypotheses that are less than established can also be very informative with regard to causal evaluation: particularly those that are provisionally established and those that are ruled out or provisionally ruled out.

§12.3.6. Assessing the correlation claim

The assessment of association studies will have yielded a preliminary determination of the status of the correlation claim. The key question now is whether the assessment of the specific mechanism hypotheses provides grounds to modify this determination. Specific mechanism hypotheses may identify potential confounders that have not been controlled for by association studies, for example, and this can undermine the credibility of the correlation claim. On the other hand, if the key features of the relevant mechanisms are well established and it can be seen that the mechanism of action is not masked by counteracting mechanisms, this can increase the credibility of the correlation claim. The task here is thus to reach a more nuanced determination of the status of the correlation claim.

§12.3.7. Assessing the general mechanistic claim

The status of the general mechanistic claim is based on an assessment of channels α_2 and μ_2 of Figure 1.1.

With respect to channel α_2, it is important to ask whether a correlation of a similar magnitude has been detected robustly across contexts: this makes it more likely that the correlation is attributable to some underlying mechanistic connection, rather than confounding by features of the context, for example. Similarly, is the correlation large enough, and well enough controlled for, to be unlikely to be explained by bias and confounding? It is important to assess whether statistical explanations and non-causal connections can be ruled out as being responsible for the observed correlation.

The existence of a suitable mechanism can also be confirmed more directly via channel μ_2, i.e., an assessment of features of the mechanism complex. There are two main questions to address. Firstly, how well confirmed are the features of the mechanism complex? The more key features of the mechanism complex that have been established, the more confident one can be that there really is an appropriate mechanism underlying the observed correlation. Second, can the mechanism complex plausibly account for the magnitude of the observed correlation?

If the mechanism complex is very complex, it may be hard to reach a conclusive answer to this question. However, if plausible counteracting mechanisms have been shown not to cancel out the influence of the mechanism of action, then confidence can be increased.

When ascertaining the overall status of the general mechanistic claim, it is important to bear in mind that the α_2 and μ_2 channels can interact to reinforce or undermine one another. How much weight each channel should be given depends very much on how the quality of the association studies compares to that of the mechanistic studies.

§12.3.8. Assessing the causal claim

The status of the claim that intervention *A* is a cause of outcome *B* is the minimum of the status of the correlation claim and that of the mechanism claim. For example, if the correlation claim is provisionally established but the mechanism claim is arguably false then the causal claim is arguably false: this is because the correlation claim, even if true, is unlikely to be attributable to the intervention producing the outcome via some mechanism of action. Thus we have:

Causal status = minimum {correlation status, mechanism status}

Determining the status of the causal claim in the study population concludes the assessment of efficacy. The main steps of this assessment are outlined in Table 3.3.

Table 3.3 The main steps of the EBP+ approach to evaluating efficacy.

Step	*Key questions*
1. Carefully state the causal claim.	What are the cause, effect and study population? Is it the claim of interest? Is the claim single-case or repeatedly instantiable?
2. Carefully state the correlation claim.	What are the potential confounders?
3. Formulate specific mechanism hypotheses.	What are the key features of purported mechanisms of action, as well as counteracting and enhancing mechanisms?
4. Assess association studies.	How good is each individual study? What is the preliminary status of the correlation claim? What is the preliminary status of the causal claim?
5. Screen the need for a mechanistic evaluation.	Is the preliminary status of the correlation claim *ruled out*? Is it *established*? Could specific mechanism hypotheses significantly undermine this preliminary determination?
6. Search for mechanistic studies.	Which features of specific mechanism hypotheses have already been established or ruled out? Which review questions should be used to find studies relevant to remaining features?

(*Continued*)

Table 3.3 Continued

Step	*Key questions*
7. Assess mechanistic studies.	How relevant are the population and variables of each study? How reliable are its methods? Does it implement these methods well? Are the results independently verified, consistent and robust?
8. Assess specific mechanism hypotheses.	What status do the mechanistic studies confer on each specific mechanism hypothesis?
9. Assess the correlation claim.	Do specific mechanism hypotheses modify the preliminary status conferred on the correlation claim by association studies?
10. Assess the general mechanistic claim.	Have alternative explanations of the correlation, such as bias and confounding, been ruled out? How well confirmed are the features of the mechanism complex? Can it account for the magnitude of the observed correlation?
11. Assess the causal claim.	What is the minimum status of the correlation and mechanism claims?

§12.4. Assessing external validity

Having ascertained whether a causal claim holds in a source population, we turn next to the question of whether it extrapolates to a target population of interest. Recall the indirect route to establishing effectiveness:

Effectiveness = efficacy + external validity

In practice, even an efficacy assessment can require some judgements of external validity. This is because it will not normally be the case that all pertinent studies are carried out on exactly the same study population. 'The study population' may thus be a population of best fit to the various populations actually studied, and the relevance of the study results to this population of best fit will need to be assessed. Thus the distinction between efficacy and external validity is not always sharp. Nevertheless, it can help to disentangle the task of extrapolation from other aspects of an assessment of causation. Here we highlight some of the aspects of an assessment that are particular to external validity.

§12.4.1. Specifying the key claims

As in the case of an efficacy assessment, it is important to be clear as to what the putative cause and effect are, whether they are single-case or repeatedly instantiable. In the case of an external validity assessment, it is of course also crucial to clearly specify the study and target populations and to be clear about potential similarities and differences between these populations.

As explained in §4, judgements of external validity depend to a great extent on similarities and differences between the mechanisms linking the putative

cause and effect in the source population and those in the target population. The more similar the key components of the mechanisms of action in the two populations, the more confidence one can have about causation in the target population, other things being equal.

Thus, in the context of an assessment of external validity, one needs to consider all those specific mechanism hypotheses that are relevant to judgements of potential similarities and differences between key features of the mechanism complexes in the study and target populations. An assessment of these specific mechanism hypotheses can form the basis of a judgement of whether the the study and target mechanisms are sufficiently similar for causation on the study population to confirm causation on the target population.

§12.4.2. Screening the need for a mechanistic evaluation

Before searching for and assessing mechanistic studies, it is important to assess whether a full mechanistic evaluation of external validity will be worthwhile. There are two scenarios in which no such evaluation is recommended.

Firstly, there may be association studies and mechanistic studies on the target population itself that suffice to establish causation there, in which case there is normally no need to extrapolate from a different source population. The only exception would be if causation were ruled out on a source population that is hypothesised to be mechanistically very similar to the target population, as this would threaten to undermine the causal claim on the target population.

A second scenario in which a full mechanistic evaluation is unnecessary is the case in which a specific mechanism hypothesis has previously been established that implies that there is a crucial difference between the study and target populations. In such a case, the causal claim on the source population will not extrapolate straightforwardly to the target population and there is no reason to assess mechanistic studies that relate to other features of the mechanisms.

§12.4.3. Searching for mechanistic studies

As in the case of efficacy evaluation, mechanistic studies only need to be assessed where they bear on specific mechanism hypotheses that have not already been established or ruled out, or where they threaten to undermine judgements about what has been established or ruled out. Thus it is important that domain experts survey the specific mechanism hypotheses to ascertain what has already been established or ruled out and where there may be new studies that threaten to undermine such judgements.

Once domain experts have scrutinised the specific mechanism hypotheses, review questions can be formulated and a systematic literature search can be carried out. Uninformative studies can be eliminated and the review questions and the search can be refined by means of an iterative process.

§12.4.4. Assessing mechanistic studies

The next step is to assess the quality of each of the identified mechanistic studies and the credibility of their results. Evaluative criteria here are the same as in the case of efficacy: relevance of the studies; the quality of their methods and implementation; whether the results are independently verified, consistent and robust.

Needless to say, if a mechanistic study has already been assessed in the context of an assessment of efficacy in the study population or the target population, then no new assessment is required.

§12.4.5. Assessing the specific mechanism hypotheses

A status then needs to be assigned to each specific mechanism hypothesis that has not already been evaluated in the context of efficacy evaluations. Again, the procedure here mirrors that of efficacy evaluation.

§12.4.6. Assessing mechanistic similarity

There are two conditions for successful extrapolation of a causal claim from a source population to a target population. Firstly, the mechanism of action of the intervention in the source population needs to be present in the target population: the key features of the purported mechanism of action in the target population need to be similar to those identified in the source population. Second, there should be no counteracting mechanisms in the target population that are not also present in the source population, or, if there are such mechanisms, association studies in the target population should demonstrate that they do not nullify the effect of the mechanism of action.

Hence, when assessing mechanistic similarity, it is important to assess both the similarities between the mechanisms of action in study and target populations and to check the similarity of ancillary mechanisms, i.e., mechanisms that counteract or reinforce the mechanism of action.

There are some important differences between an assessment of mechanistic similarity for external validity and an assessment of the claim that there exists an appropriate mechanism to underwrite efficacy. Assessing mechanistic similarity is more demanding because it requires identifying the details of the mechanism complex, whereas the mechanistic existence claim can, in certain circumstances, be established just by means of association studies (channel α_2 in Figure 1.1). For example, while several concordant and high-quality RCTs that establish a large association might suffice to establish the mechanistic existence claim, they would be uninformative with respect to external validity. In addition, mechanistic similarity is a matter of degree, while the mechanistic existence claim is a categorical question—either a suitable mechanism exists or it doesn't. Furthermore, mechanisms can be similar with respect to certain features but not others—which features are similar matters a lot for judgements of external validity. In contrast, the mechanistic existence claim is not relativised to features.

The main steps of an external validity assessment are outlined in Table 3.4. From the point of view of EBP+, similarity of mechanisms in the study and target populations underpins external validity, so the goal is to assess this mechanistic similarity.

§12.5. Assessing effectiveness

Where the mechanisms in the study and target populations are found to be sufficiently similar, a causal claim that is established in the source population can help to confirm causation in the target population. This is helpful where the studies performed on the target population are less conclusive than those on the source population. The indirect route to establishing effectiveness thus requires an assessment of efficacy on the target population (i.e., by assessing studies on the target population rather than on the source population), but with extra channels of confirmation from the previously established causal claim on the source population, represented in Figure 1.4.

When assessing the correlation and mechanism claims in the target population, one needs to consider the extent to which the status of these claims is boosted by the fact that the causal claim holds in a source population that is mechanistically

Table 3.4 The main steps of the EBP+ approach to evaluating external validity.

Step	*Key questions*
1. Carefully state the causal claims.	What are the cause, effect, source population and target population? Is the target population causal claim the claim of interest?
2. Formulate specific mechanism hypotheses.	What are the key features responsible for the proper functioning of the mechanism of action? What are the key features of counteracting and enhancing mechanisms?
3. Screen the need for a mechanistic evaluation.	Is the status of the target causal claim determined by studies on the target population? Are there crucial differences between mechanisms on the study and target populations?
4. Search for mechanistic studies.	Which similarities and differences have already been established or ruled out? Which review questions should be used to find studies relevant to remaining features?
5. Assess mechanistic studies.	How relevant are the population and variables of each study? How reliable are its methods? Does it implement these methods well? Are the results independently verified, consistent and robust?
6. Assess specific mechanism hypotheses.	What status do the mechanistic studies confer on each specific mechanism hypothesis?
7. Assess mechanistic similarity.	How similar are the study and target mechanisms of action? Are any new counteracting mechanisms in the target population likely to nullify the effect of the mechanism of action?

similar to the target population. As noted here, the extent to which this status is boosted depends on the salience of the features with respect to which there is similarity, as well as the degree of similarity of each feature.

One can combine the direct and indirect routes to evaluating effectiveness by means of the steps outlined in Table 3.5.

§13 EBP+ in Comparison to Existing Approaches

In this section, we compare EBP+, as motivated by Evidential Pluralism, to other approaches that are relevant to evaluation.

Campbell reviews. As noted in §11, the methods for systematic review of the Campbell Collaboration are modelled on those of the Cochrane Collaboration and focus on RCTs (see Campbell Collaboration, 2020, §2.4). These methods do not systematically scrutinise mechanistic studies to help determine whether an intervention works. However, they do admit the use of *logic models*, *conceptual frameworks* or *theories of change*: these are models of the mechanism by which an intervention produces outcomes of interest (Higgins et al., 2019, §2.5.1). From the perspective of Evidential Pluralism, these act as specific mechanism hypotheses and they ought to play an important evidential role in evidence-based policy evaluation; see also Cartwright (2020) on this point. However, they do not play an evidential role in a Campbell review—instead, they are used as a communication tool:

> The way that logic models can be represented diagrammatically . . . provides a valuable visual summary for readers and can be a communication tool for decision makers and practitioners. They can aid initially in the development of a shared understanding between different stakeholders of the scope of the review and its PICO, helping to support decisions taken throughout

Table 3.5 The main steps of the EBP+ approach to evaluating effectiveness.

Step	*Key questions*
1. Evaluate efficacy on the target population.	What preliminary status do target population studies confer on the causal claim in the target population?
2. Screen the need for an external validity evaluation.	Is the target causal claim already established or ruled out? Are there source populations in which it might be possible to establish causation?
3. Evaluate efficacy on these source populations.	For which source populations do the available studies establish causation?
4. Evaluate external validity.	For each source population within which causation is established, how mechanistically similar is it to the target population?
5. Evaluate effectiveness.	Do these source population causal claims warrant revising the status of causation in the target population?

> the review process, from developing the research question and setting the review parameters, to structuring and interpreting the results. They can be used in planning the PICO elements of a review as well as for determining how the synthesis will be structured (i.e. planned comparisons, including intervention and comparator groups, and any grouping of outcome and population subgroups). These models may help review authors specify the link between the intervention, proximal and distal outcomes, and mediating factors. In other words, they depict the intervention theory underpinning the synthesis plan.
>
> (Thomas et al., 2019, p. 26)

Importantly, a Cochrane review does not examine the evidence for these mechanism hypotheses, nor consider the extent to which they are confirmed by available evidence. Evidential Pluralism, in contrast, advocates considering these mechanism hypotheses together with hypothesised counteracting and reinforcing mechanisms, and ascertaining the status of these hypotheses in the light of concrete evidence. These hypotheses then play an important evidential role in assessments of causation.

While Campbell and Cochrane reviews do not scrutinise mechanistic studies, they can consider qualitative evidence, such as narratives or text from survey responses. Qualitative research 'can help paint a richer picture of the intervention, its effects, how or why it produced those effects (or not), and other such features that provide texture and explanatory context to a review' (Campbell Collaboration, 2020, p. 10). They can also shed light on the attitudes of stakeholders (Noyes et al., 2019, §21.1). They are not used to determine whether the intervention works. The connection between mechanistic and qualitative studies will be discussed further in the next chapter.

Realist Evaluation and EMMIE. Some practitioners have begun to question the current EBP focus on RCTs. For example, Yamey and Feachem (2011) observe that:

> while the RCT is rightly hailed as the 'pinnacle' of evidence-based medicine, in the global public health community, there is growing recognition that new research designs are desperately needed to help evaluate 'real world' programmes. Such designs would, we believe, also help to illuminate the implementation 'black box'.
>
> (Yamey and Feachem, 2011, p. 98)

Moreover, one of the What Works centres has begun to recognise the importance of mechanisms. The What Works Centre for Crime Reduction has developed the 'EMMIE' framework for systematic reviews of evidence: Effect size, Mechanism, Moderator, Implementation and Economics are all components of an evaluation (Johnson et al., 2015; Tilley, 2016; Thornton et al., 2019). From the point of view of Evidential Pluralism, considering mechanisms is an important step in the right direction. EMMIE is based not on Evidential Pluralism but on the realist evaluation approach of Pawson and Tilley (1997), and it will be

instructive to consider how their approach differs from one based on Evidential Pluralism.

While mechanisms are important to both realist evaluation and Evidential Pluralism, there are three key philosophical differences between these two approaches.

Firstly, the realist evaluation approach of Pawson and Tilley (1997) makes a firm commitment to scientific realism: specifically, a metaphysics of causation that posits causal powers (Pawson and Tilley, 1997, pp. 33, 56). This realism is accompanied by a rejection of Humean and Kantian metaphysics, which hold that causal relationships are a device we employ to structure the world, and which do not posit causal powers or causal necessitation 'out there' in the world. Evidential Pluralism, in contrast, is a purely epistemological thesis that makes no explicit metaphysical claims. As we saw in §5, it is even compatible with an anti-realist account of causation which analyses causal claims in terms of rational beliefs.

The second philosophical difference between the realist evaluation of Pawson and Tilley (1997) and Evidential Pluralism is that, on account of its metaphysical commitment, their approach involves a rejection of the experimental methodology that underpins RCTs and certain other kinds of association study. Evidential Pluralism, in contrast, takes experimental methods to have the potential to provide good evidence, relevant to the assessment of a causal claim. If Evidential Pluralism is right, one should not reject these methods—rather, one should augment them, by considering mechanistic studies alongside association studies. The aim of EBP+ is to develop and improve, rather than overturn, present-day EBP.

Third, realist evaluation proceeds from the premise that there is no logic of evaluation (Pawson and Tilley, 1997, p. xiii), while Evidential Pluralism takes there to be a logic of evaluation, portrayed by Figure 1.1. According to this logic of evaluation, causation is established by establishing correlation and mechanism, which in turn requires assessment of the confirmation channels α_1, α_2, μ_1, μ_2, μ_3, i.e., the assessment of any relevant association studies and mechanistic studies. As we saw in §12, this logic of evaluation can be broken down into a series of practical steps.

We should note that realist evaluation has been developed in a number of different directions since 1997—see Jagosh et al. (2016), for example, for some pointers. In particular, not all proponents of realist evaluation now reject the experimental methodology and RCTs. For example, Bonell et al. (2012) argue for the use of RCTs in a way that is sensitive to the concerns of realist evaluation. Moreover, the development of EMMIE can be considered to be a move towards a logic of evaluation.

Thornton et al. (2019), although proponents of EMMIE, identify some limitations of realist evaluation as implemented in the EMMIE approach. In practice, EMMIE exclusively scrutinises systematic reviews, which almost always consider association studies rather than mechanistic studies, so evidence of mechanisms tends to appear rather scant and hence to be rated as weak. From the point of view of Evidential Pluralism, it is not enough to consider systematic reviews of association studies—it is essential to articulate specific mechanism

hypotheses and to search the literature for evidence relevant to those hypotheses. As noted in §3, the International Agency for Research on Cancer provides an example of good evaluation practice here. This is because each carcinogenicity evaluation has a dedicated subgroup responsible for systematically assessing mechanistic studies.

Furthermore, since an EMMIE evaluation has five components, it is not obvious how these five aspects should combine to give an overall assessment. This opens the door to subjective judgements of relative importance to influence the overall assessment. Thus, Thornton et al. (2019) worry that realist reviews may not be replicable. This is less of a concern for Evidential Pluralism, which only has two strands to integrate, namely evidence of correlation and evidence of mechanisms (Williamson, 2021c). As we saw in the last section, one can integrate these two strands in a systematic way to come to an overall assessment.

One final point worth noting with respect to EMMIE is that 'Mechanism' is graded on a scale from 0 to 4, but only grade 4 requires concrete evidence of mechanism: grades 1 to 3 merely require some story or theory about what the mechanism might be (Thornton et al., 2019, Figure 1.2). This may stem from the important role of theory in realist evaluation (Pawson and Tilley, 1997, p. 59). In contrast, Evidential Pluralism is concerned with evidence, not theory. In the social sciences, it is often very easy to conjecture a mechanism. A story of a mechanism that is not backed up by evidence has no confirmatory value for Evidential Pluralism.

Although Evidential Pluralism differs from realist evaluation in important ways, the two approaches do share some key claims: most notably, that EBP needs to move beyond EBM's monistic focus on association studies, and that mechanisms should play a prominent role. Evidential Pluralism can be thought of as providing some motivation for these claims that is not tied to realism.

Theory-based evaluation. Realist evaluation is sometimes classed as a kind of *theory -based evaluation* approach. Theory-based evaluation tests a theory of change or logic model against available evidence. As noted earlier, a theory of change or logic model can be thought of as a specific mechanism hypothesis that articulates the structure of the mechanism of action of an intervention. From the perspective of Evidential Pluralism, then, theory-based evaluation focuses on the μ-channels of Figure 1.1.

Contribution analysis offers another example of theory-based evaluation (Mayne, 2001, 2012, 2019). It is sometimes combined with process tracing to yield *contribution tracing* (Befani and Mayne, 2014). These approaches aim to assess the contribution made by an intervention to an outcome of interest. It can be difficult to quantify this contribution, so theory-based evaluation tends to focus on the qualitative question of whether a contribution was made (HM Treasury, 2011, §3.4). Theory-based evaluation is thus sometimes viewed as appropriate where it is not possible to collect association studies that can help quantify the contribution: for example, where there is no comparable group that has not received the intervention.

From the perspective of Evidential Pluralism, there is no good reason to restrict the use of theory-based evaluation to situations in which it is not possible to assess

association studies. Evidence of correlation and evidence of mechanisms act in a complementary, mutually reinforcing way and can be combined. Evidential Pluralism provides a framework for combining the two, within which theory-based evaluation procedures can be used as a means to articulate and test specific mechanism hypotheses.

Causal modelling. We have seen that theories of change and logic models can be thought of as models of mechanisms, and thus can play a role as specific mechanism hypotheses in causal enquiry (Figure 1.1). These are not the only ways to model mechanisms, however. Structural equation models (Westland, 2015; Russo et al., 2019), causal mediation analysis (Imai et al., 2010), potential outcome models (Rubin, 1974) and graphical causal models (Pearl, 2000), for example, are also used in the social sciences.

These alternative means of modelling mechanisms can also be integrated into the framework provided by Evidential Pluralism (see, e.g., Cartwright, 2021), with two provisos. Firstly, these approaches are best thought of as appropriate for modelling only certain aspects of mechanisms: they are especially good at modelling the influence of mediating variables. However, no single kind of model is able to do justice to every kind of mechanism. Most notably, the behaviours of some mechanisms hinge on spatiotemporal organisation rather than mediating variables. A chimney mechanism offers an extreme example here, because its behaviour is largely due to its spatial structure rather than intermediary variables, and even key activities, such as the flow of smoke, are extrinsic to the mechanism. Cell wall mechanisms in biology also depend heavily on spatial organisation, as do many social mechanisms, such as mechanisms for crowd management. Such mechanisms can be hard to model using these techniques. Thus, while standard causal modelling methods can be integrated into EBP+ evaluation, none provides a panacea.

The second proviso is that while the methods themselves can be plugged into the general framework of Evidential Pluralism, some of their philosophical presuppositions may conflict with Evidential Pluralism. In particular, some causal modelling approaches identify causation with 'average treatment effect', which is a kind of probabilistic dependence. Evidential Pluralism, in contrast, is based on the premise that establishing causation should not be conflated with establishing correlation, however sophisticated the measure of correlation, and that consideration needs to be given to the question of whether there is some mechanism responsible for an observed correlation. Thus, while a modelling method may be a fruitful tool for modelling mechanisms, and quite compatible with Evidential Pluralism and EBP+ evaluation, one must be wary of certain assumptions that are sometimes propounded alongside modelling methods and which can underpin fallacious causal inferences.

Impact evaluation. An impact evaluation is used to determine the outcomes of an intervention that is implemented in practice, and to understand how the intervention contributed to those outcomes (Gertler et al., 2011). It is often undertaken

with a view to identifying ways in which an implementation of the intervention can be improved upon. A *process evaluation*, in particular, aims to shed light on the mechanism of an intervention, to help explain why the intervention produced, or failed to produce, certain outcomes. This sort of evaluation can seek the active ingredients of the intervention (Michie et al., 2009) or the common elements responsible for the effectiveness of multiple interventions (Boustani et al., 2015). Such an evaluation is normally viewed as complementary to an evaluation of efficacy or effectiveness: it seeks to answer the question of how the intervention works, rather than whether it works. An *effectiveness-implementation hybrid design* carries out an evaluation of implementation alongside an evaluation of effectiveness, but again these are viewed as different, complementary tasks (Curran et al., 2012; Landes et al., 2019).

Evidential Pluralism is distinctive because it exploits evidence of mechanisms in order to address the question of whether the intervention works—not just how it works. If Evidential Pluralism is correct, making use of evidence of mechanisms can lead to better-informed judgements of effectiveness, especially where association studies on their own are inconclusive. Thus the EBP+ approach to evaluation can be thought of as broadening the remit of techniques such as process evaluation—enabling these tools to help answer the question of whether an intervention works.

In sum, Evidential Pluralism leads to a distinctive approach to evaluating the efficacy, external validity and effectiveness of social interventions. We call this 'EBP+' because it adds the ability to scrutinise evidence of mechanisms to present-day EBP evaluation methods. But it does not only build upon present-day EBP: EBP+ provides an overarching framework within which many approaches to evaluation can play a role. In particular, existing techniques for theory-based evaluation, causal modelling and impact evaluation can all help with the assessment of evidence of mechanisms in an EBP+ evaluation. EBP+ can thus be thought of as a general evaluation framework that can help evaluators to integrate and exploit a wide range of evaluation tools to help assess what works.

Next, we turn to a second methodological consequence of Evidential Pluralism: the use of mixed methods for social science research.

4 Mixed Methods Research

Mixed methods research has become widespread in the social sciences in recent decades, with a particular preponderance of mixed methods studies in certain fields, such as educational research, family studies and anthropology. This popularity is witnessed by a sharp increase in the number of publications mentioning mixed methods in the title or abstract over the past twenty years (Creswell, 2012; Timans et al., 2019), as well as by the production of textbooks and handbooks, and by the founding of journals, e.g., *Journal of Mixed Methods Research* and *International Journal of Multiple Research Approaches*.

Despite this growth, basic questions relating to mixed methods research remain open. For example, although mixed methods research is typically construed as a methodology or a methodological orientation employing both qualitative and quantitative elements (e.g., methods, data and designs), there is still no consensus about its definition (e.g., Greene et al., 1989; Tashakkori and Teddlie, 1998; Johnson et al., 2007; Creswell and Plano Clark, 2018). In addition, there is no consensus about how to mix or combine qualitative and quantitative elements in research. There are at least seven 'levels' of research that can be mixed or combined: data, methods, design, epistemology, ontology, purposes of research and practical roles of research (Biesta, 2010). There are also multiple ways of mixing or combining these levels (e.g., Creswell et al., 2003; Teddlie and Tashakkori, 2009; Creswell and Plano Clark, 2018; van Grootel et al., 2020).

What is more, there is no consensus regarding the philosophical foundations of mixed methods research. The problem of how to motivate and justify the use of mixed methods from a philosophical point of view remains unresolved. There are a variety of positions that are employed to provide philosophical underpinnings for mixed methods research. One popular position is pragmatism (e.g., Johnson and Onwuegbuzie, 2004; Morgan and Winship, 2007; Teddlie and Tashakkori, 2009; Feilzer, 2010; Johnson et al., 2017; Creswell and Plano Clark, 2018). Other positions include the dialectical position (Greene et al., 1989; Greene, 2006; Greene and Hall, 2010), dialectical pluralism (Johnson, 2017), the transformativist position (Mertens, 2003, 2007, 2010), critical realism (Maxwell and Mittapalli, 2010),

DOI: 10.4324/9781003143000-6

the indigenous position (Chilisa, 2012), feminism (Hesse-Biber, 2010, 2015) and the performative position (Schoonenboom, 2019).[1]

In this chapter, we argue that Evidential Pluralism can provide new philosophical foundations for mixed methods research in causal enquiry: it motivates the need to pay attention to mixed methods and it can offer guidance on how to integrate quantitative and qualitative methods in practice. Thus Evidential Pluralism can help to deepen our understanding of mixed methods research. The chapter is structured as follows. §14 reviews the context of the origins of mixed methods research. §15 provides an overview of the main accounts of the philosophical foundations of mixed methods research. §16 critically examines these positions. §17 argues that Evidential Pluralism helps to justify the use of mixed methods for causal enquiry in the social sciences and that it can provide guidance on how to integrate mixed methods in practice.

§14 The Context of the Origins of Mixed Methods Research

In the twentieth century, two social science methodologies stood out: the quantitative research approach and the qualitative research approach.[2] The quantitative research approach relies on the collection of quantitative data, obtained by methods such as experiments, quasi-experiments, surveys and longitudinal studies. The qualitative research approach relies on the collection of qualitative data, obtained by methods such as narrative research, phenomenological research, ethnography and case studies. These two approaches have operated in parallel, and often in opposition to one another (Kelle, 2015). For example, sociology fostered a division between social theorists and quantitative researchers. Quantitative researchers tend to focus on statistical analyses and often dismiss the need to 'develop sociological models mirroring conceptions of mechanisms of social processes' (Sørensen, 1998, p. 239). In contrast, social theorists are 'often so concerned with their concepts and theoretical frameworks that they pay little attention to the significance of quantitative findings' (Mahoney, 2001, p. 582). Political science has also hosted a methodological divide between the quantitative and qualitative approaches. As John (2010, p. 9) notes, 'many researchers still tend to use one approach, but not the other'.

This methodological schism reflects underlying philosophical disagreement. Indeed, many social scientists view methodology and philosophy as intrinsically related (Pole and Lampard, 2002, pp. 6–8; Maxwell and Mittapalli, 2010,

1 It should be noted that feminism is not normally used to provide philosophical foundations for mixed methods research. Rather, it concerns 'why and how feminists use mixed methods research' (Hesse-Biber, 2010, p. 132).

2 The quantitative research approach and the qualitative research approach are also called 'the quantitative research paradigm' and 'the qualitative research paradigm' (Johnson and Onwuegbuzie, 2004; Feilzer, 2010; Agerfalk, 2013). Given ambiguities in the use of 'paradigm' in the social sciences (Biesta, 2010), we adopt the less controversial term 'approach' in this chapter.

p. 147; Tebes, 2012, p. 14; Johnson and Gray, 2010, p. 88; Creswell and Plano Clark, 2018, pp. 4–5). As Feilzer (2010, p. 7) puts it, 'the choice of social sciences research questions and methods . . . is a reflection of researchers' epistemological understanding of the world, even if it is not articulated or made explicit'.

The quantitative research approach is often said to be rooted in positivism (Comte, 1830; Quetelet, 1835; Pearson, 1900) and is now typically associated with so-called postpositivism (Garrison, 1986; Phillips, 1990; Phillips and Burbules, 2000), while the qualitative research approach has been generally coupled with constructivism or interpretivism (Dilthey, 1883; Weber, 1904; Guba and Lincoln, 1989).[3] Postpositivism and constructivism/interpretivism mainly differ with respect to three basic issues: ontology, epistemology and axiology.[4] Ontologically, postpositivism assumes that there is a single, mind-independent reality, though it can be only understood imperfectly (Campbell, 1974, pp. 48–49; Cook and Campbell, 1979, p. 29), while constructivism/interpretivism assumes that there are multiple, socially constructed and holistic realities (Guba and Lincoln, 1989; Guba, 1990; Guba and Lincoln, 2005). Epistemologically, postpositivism assumes that social scientific research is to a great extent objective in the sense that the researcher is usually treated as independent of the object of the research (Smith, 1983), whereas constructivism/interpretivism assumes that social scientific research is subjective in the sense that the relationship of the researcher to the object of the research is interactive (Lincoln and Guba, 1985; Guba and Lincoln, 1989, 2005).[5] Axiologically, postpositivism assumes that social scientific research is value-laden, but the influence of values can be well controlled (Cook and Campbell, 1979; Shadish et al., 2002), while constructivism/interpretivism assumes that social scientific research is essentially value-relative (Guba and Lincoln, 2005; Teddlie and Tashakkori, 2009). Accordingly, postpositivism assumes that the aim of social scientific research is to describe, explain, predict and intervene upon social phenomena (Fay, 1975; Black, 1999; Phillips and Burbules, 2000), while constructivism/interpretivism assumes that the main purpose of social scientific research is to interpret and understand social phenomena (Guba and Lincoln, 2005).

The methodological debate between the quantitative and qualitative research approaches has been intertwined with a persistent philosophical confrontation between postpositivism and constructivism/interpretivism. This is known as the

3 It should be highlighted that social scientists' characterisations of philosophical positions often differ from those of philosophers. It is also worth noting that social scientists sometimes refer to philosophical positions as 'paradigms' (Teddlie and Tashakkori, 2009), 'stances' (Greene et al., 1989; Maxwell and Mittapalli, 2010) or 'worldviews' (Creswell and Plano Clark, 2018). In order to avoid confusion, we will talk of philosophical 'positions' instead in this chapter.

4 In this context, social scientists often use the term 'epistemology' in a distinctive way, referring specifically to the relationship between the knower and the known (Lincoln and Guba, 1985, pp. 37–38).

5 According to postpositivism, the researcher and the object of the research should be independent of one another in the social sciences, as is usually thought to be the case in the natural sciences. In other words, the researcher should exteriorise the phenomenon studied, keeping detached and distant from it. However, this is not possible, according to constructivism/interpretivism, because the researcher and the object of the research are humans. It is impossible to entirely separate them and eliminate mutual interaction. The results of the research are created by this interaction between the researcher and the object of the research (Guba and Lincoln, 1989, p. 88).

'paradigm wars'. Advocates of the quantitative research approach contend that research should be centred around quantitative methods in order to develop a nomothetic body of knowledge about mind-independent reality, while supporters of the qualitative approach maintain that research should be undertaken mainly by means of qualitative methods in order to develop an ideographic body of knowledge about socially constructed realities. The contenders in this debate maintain that their disagreement over methodology is a result of their disagreement over philosophy (i.e., ontology, epistemology, and axiology). As Teddlie and Tashakkori (2009, p. 20) indicate, the paradigm wars are basically 'the conflict between the competing scientific worldviews of positivism (and variants, such as post-positivism) and constructivism (and variants, such as interpretivism) on philosophical and methodological issues'.

§15 Mixed Methods Research and its Philosophical Foundations

In the late 1980s, the heyday of the paradigm wars, mixed methods research developed as a methodological alternative to the quantitative and the qualitative approaches in order to 'overcome the speechlessness between both traditions' (Kelle, 2015, p. 603). However, an immediate difficulty arose. As we have seen, the quantitative research approach is often associated with postpositivism, while the qualitative research approach is usually coupled with the constructive/interpretivist position. The problem is that an integration of quantitative and qualitative methods appears to be thwarted by the incompatibility between their underlying philosophical positions. Thus, a key task for advocates of mixed methods research has been to develop coherent philosophical foundations that justify the use of both quantitative and qualitative methods/data/designs.

Pragmatism is often invoked to provide philosophical foundations for mixed methods research. Pragmatism in this context is rooted in American pragmatism, especially the works of John Dewey, Charles Sanders Peirce and Richard Rorty (Cherryholmes, 1992; Johnson and Onwuegbuzie, 2004; Feilzer, 2010; Tebes, 2012). This position takes knowledge, as a product of person-environment interaction, to be both constructed and based on a mind-independent reality, and it highlights the instrumental feature of theories in enquiry (Johnson and Onwuegbuzie, 2004; Johnson and Gray, 2010; Morgan, 2014). According to pragmatism, both the mind-independent physical world and the constructed social and psychological world exist, and this reality is complex (Johnson and Onwuegbuzie, 2004, p. 18; Johnson and Gray, 2010, p. 88, Creswell and Plano Clark, 2018, pp. 10–11); social scientific research is value-oriented (Johnson and Onwuegbuzie, 2004, pp. 16–18; Johnson and Gray, 2010, p. 88; Creswell and Plano Clark, 2018, pp. 10–11); and the aim of social scientific research is to solve problems (Johnson and Onwuegbuzie, 2004, p. 18; Teddlie and Tashakkori, 2010, pp. 17–18). Social scientists do not have to make an either-or choice between postpositivism and constructivism/interpretivism (Johnson and Onwuegbuzie, 2004; Teddlie and Tashakkori, 2009; Creswell and Plano Clark, 2018)—they are free to choose the methods, data and procedures that best meet their needs, and can employ both quantitative and qualitative elements (e.g., methods and data) when designing and

conducting research (Johnson and Onwuegbuzie, 2004; Greene, 2006; Teddlie and Tashakkori, 2009; Creswell and Plano Clark, 2018).[6]

The dialectical position provides an alternative to pragmatism (Greene et al., 1989; Greene, 2007; Greene and Hall, 2010). Like pragmatism, the dialectical position does not assume that social scientists have to make a choice between postpositivism and the constructive/interpretivist position. Unlike pragmatism, the dialectical position recognises and accepts the legitimacy of all other philosophical positions (e.g., postpositivism and constructivism/interpretivism). The dialectical position maintains that different philosophical positions play an important role in leading to different lines of enquiry. Mixed methods research is justified on the grounds that diverse methods engender a better understanding of the phenomena being studied. The dialectical position 'actively welcomes more than one philosophical position, along with more than one methodology and type of method, into the same inquiry space and engages them in respectful dialogue one with the other throughout the inquiry' (Greene and Hall, 2010, p. 124).

More recently, based on an appeal to pragmatism and the dialectical position, Johnson (2017) has developed 'dialectical pluralism'. According to dialectical pluralism, there are multiple realities and multiple ways of conceptualising reality; knowledge in the social sciences is fallible and contextual; and social scientific research is value-laden. Therefore, social scientists 'should dialectically listen and consider multiple methodological concepts, issues, inquiry logics, and particular research methods and construct the appropriate mix for each research study' (Johnson, 2017, p. 167). Johnson argues that dialectical pluralism complements and extends the dialectical position by articulating its philosophical assumptions.

Another influential position is the transformative position, developed largely by Mertens (2003, 2007, 2010). The transformative position assumes that there are multiple realities that are socially constructed and defined by social, political, cultural, economic, ethnic, racial, gender, age and disability values; knowledge is socially and historically located within a complex cultural context; and an important aim of social scientific research is to 'serve the ends of creating a more just and democratic society' (Mertens, 2003, p. 159). Mertens argues that a careful mixture of quantitative and qualitative methods enables one to represent a variety of perspectives, including those that have been traditionally overlooked. According to the transformative position, social scientists should prefer mixed methods 'for working toward increased social justice' (Mertens, 2007, p. 224).

6 It is worth noting that there are different versions of pragmatism, ranging from 'dialectical pragmatism' (Teddlie and Tashakkori, 2009; Johnson and Gray, 2010; Johnson and Christensen, 2014), which offers a wholesale justification of mixed methods research, to Gert Biesta's 'Deweyan pragmatism', which only '[helps] us to have a more precise discussion about the strengths and weaknesses of mixed methods approaches' (Biesta, 2010, p. 97). Pragmatism is also sometimes associated with perspectivism (Tebes, 2012) and pluralism (Johnson and Onwuegbuzie, 2004; Johnson, 2017).

In contrast, Maxwell and Mittapalli (2010) motivate mixed methods research by appeal to critical realism.[7] This position assumes that there is a mind-independent physical world and 'there can be more than one scientifically correct way of understanding reality in terms of conceptual schemes with different objects and categories of objects' (Lakoff, 1987, p. 65). Maxwell and Mittapalli argue that these critical realist assumptions imply a mechanistic account of causality, a realist account of mental phenomena, a realist concept of validity and a realist account of diversity. Moreover, they argue that these 'realist assumptions' justify a process-based qualitative approach to causality, a critical approach to qualitative research and a mixed methods approach to validity, and overcome the neglect of methodological diversity in practice. In these ways, Maxwell and Mittapalli claim that mixed methods research can be justified from a critical realist perspective.

Recently, Schoonenboom (2019) has developed a performative position. Its ontological and epistemological assumptions are borrowed from dialectical pluralism: the performative position assumes the existence of multiple realities that can be known and investigated in various ways. Instead of claiming that people adopt different perspectives on a single objective reality, the performative position supposes that people to some extent live in different realities, with different concepts, habits, interests and values. These different realities are dynamic and changeable. Moreover, the concepts that researchers use to speak about these realities are not fixed, but change as these realities change. Realities are thus multiple, varied and changing. The axiological assumptions of the performative position are similar to those of pragmatism. However, the performative position differs from pragmatism in ontology and epistemology, while it is distinct from dialectical pluralism in that 'mixed methods researchers do not switch between ontologies, epistemologies, or research stances. Rather, they stay and live in multiple realities that can be known in various ways' (Schoonenboom, 2019, p. 289).

These positions can be classified into two types: monistic positions and pluralistic positions. A monistic position tries to justify the use of mixed methods in social scientific research by appeal to a particular set of philosophical assumptions.[8] Pragmatism and the transformative position are good examples of monistic positions. They each endorse a distinctive set of philosophical assumptions, which are then used to motivate the use of mixed methods in social scientific research.

A pluralistic position, on the other hand, justifies the use of mixed methods by means of different sets of philosophical assumptions.[9] As Maxwell (2011, p. 29)

7 It is worth noting that Maxwell and Mittapalli's critical realism is conceptually different from Bhaskar's late 'transcendental dialectical critical realism' which is a moral and spiritual position as well as a scientific position (Bhaskar, 2011).

8 This is similar to what Teddlie and Tashakkori (2009, p. 98) call 'the single paradigm thesis', or what Creswell and Plano Clark (2018, p. 26) call 'one "best" worldview' for mixed methods.

9 This is similar to what Teddlie and Tashakkori (2009, p. 99) call 'the multiple paradigms thesis', or what Creswell and Plano Clark (2018, p. 27) call 'multiple worldviews' for mixed methods.

says, 'I do not think it is generally appropriate or useful to attempt to synthesize different philosophical approaches or assumptions into a single, logically consistent paradigm for mixed methods research. Different situations and research problems may require different sets of assumptions and models, as well as different combinations of methods'. The dialectical position, dialectical pluralism and the performative position are pluralistic positions. All maintain that different philosophical assumptions guide the use of different methods. Note that the performative position differs from the dialectical position and dialectical pluralism in that it holds that different philosophical positions appeal to 'sharable' concepts (Schoonenboom, 2019, p. 289).

Critical realism involves aspects of both a monistic and a pluralistic position. On the one hand, Maxwell and Mittapalli (2010, p. 147) are explicitly sceptical of 'the entire concept of unified paradigms in research, a concept that has dominated the discussion of the relationship between philosophical assumptions and research methods'. This speaks against a monistic position. On the other hand, Maxwell and Mittapalli argue that their critical realist position merely motivates the use of mixed methods in some cases—it does not justify the universal use of mixed methods in the social sciences. Therefore, it is perhaps appropriate to regard Maxwell and Mittapalli's critical realist position as a single set of realist assumptions within a pluralistic position.

§16 A Critical Analysis

In order to analyse these positions, we need to revisit the concept of 'philosophical foundations': what are philosophical foundations and what are they expected to provide?

The debate about the philosophical foundations of mixed methods research is framed by the paradigm wars. Most philosophical foundations for mixed methods research seek a set of assumptions about ontology, epistemology and axiology that parallel those provided for quantitative research and qualitative research. There are, however, some differences in emphasis between ontology, epistemology and axiology, as we shall now see.

Both postpositivism and constructivism/interpretivism provide what we shall call *ontology-oriented* philosophical foundations for quantitative research and qualitative research respectively. For example, postpositivism consists of a set of ontological, epistemological and axiological assumptions in which ontological assumptions are privileged over other assumptions. As illustrated in Figure 4.1, the quantitative research approach is justified by postpositivist axiological assumptions, which are constrained by postpositivist epistemological assumptions, and these are ultimately underpinned by postpositivist ontological assumptions. In a similar vein, the constructivist/interpretivist ontological assumptions play a central role in the philosophical foundations of the qualitative approach. Such philosophical foundations, as Morgan and Winship (2007, p. 67) indicate, 'had a strong tendency not only to privilege epistemology over methods but also to emphasize ontological issues above all others'.

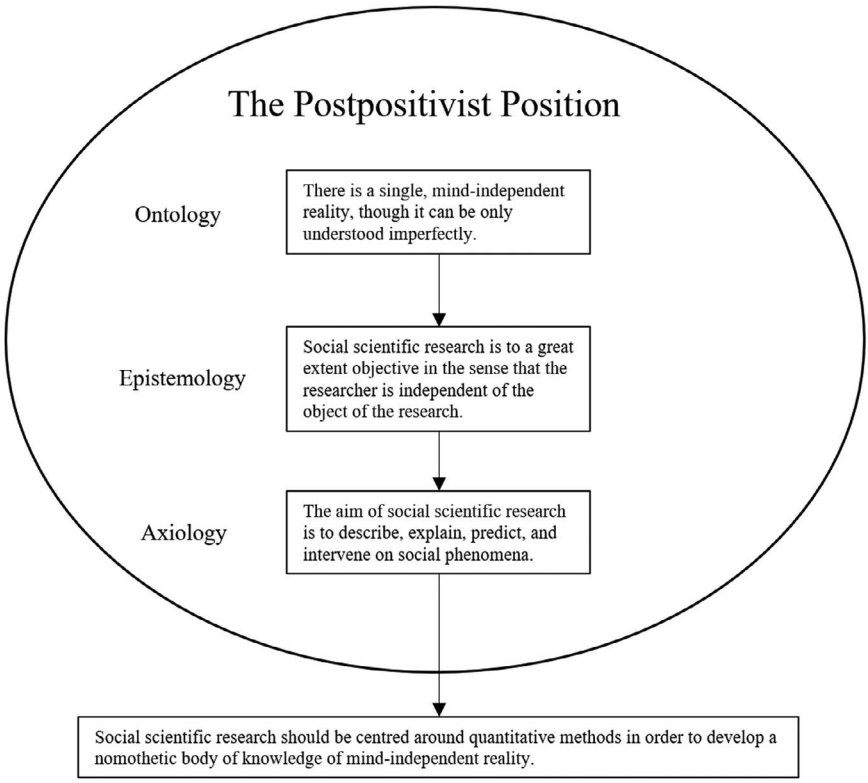

Figure 4.1 An example of 'ontology-oriented' philosophical foundations.

By contrast, pragmatism, the dialectical position, dialectical pluralism and the transformative position provide what we shall call *axiology-oriented* philosophical foundations. For example, the transformative position consists of a set of ontological, epistemological and axiological assumptions in which axiological assumptions are privileged over other assumptions. As illustrated in Figure 4.2, transformative axiological assumptions lead to transformative ontological assumptions, and both the transformative axiological and ontological assumptions underpin transformative epistemological assumptions. These assumptions then support mixed methods research (Mertens et al., 2010, p. 199). As Mertens (2010, p. 470) puts it, 'The axiological belief is of primary importance in the transformative paradigm and drives the formulation of the three other belief systems (ontology, epistemology, and methodology)'. In a similar vein, axiological assumptions play a central role in pragmatism, the dialectical position, dialectical pluralism and the performative position. Critical realism is the only position that provides ontology-oriented foundations for mixed methods research specifically.

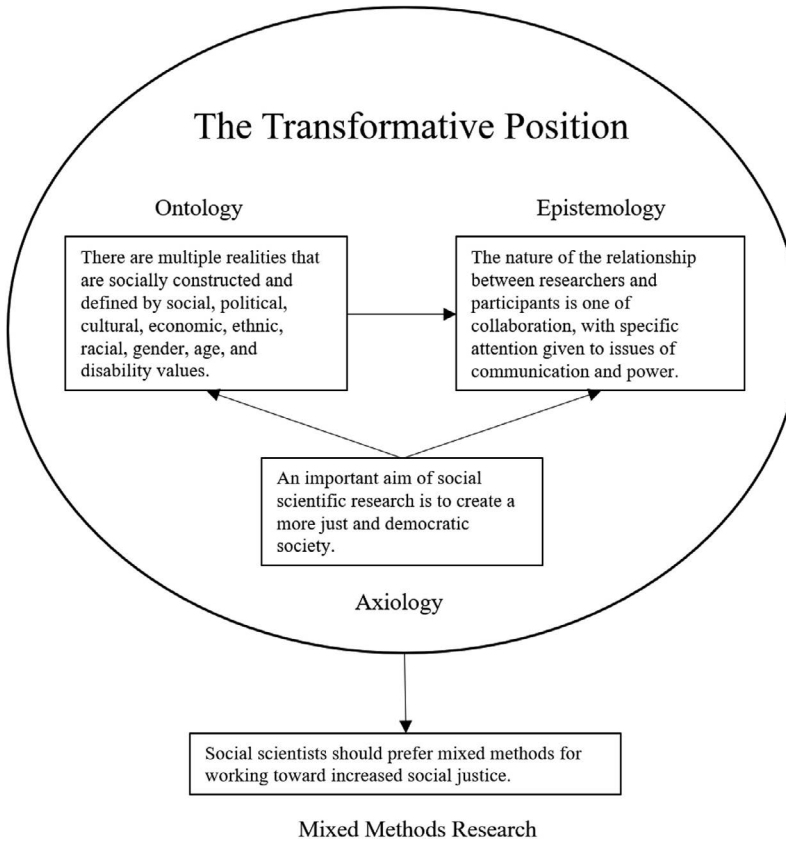

Figure 4.2 An example of 'axiology-oriented' philosophical foundations.

We now turn to a second distinction. Philosophical foundations for mixed methods research need to motivate the mixing of quantitative and qualitative methods/data/designs. There are three ways in which this motivation can be provided:

Weak. Weak philosophical foundations admit the possibility of the integration of both quantitative and qualitative methods/data/designs, in some context of enquiry.

Moderate. Moderate philosophical foundations provide good reason to use mixed methods in social scientific research, in some context of enquiry.

Strong. Strong philosophical foundations justify the claim that mixed methods research ought to be used instead of other approaches in social scientific research, in some context of enquiry.

It is clear that weak philosophical foundations achieve less than moderate philosophical foundations, while moderate philosophical foundations achieve less than strong philosophical foundations. Showing that the integration of both quantitative and qualitative elements is possible does not necessarily provide good reason to mix these elements, while having good reason to integrate or mix quantitative and qualitative elements does not imply that, on balance, these elements ought to be combined.

Pragmatism provides weak philosophical foundations for mixed methods research. It warrants considering mixed methods research in a portfolio of different research designs. As Greene and Hall (2010, p. 138) summarise, 'whatever works; whatever can best engage and usefully inform the important practical problem at hand' is a viable methodology, according to pragmatism. It makes sense for social scientists to use mixed methods research wherever this kind of research design works. Pragmatism does not provide grounds to prefer mixed methods to purely quantitative methods, nor to purely qualitative methods, however. This is because there are situations in which the quantitative approach, or indeed the qualitative approach, might be perfectly adequate. In addition, it is not clear in advance which situations provide appropriate contexts for the use of mixed methods. Overall, pragmatism merely justifies considering mixed methods, rather than requiring their use. Therefore, pragmatism is at best, as Johnson and Onwuegbuzie (2004, p. 14) put it, 'an attractive philosophical partner for mixed methods research'.

The transformative position provides moderate philosophical foundations. If the aim of research is to increase social justice and there are multiple realities that are socially constructed and defined by social, political, cultural, economic, ethnic and racial features, as well as gender and age, there may be good reason to use a variety of quantitative and qualitative elements in (at least some) social scientific research. That said, as Mertens et al. (2010, p. 196) indicate, 'researchers who situate themselves within the transformative worldview do not necessarily use mixed methods'. Mixed methods research is just 'reflective of' the transformative position (Mertens et al., 2010, p. 199). Thus, the transformative position does not justify the stronger claim that an integration of quantitative and qualitative methods/data/designs is to be preferred over alternative methodologies.

The dialectical position, dialectical pluralism and the performative position all provide strong philosophical foundations. According to the dialectical position, any particular position (e.g., postpositivism) provides but one perspective, inevitably partial, on human phenomena. Given that social phenomena are complex, 'better understanding of this complexity can be attained with the use of more than one perspective' (Greene and Hall, 2010, p. 124). Therefore, the integration of both quantitative and qualitative methods/data is not only possible but also beneficial. This is also why Greene and Hall (2010, p. 139) argue that the dialectical position and mixed methods research are 'often the best match'. Similarly, both dialectical pluralism and the performative position encourage the integration of quantitative and qualitative elements.

Each of these positions faces certain limitations. Pragmatism provides philosophical foundations that are arguably too weak: that pragmatism is compatible

with mixed methods research provides scant motivation to use mixed methods research. However, it is hard to see how pragmatism could provide stronger foundations, because it is an empirical question as to which methodology works best in which context.

The transformative position provides good reason to use mixed methods in cases in which the purpose of social scientists is to increase social justice. However, there are other goals to social scientific research. The transformative position does not ground the use of mixed methods where there are other goals: it does not explain why diverse perspectives are important where the goal is epistemic, say, rather than transformative. Advocates of the transformative position can either accept this limitation or explore what motivates and justifies mixed methods in these other contexts.

Pragmatism, the dialectical position and performative position all face a problem of scope. Specifically, they face the challenge of specifying the precise range of circumstances in which the use of mixed methods research is beneficial. Or should mixed methods research be universally applied? It is far from obvious that it is always helpful to consider different ontologies, epistemologies, axiologies, and methods—it may be that doing so merely increases confusion and reduces the chance of progress.

Critical realism also faces an important limitation. Maxwell and Mittapalli's central argument is that critical realism presupposes a mechanistic account of causality, a realist account of mental phenomena, and a realist concept of validity, which in turn justify a process-based qualitative approach to causality, a critical approach to qualitative research and a mixed methods approach to validity (Maxwell and Mittapalli, 2010, pp. 154–156). However, their argument is questionable. Maxwell and Mittapalli's argument for critical realism can be formulated as follows:

MM1. The mechanistic account of causation is a realist account.

MM2. A commitment to the mechanistic account justifies a qualitative process-based approach to causality in the social sciences.

MM3. 'A process theory of causation does not require abandoning quantitative, variance-based methods for investigating causality; it simply requires recognition that process-based approaches are as legitimate as, and often complementary to, variance-based ones' (Maxwell and Mittapalli, 2010, p. 156).

MM. Therefore, the realist position justifies the use of mixed methods in causal enquiry in the social sciences.

It is true that proponents of the mechanistic account of causation do not have to abandon quantitative variance-based methods. But the mechanistic account of causation does not seem to require such an approach. Thus it is not clear why the use of mixed methods is preferable, rather than merely permissible. The most one can conclude is that critical realism admits the possibility of mixed methods in causal enquiry in the social sciences.

§17 Evidential Pluralism and Mixed Methods Research

Causal enquiry dominates the social sciences, as social scientists are primarily interested in studying causes and effects of social phenomena. For example, Murnane and Willett (2011, p. 26) note that in educational research the key aim is 'to know the answers to questions about cause and effect'. Mixed methods research is often employed to answer these causal questions (Johnson and Christensen, 2014, p. 51). We argue in this section that Evidential Pluralism can provide coherent philosophical foundations for the use of mixed methods research in causal enquiry.

Why mix methods? Evidential Pluralism motivates mixed methods research on the grounds that it justifies the use of both association studies and mechanistic studies, where available, and this requires integrating both quantitative and qualitative methods. Typically, association studies use quantitative methods and produce quantitative data. Association studies confirm the existence of an appropriate correlation and can provide indirect evidence of the existence of an appropriate mechanism (channels α_1 and α_2 in Figure 1.1). Mechanistic studies, on the other hand, can use either quantitative or qualitative methods to test specific mechanism hypotheses (channel μ_1); see Table 1.2 for example. Moreover, it is more informative to make use of both quantitative and qualitative mechanistic studies where available. This is because different methods suit different specific mechanism hypotheses. Quantitative methods can be helpful to confirm the hypothesis that a key variable D mediates a mechanism from putative cause A to putative effect B: they can help to test for a correlation between A and D and a correlation from D to B. Qualitative methods, meanwhile, can help to fill in richer details of the mechanism complex. For example, qualitative methods can help to identify and verify potential confounders, as well as key activities and entities on mechanism pathways, and the spatiotemporal organisation of a mechanism. Thus, quantitative and qualitative methods complement one another well for the purpose of confirming features of mechanisms.

Lindsay-Smith et al.'s study of social support and health of older adults illustrates this complementarity. Lindsay-Smith et al. (2018) use a mixed methods study design to investigate the mechanism hypothesis that social support affects loneliness and that this in turn makes an impact on the health of older adults. A longitudinal analysis (a quantitative method) is used to test for a correlation between social support and loneliness, while a focus group study (a qualitative method) is employed to 'obtain a deeper understanding of the underlying reasons behind [the correlation]' (Lindsay-Smith et al., 2018, p. 11).

In sum, it is precisely because Figure 1.1 captures the confirmation pathways for causal inference that one should exploit both quantitative and qualitative methods.

Two points are worth emphasising. Firstly, Evidential Pluralism provides normative motivation for mixed methods research. One ought to scrutinise both association and mechanistic studies in causal enquiry, and one ought to scrutinise all the relevant studies—not just those that use quantitative methods, nor just those that use qualitative methods. Thus Evidential Pluralism provides strong philosophical foundations for mixed methods research: it is better to make use of mixed methods in causal enquiry than to restrict oneself to the use of a single method.

Second, Evidential Pluralism only motivates the use of mixed methods research in *causal* enquiry (i.e., for the specific purposes of establishing and assessing causal claims). This is because Evidential Pluralism is a theory of causal enquiry. Evidential Pluralism leaves open the question of whether mixed methods should be used in contexts other than causal enquiry, and how mixed methods can be best motivated in these other contexts. For example, mixed methods research can be used for hypothesis generation, or to gain a general understanding of the social context, to construct narrative case histories or to ensure that certain stakeholders have a voice that is represented in the evidence base—these purposes differ from that of establishing and assessing causal claims. Perhaps one can appeal to pragmatism in some of these other contexts, or the transformative position, for example. If so, the foundations that Evidential Pluralism provides can be viewed as complementary to those provided by other positions. For example, a research project that uses mixed methods both for the purpose of furthering social justice and for the purpose of causal enquiry might appeal to the transformative position as well as to Evidential Pluralism in order to justify its use of mixed methods. These two positions are compatible and complementary.

Integrating quantitative and qualitative research. Thus far, we have suggested that Evidential Pluralism provides grounds for using both quantitative and qualitative methods in causal enquiry. We observe next that Evidential Pluralism can also provide practical guidance on how exactly to integrate quantitative and qualitative methods. Integration is often a challenge for the mixed methods researcher.

Figure 1.1 provides the structure of the integration task. One point at which qualitative and quantitative methods need to be integrated is in the assessment of specific mechanism hypotheses: the question is the extent to which key features of relevant mechanisms are confirmed by qualitative and quantitative methods. Another point of integration is in the assessment of the general mechanistic claim that there exists a mechanism complex that explains instances of the putative effect in terms of instances of the putative cause and that can account for the extent of the observed correlation. As we see in Figure 1.1, one needs to consider quantitative data from association studies at this stage, as well as the statuses of the specific mechanism hypotheses. Then there is the assessment of the claim that the putative cause and effect are correlated, conditional on any potential confounding variables suggested by prior evidence or theory. At this stage, quantitative data will usually be most relevant, although qualitative methods may also have an influence through channel μ_3. Finally, the status of the causal claim depends on both the status of the correlation claim and the status of the general mechanistic claim, and it is at this point that all the qualitative and quantitative data are fully integrated. §12 provides a more detailed overview of all these points of integration.

Examples. We close this chapter by considering two examples of mixed methods research, viewed from the perspective of Evidential Pluralism.

Weinstein's study of insurgent violence, which we will consider in some detail in §29, provides a good example of the use of mixed methods research to establish

correlation and mechanism. Weinstein (2007) proposes a theory to explain differences between the ways in which rebel groups employ violence. Weinstein argues that resources and financing are key causal determinants of the strategies of violence:

> rebel groups that emerge in environments rich in natural resources or with the external support of an outside patron tend to commit high levels of indiscriminate violence; movements that arise in resource-poor contexts perpetrate far fewer abuses and employ violence selectively and strategically.
>
> (Weinstein, 2007, p. 7)

Weinstein's theory not only predicts a correlation between the initial resources to which rebel leaders have access and rebel groups' use of violence, but also identifies key features of a mechanism linking these two variables. For instance, Weinstein argues that resources shape the membership profile of a rebel group, which in turn affects its internal organisation and the strategies it uses in war. In order to support his theory, Weinstein integrates qualitative interview-based studies of the rebel groups and community-level social histories with statistical analysis of original newspaper data sets on patterns of violence, in four case studies of rebel groups in Mozambique, Peru and Uganda. The quantitative data provides evidence of correlation that supports his theory at the national level, while the qualitative data offers more nuanced evidence of mechanisms at the subnational level. Note that Weinstein employs mixed methods both to test the general causal claim and also to assess specific mechanistic hypotheses.[10]

Another example is Ivankova and Stick's study of PhD students' persistence. Ivankova and Stick (2007) investigate factors that contribute to students' persistence in a remote-learning doctoral programme. In order to identify these factors, they use a two-phase study, starting with a quantitative approach and followed by a qualitative approach. The roles of the two approaches are clearly stated as follows:

> In this study, the quantitative data helped identify a potential predictive power of selected external and internal factors on the distributed doctoral students' persistence and purposefully select the informants for the second phase. Then, a qualitative multiple case study approach was used to explain why certain external and internal factors, tested in the first phase, were significant predictors of students' persistence in the program. Thus, the quantitative data and results provided a general picture of the research problem, while the qualitative data and its analysis refined and explained those statistical results by exploring the participants' views regarding their persistence in more depth.
>
> (Ivankova and Stick, 2007, p. 97)

10 For example, a combination of interview-based studies and statistical analysis is used to confirm the connection between Renamo's indiscriminate abuse of noncombatant populations and the financial backing of the Rhodesian government. For an in-depth analysis of Weinstein's study, see section §29.

Interestingly, Ivankova and Stick largely avoid causal terminology. They explicitly classify their study design as a 'sequential explanatory mixed methods design', i.e., as first identifying factors associated with persistence and then finding explanations of these associations. But it is apparent that their task is really causal: they use their results to make a series of recommendations for how to improve persistence in such programmes, and this move would only be warranted if the factors they have identified are causes—rather than merely correlates—of persistence.

From the point of view of Evidential Pluralism, there is no need for any reticence here with regard to causal claims. This is because Ivankova and Stick have done what they need to do to confirm causality. They used a quantitative association study to identify factors correlated with persistence in the programme and then used a qualitative mechanistic study to provide evidence that there are mechanisms that invoke these factors to explain persistence (or drop-out). Taken together, these studies provide some good evidence for causality.[11] In situations such as this, an appeal to Evidential Pluralism might give researchers the confidence to draw causal conclusions.

We saw in the last chapter that Evidential Pluralism motivates a new approach to evidence-based policy. In this chapter we have argued that Evidential Pluralism provides strong philosophical foundations for mixed methods research insofar as it is applied to causal enquiry. Evidential Pluralism can also provide guidance on how to integrate these quantitative and qualitative methods. These two chapters help to demonstrate that Evidential Pluralism has some important methodological consequences.

11 Whether they have done enough to *establish* causality is not a question that we shall attempt to settle here.

5 Objections and Responses

Evidential Pluralism has been somewhat controversial since its introduction in 2007. In relation to the social sciences, detractors have included Reiss (2009); Claveau (2012); Beach (2021) and Runhardt (2022), for example. In this chapter, we focus on four key objections to the application of Evidential Pluralism to the social sciences: the problem of sufficiency (§18), the problem of necessity (§19), the problem of causal monism (§20) and the problem of defining mechanisms (§21).

§18 The Problem of Sufficiency

One way to object to Evidential Pluralism is to criticise object pluralism (the top part of Figure 1.1), by arguing that establishing correlation and mechanism is insufficient to establish causality.

Reiss (2009, §4) adopts this strategy, for instance.[1] He employs the following hypothetical example. Suppose association studies establish a correlation between watching violent TV (*W*) and violent behaviour (*V*). Given that a correlation on its own fails to establish the causal claim that *W* causes *V*, Evidential Pluralism suggests that one should also consider evidence of mechanisms. Now suppose that some such evidence establishes that there is a mechanism by which watching violent TV leads to increased aggression, which in turn leads to violent behaviour. (For example, there may be a psychological mechanism according to which viewers identify with the aggressive characters on TV and think of the depicted scenarios as realistic, which then results in more violent behaviour in real life.) One might then conclude that *W* is a cause of *V*, since both a correlation and a mechanism are established. But this conclusion may well be erroneous. There may be

1 Note that Reiss is a conceptual pluralist about causation: he thinks there are multiple meanings of the word 'cause', and that evidence of mechanisms and evidence of correlation may latch on to different concepts of cause. This is a view that we do not subscribe to, for reasons developed in §5 and §20, and it is worth emphasising that Evidential Pluralism need not be accompanied by either conceptual pluralism (the view that we employ different concepts of cause on different occasions) or metaphysical pluralism (the view that there are multiple causal relations in reality). Although Reiss advocates conceptual pluralism, his objection also applies to the more standard view that evidence of mechanisms and evidence of correlation are evidence for the same concept of cause.

DOI: 10.4324/9781003143000-7

a counteracting mechanism by which *W* acts as a deterrent in some individuals, reducing *V*. Suppose these two mechanisms exactly cancel out and that the correlation between *W* and *V* is in fact attributable to an unmeasured confounder—socio-economic status, say. Then *W* is not a cause of *V* after all, and Evidential Pluralism seems to have led us astray.

There are three points to make in response to this objection. Firstly, as we noted in §1, Evidential Pluralism requires establishing a correlation conditional on potential confounders, not an unconditional correlation. Socio-economic status is an obvious potential confounder, so one would need to test for its influence. Hence, Evidential Pluralism would not deem correlation in the appropriate sense to have been established, after all. Hence, nor is the causal claim established. Thus Reiss' example is not in fact a counterexample to Evidential Pluralism.[2]

Second, as we emphasised in §1, one needs to establish that the *complex of mechanisms* linking the putative cause to the putative effect can account for the observed correlation. It is not enough to consider a single pathway of action: one needs to rule out masking by counteracting mechanisms. *If* it is established that there is a mechanism of action and *if* potential counteracting mechanisms can be shown not to cancel out the influence of this mechanism of action and *if* one can infer that the putative cause and effect are correlated conditional on potential confounders, *then* one is in a position to establish the causal claim. (The larger the observed correlation, the less likely it is that there remain unmeasured potential confounding variables and unconsidered counteracting mechanisms: if they were responsible for a large correlation, it would be likely that we would know about them and have marked them down as potential confounders/counteracters.) A psychological mechanism by which watching violence discourages violence is an obvious potential counteracting mechanism, so one would need to examine its influence. This is another respect in which Reiss' example fails to undermine Evidential Pluralism.

Thus, Reiss' objection is really an objection to a caricature of Evidential Pluralism, rather than to Evidential Pluralism itself. But there is a third point to make about Reiss' objection, which highlights an important aspect of Evidential Pluralism: Evidential Pluralism is an account of the epistemology of causality, not an analysis of causality. If it were an analysis of causality of the form *A causes B if and only if A and B are appropriately correlated and linked by an appropriate mechanism complex*, then a single hypothetical counterexample would refute it—and there are such counterexamples, as we noted in §1. But Evidential Pluralism is an account of causal enquiry, rather than an analysis of causality. Establishing

2 It is worth emphasising that Evidential Pluralism does not demand that one establish a correlation conditional on all *possible* confounders, including unforeseen confounders as well as those that, insofar as one can tell from current evidence, are potential confounders. It is arguably impossible to establish such a conditional correlation by means of non-randomised studies, as that would require conditioning on almost every variable. However, Evidential Pluralism does demand that one establish a correlation conditional on *potential* confounders, which, in Reiss' example, includes socio-economic status.

a claim is a fallible activity; that one might sometimes be mistaken about whether *A* causes *B* does not refute the thesis that the best method for establishing that *A* causes *B* requires establishing that *A* and *B* are appropriately correlated and linked by an appropriate mechanism complex. As long as errors are sufficiently rare, this method may yet lead to the optimal advancement of science. Requiring immunity from all possible scepticism when establishing a proposition would prevent that proposition from ever being established. To reiterate a point from §1: when establishing a claim, we must balance the need to avoid falsity with the need to establish truths.

§19 The Problem of Necessity

We now turn to a second objection to the application of Evidential Pluralism to the social sciences: an objection to the claim that establishing both correlation and mechanism is necessary for establishing causation.

Specifically, the necessity of establishing correlation has been challenged by Claveau (2012). Claveau considers an example drawn from economics and argues that the following causal claim was established without establishing correlation:

(C) The strictness of unemployment benefit eligibility (*S*) reduces the unemployment rate (*U*).

Claveau suggests that there is a consensus amongst economists that a monitoring and sanctions system for unemployment benefits has a powerful effect on the unemployment rate. Various studies have sought to show that benefit sanctions induce a sharp increase in the exit rate from unemployment to employment across different countries. In other words, the strictness of unemployment benefit eligibility is a negative cause of the unemployment rate across countries. Claveau (2012) suggests that this causal claim was established by mechanistic evidence alone:

> The reason why difference-making evidence was not relied on is rather trivial: there was no measure of [*S*] comparable across countries (and there is still none as far as I am aware). Economists drew on the available evidence, i.e. mechanistic evidence. And it seems that this evidence—including clear model predictions and micro-data evidence in line with them . . .—was sufficient to gather general support to the claim.
>
> (Claveau, 2012, p. 812)

Claveau's argument rests on two claims: (i) there is no universal quantitative measure of the strictness of the monitoring and sanctions; (ii) there is only mechanistic evidence for *C*. Because of (i), Claveau argues that a correlation between *S* and *U* across countries is difficult to establish. And thus, given (ii) and given that *C* is established, it can be concluded that establishing the existence of a correlation is not required to establish a causal claim.

We would resist Claveau's conclusions as follows.

Firstly, there was good evidence of correlation in this case. That there was no good universal measure of *S* readily available does not imply that *S* is unmeasurable and that there is no way of establishing correlation. Claveau (2012) challenges the use of a coefficient associated with the overall cost of 'active labour market policies' (ALMPs) as a good universal measure across the countries. He argues that ALMPs include many factors which are unrelated to *S*, such as placement services, subsidised training, and subsidised employment. However, this does not eliminate the possibility of a good measure of *S* in each particular country, even if a good universal measure is not to hand. Moreover, local measures of *S* can be used to establish a correlation with *U* that holds across countries. For example, van den Berg et al. (2004) show that the imposition of a particular sanction, namely additional job search assistance, is significantly positively associated with the transition rate from welfare to work in the Netherlands, while Lalive et al. (2005) show that warning about benefits reduction is positively correlated with the transition rate in Switzerland. These studies provide strong evidence of a correlation between *U* and (some realisation of) *S* across the Netherlands and Switzerland. Thus, a correlation can be established across countries even if (i) holds.

Second, in this case the evidence of mechanisms was arguably not enough on its own to establish causation generally. At best, the existence of a mechanism from benefit sanctions to employment was established in the USA, UK, the Netherlands, and Switzerland (Johnson and Klepinger, 1994; Dolton and O'Neill, 1996; van den Berg et al., 2004; Lalive et al., 2005). More would need to be done to establish a general mechanistic claim that holds more widely across countries: it would need to be shown that the mechanisms are extrapolable to other countries. Thus it appears that the causal claim has not been established generally, but only in specific countries. It seems to be too hasty to infer that *S* reduces *U* generally.[3]

In sum, it appears, contra Claveau, that claim *C* may not have been generally established, and that correlation was established in those countries in which *C* was established (with the help of mechanistic studies, as per channel M_3 in Figure 1.1). Thus Evidential Pluralism fits this case after all.

Even if Claveau's proposed counterexample is unsuccessful, one might think that there are other counterexamples to the necessity of establishing correlation. For example, one might think that the use of process tracing allows researchers to establish causal relationships in political science without needing to establish correlation. We argue that this is not the case in §30.2.

3 It is worth noting that Claveau's 'mechanistic evidence' does not coincide with what we call 'mechanistic studies'. For Claveau (2012, p. 810), '[t]he first step in getting to mechanistic evidence is to redescribe the two relata [i.e., cause and effect] at a lower level'. In the case of unemployment benefits and unemployment rates, Claveau takes mechanistic evidence for *C* to be evidence that *S* reduces *U* in the particular countries, say, the USA. However, such evidence would not on its own constitute a mechanistic study, as it does not shed light on features of a mechanism linking *S* to *U*.

§20 The Problem of Causal Monism

Another objection to Evidential Pluralism challenges causal monism, i.e., the view that there is a single concept of cause. Causal pluralists hold that there are multiple concepts of cause, and, typically, that different kinds of evidence are used to establish different kinds of causation. Evidential Pluralism, on the other hand, seems to presuppose causal monism: it takes different kinds of evidence to confirm a single causal claim. If causal pluralism were correct, that would threaten to undermine Evidential Pluralism.

Some have argued that causal pluralism is more appropriate to the social sciences (e.g., Reiss, 2009; Goertz and Mahoney, 2012; Vaidyanathan et al., 2016; Crasnow, 2019; Maziarz, 2020; Rohlfing and Zuber, 2021). For example, Goertz and Mahoney (2012) argue that there are two concepts of cause in the social sciences—one that underpins the quantitative approach and the other the qualitative approach. Goertz and Mahoney's argument rests on their own interpretation of David Hume's theory of causation. Hume defines cause as follows:

> [W]e may define a cause to be an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed.
>
> (Hume, 1748, 7.2.29)

Typically, the first part of this definition is regarded as the original formulation of the regularity theory of causation, while the second part is viewed as a precursor to the counterfactual theory of causation (Lewis, 1973). Goertz and Mahoney (2012, p. 76) follow this received view by calling the first part 'the constant conjunction definition' and second part 'the counterfactual definition'. However, Goertz and Mahoney also argue that these two definitions correspond to the views on causation of the quantitative and qualitative traditions in the social sciences respectively:[4]

> Hume's famous quotation contains two definitions of causation. Definition 1 suggests a constant conjunction between cause and effect, such that effects always follow causes. This definition assumes many cases and has affinities with quantitative views on causation. Definition 2 suggests a counterfactual view of causation, in which the absence of a cause leads to the absence of an outcome. This definition is built around a single case and has important linkages to qualitative views of causation.
>
> (Goertz and Mahoney, 2012, p. 81)

In addition, Runhardt (2022) distinguishes two concepts of causality in political science: statistical correspondence and mechanistic generality, while Maziarz

4 Goertz and Mahoney (2012, pp. 81–82) suggest, however, that some qualitative researchers, especially those who use qualitative comparative analysis, 'may gravitate' towards the constant conjunction definition.

(2020) argues that economists are causal pluralists.[5] If there were different concepts of cause in the social sciences, the application of Evidential Pluralism to the social sciences would be challenged, because it would no longer be clear that different objects of evidence and different kinds of study confirm the same causal claim.

In response, we would suggest that the idea that causal pluralism is required in the social sciences often stems from an unwarranted inference from methodological diversity to causal pluralism:

> *Methodological diversity.* There are a variety of methods for causal enquiry in the social sciences.
>
> *Causal pluralism.* There are multiple concepts of cause in the social sciences.

For example, Goertz and Mahoney argue that two different concepts of cause in the social sciences are required to account for the presence of two distinct approaches to establishing causal claims in the social sciences (the quantitative approach and qualitative approach). We would argue that although there is indeed methodological diversity in the social sciences, this is perfectly well explained by Evidential Pluralism, which appeals to a single concept of cause. This is because Evidential Pluralism motivates considering both association studies and mechanistic studies, and these studies are produced by a range of different methods (§17). Thus methodological diversity is a result of the need for different kinds of study, not the result of any ambiguity in the word 'cause'. Causal pluralism is not needed to account for methodological diversity.

Consider Weinstein's study on rebellion and violence. Weinstein (2007) uses both statistical techniques and ethnographic methods to support the claim that the initial conditions facing rebel leaders are a cause of their strategy of violence. As we explain in §29, statistical techniques and ethnographic methods are not used to establish different causal claims. Rather they are used to obtain different types of evidence—evidence of correlation and evidence of mechanisms—for the same causal claim.

Thus, any inference from methodological diversity to causal pluralism is deeply problematic. As Crasnow (2011, p. 47) puts it, '[P]luralism about methodology need not commit us to a conceptual pluralism about causes'.[6]

Thus far, we have argued that methodological diversity can be explained by Evidential Pluralism, obviating the need to posit causal pluralism. This at least puts causal monism on an equal footing with causal pluralism. But there are several reasons to think that causal monism is preferable to causal pluralism. We

5 See §27 for further discussion of causal pluralism in economics and §30 on causal pluralism in political science.

6 Interestingly, Crasnow changed her mind about this. She later argued for methodological pluralism in political science and claimed that her 'methodological pluralism . . . minimally requires a conceptual pluralism about causality' (Crasnow, 2019, p. 41). We side with the early Crasnow.

encountered some reasons to doubt causal pluralism in §5: causal pluralism conflicts with our usage of causal claims, which does not treat the word 'cause' as ambiguous; it conflicts with Evidential Pluralism itself, which has independent support; it faces counterexamples; and it faces the problem of explaining what it is that different concepts of cause have in common that qualifies them as concepts of *cause*. Causal monism, on the other hand, is not susceptible to these problems, at least when epistemic causality is used to underpin causal monism. Causal monism is clearly also more parsimonious than causal pluralism: it only posits one concept of cause, rather than two or more. This provides a further reason to prefer causal monism over causal pluralism.

Arguably, causal pluralism also provides a less coherent account of the actual practice of social science research. Suppose there are two different concepts of causation, say, difference-making causation and mechanistic causation. If so, when social scientists say '*A* causes *B*', the question arises as to whether they are really saying '*A* $\text{causes}_{\text{DM}}$ *B*' or '*A* $\text{causes}_{\text{Mech}}$ *B*'. For example, Donohue and Levitt (2001) show both that there is a robust correlation between the legalisation of abortion in the USA in 1973 and the drop of crime rates in the USA in the early 1990s, and that there are mechanisms linking the former to the latter.[7] What can we conclude from this? Is the legalisation of abortion in the USA in 1973 a cause of the drop of crime rates in the USA in the early 1990s? If so, in what sense?

1 Does the legalisation of abortion in the USA in 1973 cause_{DM} the drop of crime rates in the USA in the early 1990s?
2 Does the legalisation of abortion in the USA in 1973 $\text{cause}_{\text{Mech}}$ the drop of crime rates in the USA in the early 1990s?
3 Or, does Donohue and Levitt's study suggest a new concept of causation?

It seems that a commitment to a plurality of concepts of cause in the social sciences leads to a confusion of further questions. What is worse, causal pluralism may lead to a problem of incommensurability. As Gerring argues,

> If causation means different things to different people then, by definition, causal arguments cannot meet. If A says that X_1 caused Y and B retorts that it was, in fact, X_2 or that Y is not a proper outcome for causal investigation, and they claim to be basing their arguments on different understanding of causation, then these perspectives cannot be resolved; they are incommensurable.
>
> (Gerring, 2005, p. 165)

As Gerring (2005, p. 165) argues, causal pluralism 'may over-state the ontological, epistemological, and/or logical different-ness of causal explanations in the social sciences'.

In sum, there are a wide variety of considerations to which one can appeal to fend off the challenge of causal pluralism. We will revisit the question of causal

7 See Section §24 for more discussion of this case.

pluralism in the context of economics (§27) and political science (§30.1), arguing against causal pluralism there. In Chapter 9, on the other hand, we will see that the concept of cause employed in law does differ from our usual concept of cause in a key respect, but we will argue that Evidential Pluralism can provide an account of causal enquiry in law that mirrors its account of causal enquiry elsewhere.

§21 The Problem of Defining Mechanisms

Evidential Pluralism faces the problem of how to understand mechanisms. As Mahoney (2001, p. 578) complains, 'a good deal of confusion currently surrounds the precise meaning of causal mechanism'.

The following quotes give a flavour of the lack of consensus on the question of how to define mechanisms in the social sciences:

> [Social mechanisms are] social processes having designated consequences for designated parts of the social structure—which articulate the expectations of those in the role-set sufficiently to reduce conflicts for the occupant of a status. (Merton, 1968, p. 43)
>
> [A mechanism is what] describes the process by which one variable influences the other, in other words, how it is that *X* produces *Y*. (Kiser and Hechter, 1991, p. 5)
>
> Mechanisms in a theory are defined here as bits of theory about entities at a different level (e.g. individuals) than the main entities being theorized about (e.g. groups), which serve to make the higher-level theory more supple, more accurate, or more general. (Stinchcombe, 1991, p. 367)
>
> A causal mechanism . . . is a series of events governed by lawlike regularities that lead from the explanans to the explanandum. (Little, 1991, p. 15)
>
> A social mechanism is an integral part of an explanation which (1) adheres to the four principles [i.e., action, precision, abstraction, reduction], and (2) is such that on the occurrence of the cause or input, *I*, it generates the effect or outcome, *O*. (Hedström and Swedberg, 1998, p. 25)
>
> A [social mechanism] is . . . the well-articulated set of causes responsible for a given social phenomenon. (Boudon, 1998, p. 172)
>
> Roughly speaking, mechanisms are frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences. (Elster, 1998, p. 45)
>
> [A] causal mechanism is a plausible hypothesis, or set of plausible hypotheses, that could be the explanation of some social phenomenon, the explanation being in terms of interactions between individuals and other individuals, or between individuals and some social aggregate. (Schelling, 1998, pp. 32–33)

> Social mechanisms in particular are usually thought of as complexes of interactions among individuals that underlie and account for aggregate social regularities. (Steel, 2004, pp. 57–58)

This lack of consensus might lead to concerns about the notion of mechanism to which Evidential Pluralism appeals. Does the concept of mechanism encapsulated in Evidential Pluralism capture all these different senses of mechanism in the social sciences? If not, is that a problem?

We would suggest that the concept of mechanism in Evidential Pluralism captures most of the usages of 'mechanism' in the social sciences. As we can see from the examples quoted earlier, many definitions of 'mechanism' can be classified into two groups. One group of definitions takes mechanisms to be complexes consisting of interactions among lower-level entities (e.g., individuals) that produce aggregate patterns at some higher level (e.g., at the population level). The other group takes a mechanism to be a process or a chain of variables. It is evident that Stinchcombe's, Schelling's, and Steel's definitions fall into the first group, while Kiser and Hechter's and Little's definitions fall into the second group.

These two groups of definitions correspond well to the two main senses of mechanism to which we appeal in §1: complex-systems mechanism and mechanistic process. A complex-systems mechanism consists of entities and activities organised in a way that they are responsible for some phenomenon to be explained. This notion fits well with the definitions in the first group. A mechanistic process, defined as a spatiotemporally contiguous process along which a signal can be propagated, fits the definitions in the second group. Thus, the plurality of definitions of 'mechanism' in the social sciences does not seem to pose a serious challenge to Evidential Pluralism, which construes a mechanism to be a complex-system mechanisms or a mechanistic process or a combination of the two.[8]

Nevertheless, a further worry may remain: if a mechanism is construed as a mechanistic process, and a mechanistic process is itself understood as a chain of causal relationships, does Evidential Pluralism offer an account that is conceptually distinct from other accounts? Marchionni and Reijula (2019), for example, express this kind of concern. Suppose then that a mechanism is defined as a chain of pairwise correlated mediating variables that link an independent and dependent variable (Kiser and Hechter, 1991). One might wonder whether such a definition offers anything beyond the concept of correlation—such mechanisms seem to be conceptually reducible to correlations. For example, Mahoney suggests,

> Causal mechanisms as intervening variables must be identified and analyzed with correlational tools. Indeed, a variable's status as a "mechanism" as opposed to an "independent variable" is arbitrary. With this definition, then, a correlation is "explained" simply by appealing to another correlation of observed variables. (Mahoney, 2001, p. 578)

8 For further discussion of the problem of defining 'mechanism', see §26.

If Mahoney is correct, it seems that there is no difference between mechanism and correlation from a conceptual point of view. This appears to pose a threat to Evidential Pluralism, which rests heavily on the distinction between mechanism and correlation.

In response, it suffices to observe that Evidential Pluralism rests on a distinction between *evidence of* correlation and *evidence of* mechanisms, and that this distinction is tenable whether or not correlation and mechanism are conceptually independent of one another. The rationale behind Evidential Pluralism is that establishing correlation is insufficient on its own to establish a causal claim and that evidence of correlation and evidence of mechanisms are complementary in causal enquiry. Evidence of mechanisms remains complementary to evidence of correlation, even if mechanisms were to be definable in terms of correlations. Consider a simple sketch of a social mechanism: one's socioeconomic status affects one's eating habits which in turn influence one's health. Some might suggest that this mechanism is merely a chain of two correlations from a conceptual point of view: one is between socioeconomic status and eating habits, the other between eating habits and health. This view of mechanisms is clearly questionable, but even if it were true, it would not undermine Evidential Pluralism. Consider two scenarios. In the first scenario, there is only evidence of a net correlation between socioeconomic status and health. In the second scenario, there is evidence of two further correlations: a correlation between socioeconomic status and eating habits and a correlation between eating habits and health. It is clear that the second scenario provides stronger support for the claim that socioeconomic status is a cause of health, because it provides some evidence that the correlation between socioeconomic status and health is not simply spurious, or attributable to considerations other than causation (see Table 1.1). Thus, Evidential Pluralism is independent of any views of the conceptual or reductive basis of mechanisms.

Beach (2021) puts forward a related objection to Evidential Pluralism. He identifies three different approaches to mechanisms in the social sciences: a minimalist approach, an in-depth approach and a realist approach. The minimalist approach assumes that a mechanism is a chain of causal relationships. Accordingly, a mechanism linking X to Y is assessed by 'analyzing the difference that variation in the intervening variable(s) has for values of Y across a large set of cases, controlled for other variables' (Beach, 2021, pp. 8905–8906). The in-depth approach assumes that a mechanism in the social sciences consists of entities and activities. Accordingly, a mechanism is examined by unpacking 'distinct parts composed of social actors (aka entities) engaging in activities' (Beach, 2021, p. 8909). The realist approach assumes that studying the social world is fundamentally different from the natural world because the social world is constructed and reconstructed by its social actors. Accordingly, it aims at developing detailed explanatory accounts of how processes actually play out in some particular contexts with an emphasis on assessing 'the intentions and motivations that actors had with the activities they performed, as well as their understanding of the underlying situational logic' (Beach, 2021, p. 8914).

Beach argues that each of these accounts of mechanism poses a problem for Evidential Pluralism. If establishing a mechanism means establishing a mechanism in the minimalist sense, then applying Evidential Pluralism to the social sciences is easy but superficial, because the detail of the mechanism under investigation remains hidden—little is said about social actors and their activities. If establishing a mechanism means establishing a mechanism in the in-depth sense, then applying Evidential Pluralism to the social sciences is fairly limited, because the evidence from an in-depth mechanistic study says little in relation to the overarching causal claim, making evidence of mechanisms 'difficult to communicate meaningfully with evidence of causal effects, which typically has a much broader scope' (Beach, 2021, p. 8902). Beach contends that evidence of mechanisms acquired by the in-depth approach only applies to the specific cases under investigation. Given that evidence of correlation 'typically deals with an average effect across cases, it is difficult to see what evidence of causation from a single case tells us about the cross-case effect' (Beach, 2021, p. 8911). If establishing a mechanism means establishing a mechanism in the realist sense, then applying Evidential Pluralism to the social sciences is almost impossible, according to Beach, because it is difficult to connect the detailed social context of action with evidence of correlation, which hinges on the assumption that there are stable, regular effects in the world (Beach, 2021, p. 8915). Thus for Beach, evidence of mechanisms is difficult to generalise under the realist approach.

Three points are pertinent here. Firstly, it is not necessary to opt for precisely one of Beach's views of mechanisms. Evidential Pluralism appeals to both mechanistic processes and complex-systems mechanisms, and so accords well with the first two senses of mechanism that Beach identifies. These mechanisms can include intentions, motivations and understandings of social actors, where necessary, so the realist approach can be accommodated too.

Second, establishing the existence of a mechanism does not require unpacking the black-box of a mechanism completely. Indeed, route α_2 of Figure 1.1 provides a means of confirming the existence of a mechanism without unpacking *any* detail of the mechanism, while confirmation channel μ_2 works by identifying key features of the mechanism complex, rather than its full detail. It is by no means impossible to identify key features of social mechanisms, and doing so clearly confirms the existence of these mechanisms. Evidential Pluralism concerns what is required for causal enquiry. Unpacking a mechanism in great detail is a different task entirely.

Third, correlation should not be conflated with a generic probabilistic dependence between variables at the population level. As we emphasised in §1, single-case causation requires considering probabilistic dependence with respect to a reference class which contains only a single individual. In other words, evidence of correlation does not necessarily have broader scope than a single case. Thus, it is perfectly possible to mesh evidence of correlation with evidence of mechanisms even when a mechanism is unpacked in detail, as per the in-depth or realist approaches.

In this chapter, we have responded to four principal objections to the application of Evidential Pluralism to the social sciences. This concludes our discussion of Evidential Pluralism insofar as it cuts across the social sciences. Next, we turn to questions that arise as Evidential Pluralism is applied to particular social sciences. As we do so, we will encounter and respond to further objections to Evidential Pluralism, and we will reconsider some of the earlier objections in more specific contexts.

Part III

Particular Social Sciences

6 Sociology

In sociology, as Morgan and Winship (2015, p. 3) observe, 'simple cause-and-effect questions are the motivation for much research'. Causal questions include: Does obtaining a university degree increase one's earnings? Does family dissolution have an impact on children's educational performance? Does neighbourhood of residence influence youth development? Because causal enquiry is so central to sociology, Evidential Pluralism can be fruitfully applied there, as we see in this chapter.

§22 Causal Enquiry in Sociology

In this section, we appeal to examples to illustrate the important roles of evidence of correlation and evidence of mechanisms in causal enquiry in sociology.

A clear case concerns the relationship between socioeconomic status and health status (e.g., House et al., 1994; Link and Phelan, 1995; Adler and Newman, 2002; Pampel et al., 2010; Phelan et al., 2010). There is a strong association between socioeconomic status and health status. For example, lower socioeconomic status is associated with the 14 major causes of death in the International Classification of Diseases (Illsley and Mullen, 1985). In addition, lower socioeconomic status has been shown to be associated with lower life expectancy, higher overall mortality rates and higher rates of infant and perinatal mortality (e.g., Dutton, 1986; Adler et al., 1994; Bosworth, 2018). However, it remains a question of debate whether socioeconomic status is a cause of health status. Sceptics typically argue that socioeconomic status is a placeholder variable for real causes of diseases that have not yet been identified. For example, Rothman (1986, p. 90) contends that socioeconomic status is 'a correlate of many causes of diseases'.

For some sociologists who argue in favour of a causal relationship between socioeconomic status and health, a strong and pervasive association between socioeconomic status and health merely provides 'a description of the social patterning of disease' (Link and Phelan, 1995, p. 82). These researchers acknowledge that in order to establish the causal claim that socioeconomic status is a cause of disease, one has to consider mechanisms as well as correlation (House et al., 1994; Phelan et al., 2004). As Link and Phelan (1995, p. 82) suggest, it is necessary to identify 'the direction of causation between social conditions and health and the

DOI: 10.4324/9781003143000-9

mechanisms that explain observed associations' for the purpose of 'establishing a causal role for social factors'.

With their collaborators, Link and Phelan identify a variety of mechanisms linking socioeconomic status to health status (Link and Phelan, 1995; Phelan et al., 2004, 2010). They show that people of higher socioeconomic status possess a wide range of resources, including money, knowledge, power and beneficial social connections, which shape health-enhancing behaviours (such as getting flu jabs, eating fruit and vegetables and exercising regularly) and which provide access to contexts that are associated with protective factors of health. In contrast, those who have lower-status jobs more commonly have 'job strain' (i.e., a combination of high job demands and low decision latitude), which is associated with coronary heart disease (Schnall et al., 1990); people with lower socioeconomic status are more likely to smoke and be overweight, with ensuing health problems (Lantz et al., 1998); and people with lower socioeconomic status experience greater residential crowding and noise, which is linked to poorer long-term memory and to reading deficits (Evans and Saegert, 2000).

Moreover, Link and Phelan argue that although there are various mechanisms linking socioeconomic status and health status, no individual mechanism is so dominant that it alone is responsible for the bulk of the observed association. In other words, multiple mechanisms are required to account for the full extent of the association between socioeconomic status and health status. Lutfey and Freese (2005, p. 1328) note that 'the association persists even while the relative influence of various proximate mechanisms changes'. It is in this sense that socioeconomic status is a fundamental cause of health status. This is the key idea of the so-called 'theory of fundamental causes' (Link and Phelan, 1995; Phelan et al., 2010).

Other sociologists argue against this causal claim. For example, Adams et al. (2003) question the causal relationship between socioeconomic status and health status by challenging the suggestion that an appropriate conditional correlation has been established. The debate about socioeconomic status and health shows that sociologists take both correlation and mechanism into account when they try to establish or assess a causal claim. The proponents of the theory of fundamental causes maintain that socioeconomic status is a fundamental cause of health status because there is both an established correlation and multiple established mechanisms which together can account for the magnitude of the correlation. On the other hand, its opponents challenge the causal claim by questioning the truth of the mechanism hypotheses (e.g., Cutler et al., 2011) or the existence of a correlation conditional on potential confounders (e.g., Adams et al., 2003; Adda et al., 2003). That both sides of the debate focus on evidence of correlation and evidence of mechanisms indicates that Evidential Pluralism captures the structure of causal enquiry in sociology.

Another example concerns the relationship between family background and educational attainment (Blau and Duncan, 1967; Bourdieu, 1973; Breen and Jonsson, 2005; Morgan et al., 2013). It has been argued that family background (including socioeconomic status, parental educational attainment and parental marital status) is a cause of educational attainment, alongside mental ability. This

causal claim is established by appeal to evidence of correlation as well as 'the purported existence of a specific causal mechanism that relates individuals' expectations and aspirations for the future to the social contexts that generate them' (Morgan and Winship, 2015, pp. 14–15). This mechanism was first confirmed by the Wisconsin Longitudinal Survey, led by Sewell, Haller and their collaborators (Sewell et al., 1969, 1970). The key idea of the mechanism hypothesis, also known as the Wisconsin model, is that family socioeconomic status and mental abilities affect the influence of significant others' (e.g., parents, teachers and friends) on the youth and the youth's own understanding of her ability. These in turn affect the youth's educational and occupational aspirations, which ultimately influence their subsequent educational attainment (see Figure 6.1).

Based on surveys of the educational plans of all high school students in Wisconsin in 1957 and 1964, Sewell et al. (1969, p. 90) conclude that 'aspirations are in fact performing mediational functions in transmitting anterior factors [e.g., family socioeconomic status and mental abilities] into subsequent behaviors'.

The claim that family background is a cause of children's educational attainment is further supported by other studies. It has been argued that the education of parents is a cause of a child's success in school (Haveman and Wolfe, 1995; Holmlund et al., 2011). Not only is the statistical correlation well confirmed, but key underlying mechanisms have also been identified. For example, Holmlund et al. (2011) argue that higher levels of educational attainment of parents leads to greater parental resources (e.g., money) that can be used to invest in children's education.

In addition, parental marital status has been cited as a cause of children's educational attainment. In many developed countries (e.g., OECD countries), both longitudinal and cross-sectional studies have shown that there is a correlation between family dissolution and children's low educational attainment (e.g., Jonsson and Gähler, 1997; Albertini and Dronkers, 2009; Hampden-Thompson, 2013). Those who claim that family dissolution causes children's low educational attainment typically appeal to these association studies. This has led to a persistent

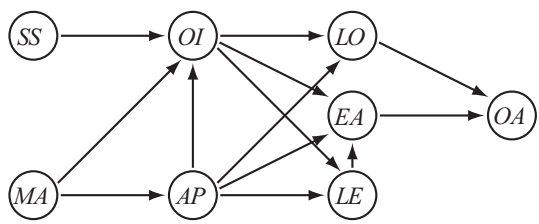

Figure 6.1 A simplified representation of the mechanism of Sewell et al. (1970). *SS*: Socioeconomic status; *OI*: Significant Others' Influence; *LO*: Level of Occupational Aspiration; *EA*: Educational Attainment; *OA*: Occupational Attainment; *MA*: Mental Ability; *AP*: Academic Performance; *LE*: Level of Educational Aspiration.

worry that there may be some unmeasured characteristics of families that also affect children's educational performance (e.g., Sandefur and Wells, 1999; Bernardi and Boertien, 2017). In addition, there are some discrepancies between different cohorts. For example, Erman and Härkönen (2017) find that the association varies across ancestries, being weakest among children with Chilean-born mothers and strongest among children with mothers born in Bosnia and Herzegovina. Thus, although the available association studies do establish a correlation, they arguably fail to establish causation. As Sandefur and Wells (1999, p. 331) argue, 'The existence of this association . . . does not establish causality'. By appeal to Evidential Pluralism, one can argue that this failure to establish causality stems from concerns that undermine confirmation along the α_2 channel of Figure 1.1.

Recently, efforts have been made to identify the mechanisms linking family dissolution to children's low educational attainment. Jonsson and Gähler (1997) argue that a longitudinal analysis of Swedish data supports the downward social mobility mechanism, the basic idea of which is that a child's lower educational attainment can be explained by the parent with the higher level of education leaving the household after divorce. They show that when educational and occupational attainment is low for custodial mothers but high for absent fathers, children are most negatively affected by a separation: about 54 percent of children in this category reach upper secondary education, compared to 78 percent when the mother also has a high educational level or social position. (Here the longitudinal analysis is used to provide evidence for the specific mechanism hypothesis.) Turunen (2014) identifies another mechanism: in Sweden, family reconstitution increases the complexity of family structure, which then has a negative impact on children's educational performance. Zeratsion et al. (2015) show that in Norway, parental divorce reduces children's educational ambitions, which negatively influences educational attainment.

As these examples show, sociologists pay serious attention to both evidence of correlation and evidence of mechanisms in causal enquiry. Although they do not conceptualise causal enquiry in terms of Evidential Pluralism, examples like the two considered here support object pluralism (the need to establish the existence of correlation and of mechanisms) as well as study pluralism (the need to consider both association studies and mechanistic studies). Using Evidential Pluralism to frame examples such as these can help us to understand the structure of cases of causal enquiry in sociological research.

§23 Sociologists' Methodological Reflections on Causal Enquiry

In this section, we observe that Evidential Pluralism fits well with certain methodological views on causal enquiry in sociology.

As discussed in §9, Goldthorpe's approach to causal enquiry in sociology highlights the significance of both correlation and mechanisms. Recall that Goldthorpe suggests that establishing causation in sociology requires three steps:

(i) establish the phenomena that form the explananda;

(ii) hypothesise generative processes at the level of social action;
(iii) test the hypotheses.

(Goldthorpe, 2001, p. 10)

This account clearly accords with Evidential Pluralism, as it can be thought of as requiring both evidence of correlation and evidence of mechanisms for causal enquiry.

Evidential Pluralism is also compatible with a methodology put forward by Morgan and Winship. Morgan and Winship (2015, pp. 447–448) propose a 5-mode account of causal enquiry:

Mode 1: Associational analysis.

Mode 2: Conditional associational analysis.

Mode 3a: Targeted analysis of the effects of one or more focal causes.

Mode 3b: Mechanism-based analysis.

Mode 4: All-cause structural analysis.

Morgan and Winship maintain that causal enquiry typically begins with an assessment of whether there is an association between the putative cause and effect of interest. In their words, establishing association is 'a precondition for subsequent causal enquiry' (Morgan and Winship, 2015, p. 447). They also suggest that after establishing an unconditional association, it is customary to re-evaluate the association after conditioning on other observed variables, including those that are thought to determine the effect and may be related to the putative cause, and those that are thought to determine the cause and may be related to the effect. Next, one needs to undertake a targeted analysis by elaborating a directed graph so as to eliminate confounding via back-door paths.[1] Thereafter, one needs to identify a mechanism linking the cause to the effect by introducing intervening variables between the cause and the effect. Ideally, causal enquiry is concluded by identifying 'all causes in chains of causality from the causal variables to the outcome variable, eliminating or neutralizing all confounding in order to identify all parameters of full causal systems' (Morgan and Winship, 2015, p. 448).

This 5-mode approach echoes Evidential Pluralism in many respects. First, it highlights the significance of both evidence of correlation and evidence of mechanisms. Morgan and Winship also emphasise the significance of conditional correlation. Mode 2 clearly requires the establishing of a correlation in order to 'eliminate obvious sources of confounding' (Morgan and Winship, 2015, p. 447).

1 A back-door path is a path from *A* to *B* that proceeds backwards from *A* to one of its causes, and then forwards to *B* (Morgan and Winship, 2015, §4.1).

This is exactly what Evidential Pluralism maintains: an unconditional association is not enough; what is required is a correlation conditioned on potential confounding variables. Moreover, for Morgan and Winship, all the five modes are 'valuable' (Morgan and Winship, 2015, p. 447). In particular, mechanism-based analysis is complementary to associational analysis. Thus, both the 5-mode approach and Evidential Pluralism highlight that the complementarity of association studies and mechanistic studies.

However, it might seem that not all sociologists' methodological reflections fit perfectly with Evidential Pluralism. For example, Klein (1987, pp. 22–25) identifies eight criteria for causality in sociology:

1 The causal relata must be variables.
2 The variables should vary over time.
3 The relationship between variables must approach a constant conjunction.
4 The relationship must be stronger under specified conditions.
5 The relationship must be asymmetrical.
6 The time lag between cause and effect must be reasonably short.
7 The relationship must not be spurious.
8 The relationship must weigh the relative contributions of causal elements.

Criterion 3 stresses the importance of association, but none of these principles explicitly invokes evidence of mechanisms. That said, when discussing Criterion 4, Klein (1987, pp. 23–24) does suggest that establishing correlation alone is insufficient. He argues that some 'specified conditions' must be taken in consideration, though he does not say exactly what these conditions are. Klein (1987, p. 24) does highlight the significance of mechanisms when 'asserting causes': in order to establish the time lag between cause and effect, sociologists are encouraged to 'posit or discover intervening mechanisms'. Thus for Klein, establishing the existence of a mechanism can help to establish a causal claim. Consequently, Klein's principles are somewhat compatible with Evidential Pluralism.

In sum, Evidential Pluralism fits a range of methodological reflections on causal enquiry in sociology. That said, there are some substantial differences between Evidential Pluralism and these sociologists' approaches (see also section §10 on this point). For example, both Goldthorpe's and Morgan and Winship's approaches are forthright in providing a sequential procedure for causal enquiry. For Goldthorpe, there is a particular order of activities in sociological causal enquiry, namely, from (i) to (iii). Likewise, Morgan and Winship also present a step-by-step recipe for causal enquiry, which proceeds from associational analysis, through conditional associational analysis, and targeted analysis/mechanism-based analysis, to structural analysis. Evidential Pluralism does not insist upon any particular ordering of steps of causal enquiry. What Evidential Pluralism provides is a normative account of the evidence that is needed in order to establish or assess a causal claim: this evidence can be obtained and processed in a wide variety of ways. In particular, whether evidence of correlation or evidence of mechanisms is obtained first does not really matter. This can be construed as an advantage of

Evidential Pluralism over Goldthorpe's and Morgan and Winship's methodological reflections: Evidential Pluralism captures a wider variety of actual practice of causal enquiry in sociology—practice which does not conform to the clearly delineated procedures suggested by Goldthorpe or Morgan and Winship. Nevertheless, Evidential Pluralism, as a normative account of the epistemology of causation, does explain why both these particular methodological approaches are basically sound. By making the role of evidence of mechanisms explicit, it can also explain the importance of several of Klein's criteria, and might be used as a way to flesh out Klein's approach to causal enquiry.

In this chapter, then, we have argued that Evidential Pluralism can be fruitfully applied to sociology. We have shown that Evidential Pluralism validates good examples of causal enquiry in sociology, and that it is largely compatible with certain sociologists' methodological reflections on causal enquiry. Overall, Evidential Pluralism motivates a unified approach to sociology. In sociology, there has been a division between social theorists and quantitative researchers. Quantitative researchers focus on statistical models and analyses, and pay rather little attention to theories and mechanisms. In contrast, social theorists are primarily concerned with concepts and theoretical frameworks and downplay the significance of quantitative findings. As Mahoney argues,

> Clearly, neither statistical researchers nor social theorists can afford to live without one another. Correlational findings are incomplete and not fully intelligible without an understanding of the mechanisms that generate those findings; by contrast, theories of causal mechanisms are entirely speculative until their power is revealed through empirical correlations. Hence, a complete science must strive to identify both correlations and causal mechanism.
>
> (Mahoney, 2001, p. 582)

Evidential Pluralism accords well with these remarks, as it justifies bringing both strands of sociological research together for the specific purposes of establishing and assessing causal claims.

7 Economics

The application of Evidential Pluralism to economics has already attracted some debate. Moneta and Russo (2014) and Maziarz (2021), for example, argue that Evidential Pluralism can be applied to economics, while Reiss (2009) and Claveau (2012) are more sceptical.[1] This chapter defends the application of Evidential Pluralism to economics. We argue that economists take both association studies and mechanistic studies into account when they try to establish causal claims, and we illustrate this study pluralism in §24 by means of two examples. In §25, we show that there are significant advantages to taking Evidential Pluralism seriously in economics. We examine the notion of mechanism in economics in §26 and argue that it is captured by the notion of mechanism to which Evidential Pluralism appeals. In §27, we respond to an objection to Evidential Pluralism that appeals to causal pluralism in economics.

§24 Causal Enquiry in Economics

As an example of the compatibility of economic research with Evidential Pluralism, consider Donohue and Levitt's study on legalised abortion and crime rates.[2]

Donohue and Levitt (2001) argue that the legalisation of abortion in the early 1970s was a cause of the decline in crime rates in the early 1990s in the United States. In order to establish this causal claim, they provide evidence of mechanisms as well as evidence of correlation. They find two mechanisms of action. The first is that legalising abortion reduces crime through smaller cohort sizes. The smaller cohort that results from legalised abortion means that when that cohort reaches the late teens or early twenties, there are fewer young males in their highest-crime years, and thus less crime.[3] The second mechanism stems

1 See §18 and §19 respectively for responses to the criticisms of Reiss and Claveau.

2 Although abortion and crime rates are not typical economic variables, this research is viewed as a classic example of economic research because of its use of research methods that are standard in economics.

3 It is shown that the legalisation of abortion leads to a drop in birth rates and that the crime rate would be expected to fall accordingly (Levine et al., 1999). For example, consider a town with a population of 10,000. Suppose that the birth rate before the legalisation of abortion is constant at 2%, while after the legalisation of abortion it is 1%. Also assume that the overall crime rate is initially 5%, and those who are aged 18–24 commit half of all crimes. All other things being equal, one can infer that the overall crime rate will drop by approximately 1%, 24 years after the legalisation of abortion.

DOI: 10.4324/9781003143000-10

from the fact that abortion has a disproportionate effect on the birth of those who are most at risk of engaging in criminal behaviour. Teenagers, unmarried women and the economically disadvantaged are all substantially more likely to seek abortions (Levine et al., 1999). Recent studies have found children born to these mothers to be at higher risk of committing crime in adolescence (Comanor and Phillips, 2002). Thus, the two mechanisms form a mechanism complex linking the legalisation of abortion in the early 1970s to the drops in crime in the early 1990s.

Donohue and Levitt look for evidence of correlation by focusing on the variations of national time series of crime and abortion, of differential crime patterns between states which legalised abortion early and other states, and of state abortion rates and the state crime rates. They show that the legalisation of abortion was associated with a subsequent drop in crime. All of violent crime, property crime and murder have fallen steadily since 1991, roughly the time the first cohort born would hit its criminal prime. Additionally, the five states that legalised abortion in 1970 saw drops in crime before the other 45 states and Washington DC, which legalised abortion in 1973. Also, higher rates of abortion in a state in the 1970s and early 1980s are strongly linked to lower crime over the period from 1985 to 1997. Moreover, the observed association holds conditional on various potential confounders, such as the level of incarceration, the number of police and measures of the state's economic well-being (e.g., the unemployment rate, income per capita and poverty rate). It is shown that there is no relationship between abortion rates in the mid-1970s and crime rate changes between 1972 and 1985, when the cohort directly affected by abortion legislation would have been very young. Almost all of the abortion-related crime decrease can be attributed to reductions in crime among cohorts born after abortion legalisation. In contrast, there is little change in crime among older cohorts, who were not affected by abortion legalisation. The correlation is further supported by the more recent study of Donohue and Levitt (2019).

It is clear that the way in which Donohue and Levitt confirm the causal claim about legalised abortion and crime rates accords well with Evidential Pluralism. Not only do they look for a conditional correlation, but they also seek relevant mechanisms and they show that these mechanisms can account for the extent of the observed correlation. What is more, Donohue and Levitt's criticisms of alternative causal explanations are compatible with the epistemological picture provided by Evidential Pluralism. For example, the reason that Donohue and Levitt dismiss factors such as the increasing use of incarceration and the rise in police numbers as the causes of the drop in crime rates is that these trends fail to exhibit an appropriate conditional correlation.

Donohue and Levitt's work sparked debate and controversy in the literature. Both their evidence of correlation and their evidence of mechanisms have been disputed (e.g., Joyce, 2004; Lott and Whitley, 2007; Chamlin et al., 2008; Foote and Goetz, 2008; Dills et al., 2010). Chamlin et al. (2008), for example, argue that there is no evidence that the legalisation of abortion led to a decline in the birth rate for teenage or unmarried women: i.e., they are sceptical of the evidence of mechanisms. Moreover, Lott and Whitley (2007) question whether Donohue and

Levitt adequately capture the mechanism complex from abortion to crime. As they point out, 'abortion can eliminate unwanted children and can benefit many women, but it can also make other women who are unable to bring themselves to have an abortion worse-off and more likely to have out-of-wedlock births' (Lott and Whitley, 2007, p. 324). Joyce (2004) challenges both the correlation and mechanisms. In response, Donohue and Levitt (2019) defend their causal claim with updated evidence, which is relevant to both correlation and mechanisms. That both sides of this debate focus on evidence of correlation and evidence of mechanism supports the view that Evidential Pluralism captures the structure of causal inference in economics. Whether or not Donohue and Levitt are correct, the debate indicates that good research needs to consider both association studies and mechanistic studies in causal enquiry.

Other cases in economics might seem to sit less well with Evidential Pluralism. Consider Raphael and Winter-Ebmer's study of unemployment and its relationship to crime. It has been shown that crime rates are strongly associated with unemployment rates. However, it is questionable whether unemployment causes criminal activity. Raphael and Winter-Ebmer (2001) argue that the decline in the unemployment rate is a cause of the drop in the property crime rate. They use ordinary least squares (OLS) regression to estimate the influence of unemployment rates on the rates of seven felony offences, by appeal to the following equation:

$$\text{Crime}_{it} = \alpha_t + \delta_i + \psi_i \,\text{time}_t + \omega_i \,\text{time}_t^2 + \gamma \text{Unemployed}_{it} + \beta X_{it} + \eta_{it}, \qquad (1)$$

where i and t index states and years, Crime_{it} is the log of the number of crimes per 100,000 state residents, Unemployed_{it} is the unemployment rate, X_{it} is a vector of standard controls, α_t is a year fixed effect, δ_i is a state fixed effect, time_t and time_t^2 are linear and quadratic time trends, ψ_i gives the state-specific coefficient on quadratic time trend, γ is the semielasticity of the crime rate with respect to the unemployment rate, β is the vector of parameters for the control variables in X_{it}, and η_{it} is the residual. Raphael and Winter-Ebmer control for a number of variables, including alcohol consumption, average income, drug consumption, and gun availability. The empirical results show that unemployment rates are positively associated with property crime rates. Based on this, Raphael and Winter-Ebmer (2001, p. 271) put forward a causal claim: 'a 1 percentage point drop in the unemployment rate causes a decline in the property crime rate of between 1.6 and 2.4 percent'. In contrast, the coefficient with respect to unemployment is negative for most violent crimes.

In addition, Raphael and Winter-Ebmer (2001, p. 267) use two-stage least squares (2SLS) method to re-estimate Equation 1 for the purpose of examining 'instrumental variables that determine state unemployment rates yet are unrelated to possible contaminating omitted factors', such as military contracts and oil costs. The results confirm that the unemployment rate is a positive and significant factor of the total property crime rate. Raphael and Winter-Ebmer contend that the 2SLS results reinforce the OLS results for property crime. On the other hand, the 2SLS results for violent crime are not strong. Therefore, Raphael and Winter-Ebmer conclude that unemployment is a cause of property crime but not

of violent crime. We refer to this causal claim as *UCPC* (Unemployment is a Cause of Property Crime).

It might seem that Raphael and Winter-Ebmer's study provides a counterexample to Evidential Pluralism because *UCPC* seems to have been established on the basis of evidence of correlation alone. However, it is controversial as to whether *UCPC* has been successfully established (see, e.g., Hojman, 2004; İmrohoroğlu et al., 2004). In their study of property crimes in the US, for example, İmrohoroğlu et al. (2004, p. 709) argue that 'the effect of unemployment on crime is negligible'. Thus, the correlation between unemployment and crime has been challenged.

Evidential Pluralism suggests a strategy for establishing *UCPC*: combine Raphael and Winter-Ebmer's association studies with suitable mechanistic studies (Cantor and Land, 1985; Greenberg, 2001; Burdett et al., 2003). Cantor and Land (1985) propose a mechanism hypothesis that links the unemployment rate to the crime rate through two main channels: criminal motivation and criminal opportunity (see Figure 7.1). Unemployment affects both criminal opportunity and criminal motivation, each of which has a positive effect on the other, and these in turn affect the crime rate.

There is some evidence for this mechanism complex. For example, the hypothesis that unemployment affects crime through criminal opportunity is confirmed by a 'search equilibrium framework' (Burdett et al., 2003). A different, simpler model that predicts a positive effect of unemployment on crime via criminal opportunity has been put forward by Calvó-Armengol and Zenou (2003) as follows.[4] Suppose that the individual expected returns on being involved in criminal activities are

$$\beta V - (1 - \beta)P - z \tag{2}$$

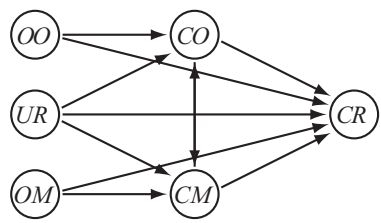

Figure 7.1 A simplified representation of the unemployment-crime mechanism complex of Cantor and Land (1985, p. 321). *OO*: Other Criminal Opportunity Factors; *CO*: Criminal Opportunity; *CR*: Crime Rate; *UR*: Unemployment Rate; *OM*: Other Criminal Motivation Factors; *CM*: Criminal Motivation.

4 Note that Burdett et al. (2003) do not use the word 'mechanism'. Nor do Calvó-Armengol and Zenou (2003) conceptualise Burdett et al.'s study in terms of mechanisms or mechanistic studies. From the perspective of Evidential Pluralism, however, Burdett et al.'s study provides evidence of mechanisms and can be classified as a mechanistic study. See §26 for more discussion.

where V is the plunder, $1 - \beta$ is the probability of being caught, P the corresponding penalty, and z is an idiosyncratic reluctance to commit crime, uniformly distributed on $[0, \bar{z}]$ where $0 < \bar{z} < 1$. Suppose b is the unemployment benefit. Then an unemployed individual accepts the offer of participating in a criminal activity whenever

$$\beta V - (1 - \beta)P - z > b, \tag{3}$$

that is, with probability

$$\pi = \beta V - (1-\beta)P - b, \tag{4}$$

assuming that $\bar{z}$ is high enough, that is,

$$\bar{z} > \beta V - (1-\beta)P - b, \tag{5}$$

and that

$$\beta V - (1 - \beta)P > b. \tag{6}$$

Suppose that unemployed workers learn of some available criminal activity that they could take part in with some exogenous probability α, that employed workers have no access to this information, and that criminals remain so unless they are in jail. Then, the crime rate evolves as follows:

$$C_{t+1} = \beta C_t + \alpha \pi u_t, \tag{7}$$

where u_t is the unemployment rate. Equation 7 implies that current unemployment ultimately leads to crime, at the steady state,

$$C = \frac{\alpha}{1-\alpha} \pi u \tag{8}$$

With this 'theoretical motivation' (which provides evidence of mechanisms) and evidence of correlation, Calvó-Armengol and Zenou (2003, p. 184) conclude that unemployment causes crime. Thus, Calvó-Armengol and Zenou's strategy accords with Evidential Pluralism.

To sum up, both Donohue and Levitt's study of legalised abortion and crime and Calvó-Armengol and Zenou's study of unemployment and crime indicate that good economic research does indeed consider both association studies and mechanistic studies when assessing causal claims.

§25 Benefits of Evidential Pluralism in Economics

We now move on to argue that there are two main benefits of Evidential Pluralism for economics: Evidential Pluralism can aid understanding and it can help to improve practice.

Firstly, Evidential Pluralism helps us to better understand causal enquiry in economics. As we saw in the abortion and crime rate example of the previous section, Evidential Pluralism can help us to identify the locus of disagreement where a causal claim is disputed—it can help us to understand whether it is a dispute about correlation or mechanisms. It can tell us whether the available evidence is informative about causation (see Chapter 3) and can tell us what sorts of evidence to obtain when evidence is lacking. Similar points can be made about the example of unemployment and crime.

Moneta and Russo (2014) have argued that Evidential Pluralism can also help us to interpret statistical models in econometrics. More specifically, they argue that Evidential Pluralism can help to determine whether a particular statistical model can be interpreted in causal terms. They classify statistical models in econometrics into two groups: associational models and causal models. Associational models are those statistical models which 'aim to provide an accurate and reliable description of how certain phenomena (chance events) are regularly associated among each other in their occurrence' (Moneta and Russo, 2014, p. 56), while causal models are what 'aim at uncovering causal relations' (Moneta and Russo, 2014, p. 55). For Moneta and Russo, associational models and causal models differ in three main respects: background knowledge, assumptions and methodologies (for a summary, see Table 7.1). Typically substantive extra-statistical knowledge plays little role in building associational models, whereas causal models are based on a variety of kinds of extra-statistical knowledge, such as general knowledge of the sociopolitical context and of the demographic characteristics of the population under analysis, and institutional knowledge. The assumptions of an associational model are statistical assumptions. For example, one might assume that the data generating process is such that observations result from independent and identically distributed random variables. In contrast, causal models have both statistical and extra-statistical assumptions (e.g., the direction of the causal relation). Regarding methodology, associational models are generally used to establish statistical dependencies in an inductive way by deriving conclusions about some general stochastic process from a specific set of data, while causal models can be used to draw causal inferences via a hypothetico-deductive approach, by formulating hypotheses out of background knowledge and data analysis, and then testing them. The fundamental difference between associational models and causal models is that associational models

Table 7.1 Associational models versus causal models (Moneta and Russo, 2014, p. 56).

	Associational models	*Causal models*
Background knowledge	Choice of variables	Causal context; theoretical knowledge; institutional knowledge, etc.
Assumptions	Statistical	Statistical; extra-statistical; causal
Methodology	Model-based statistical induction and H-D method	Model-based induction and H-D method

only convey statistical information, while causal models convey both statistical information and causal information.

Equation 1 in Raphael and Winter-Ebmer's study can be interpreted as an associational model. Extra-statistical background knowledge does play a role in choosing variables, especially the control variables X_{it}. However, once variables are chosen, the estimation of the regression function (i.e., Equation 1) is based only on statistical information. In other words, Equation 1 can be interpreted as devoid of causal information. On the other hand, Equation 1 can be interpreted as a causal model when it is augmented with theoretical knowledge of unemployed workers' behaviour given by Calvó-Armengol and Zenou (2003). It is in this way that, as Moneta and Russo (2014, p. 72) argue, 'a causal interpretation of statistical models is justified to the extent that statistical information is augmented into causal information, or in other words that evidence of difference-making is integrated with evidence of mechanisms'. Thus, a statistical model can be interpreted as a causal model when it is supported by both evidence of correlation and evidence of mechanisms. In this way, Evidential Pluralism can help us to interpret models in economics.

Williamson (2017) refines this distinction between kinds of models. One can distinguish association models, which model associations, causal models, which model causal relationships, and mechanistic models, which model mechanisms. As we saw in §13, some mechanisms—such as those that appeal to spatiotemporal organisation—cannot be captured straightforwardly as a network of causal relationships. Thus it can be important to distinguish mechanistic models from causal models. This tripartite distinction is also helpful for understanding the roles of models in causal enquiry. From the point of view of Evidential Pluralism, an association model can aid the process of establishing a correlation between *A* and *B*, while a mechanistic model can act as a specific mechanism hypothesis, positing key features of a mechanism linking *A* and *B* that can be tested by mechanistic studies. When correlation and mechanism are established, a causal relationship between *A* and *B* is established. This can then contribute to a causal model that includes other established causal relationships. Or the established causal relationship may contribute to a further mechanistic model that acts as a specific mechanism hypothesis for establishing some other causal relationship. Thus, Evidential Pluralism and this tripartite distinction can help us understand three possible roles that models can play in causal enquiry in economics.

Evidential Pluralism can also offer practical guidance in economics. This is because it provides a general account of how to establish and assess causal claims in the social sciences. To some extent, economics has been overwhelmed by association studies, and considering mechanistic studies in a systematic way can help to resolve debates that have not been settled by means of association studies. As we saw when considering the study of unemployment and crime in the previous section, mechanistic studies can provide crucial evidence for *UCPC*. Another case surrounds the enduring debate about the causal relationship between money and prices. Some contend that prices are causally determined by the amount of money in circulation (e.g., Black, 1970; Fama, 1980; Schmidt, 2003), while others argue

for the opposite conclusion: prices influence money supply (e.g., Hoover, 1991; Engle and Hendry, 1993; Fisher and Nicholetti, 1993). This controversy remains unresolved. Introducing mechanistic studies might be helpful here. For example, if one is able to propose a mechanism with some intermediate variables between money and prices, and the mechanism is well evidenced, this would help to identify the arrow of causality.

In addition, Maziarz (2021) has argued that Evidential Pluralism can resolve controversies when inconsistent causal hypotheses emerge from similar econometric models. Consider a controversy over tax and smoking intensity. Based on data from the National Health and Nutrition Examination Survey (NHANES) in the US, Adda and Cornaglia (2006) use OLS regressions to estimate the influence of the level of taxation, age, sex, race, education attrition, household size and the number of smokers in the family on the level of cotinine[5] (*cot*) measured in the *i*th sampled smoker from state *s* at time *t*, and on estimated cigarette consumption (*c*), by means of the following equations, where the α_n and β_n are estimated parameters:

$$\log cot_{ist} = \beta_0 + \beta_1 \log tax_{st} + \beta_2 X_{ist} + \beta_s + \beta_t + v_{ist}, \quad (9)$$

$$\log c_{ist} = \alpha_0 + \alpha_1 \log tax_{st} + \alpha_2 X_{ist} + \alpha_s + \alpha_t + u_{ist}, \quad (10)$$

OLS results show that a 1% increase in taxes leads to a 0.47% increase in the intensity of smoking. The tax elasticity of the number of cigarettes (α_1) is -0.2, while the elasticity of cotinine (β_1) is approximately 0. In other words, an increase in taxes has a negative influence on cigarette consumption but has no impact on nicotine consumption. Thus, Adda and Cornaglia argue that the efficacy of cigarette tax increases as a policy tool aimed at reducing nicotine consumption is limited. However, Abrevaya and Puzzello (2012) challenge Adda and Cornaglia's conclusion. By conducting a replication study of Adda and Cornaglia's research with a larger sample from NHANES (covering the same period but including more states), Abrevaya and Puzzello obtain different results: the estimated cigarette and cotinine elasticities are both negative (-0.3271 for cigarettes and -0.2091 for cotinine) for the restricted sample of long-term smokers, whereas they remain statistically insignificant for the unrestricted sample of smokers. Abrevaya and Puzzello argue that the limited variability of taxation across the states only suggests that the NHANES dataset used by Adda and Cornaglia is not suitable for estimating tax elasticities of cigarette consumption and nicotine levels.

Such a controversy, argues Maziarz, can be resolved by taking evidence of mechanisms into consideration. There are two hypotheses concerning the mechanism of an addict's decision process. One is derived from the classical theory of rational addiction, developed by Becker and Murphy (1988). According to this mechanistic hypothesis (*RAM*1), an agent's decision regarding the consumption of an addictive good is caused by external determinants (e.g., the price of the

5 Cotinine is an alkaloid found in tobacco and the major metabolite of nicotine. It is widely used as a measure to test recent tobacco consumption.

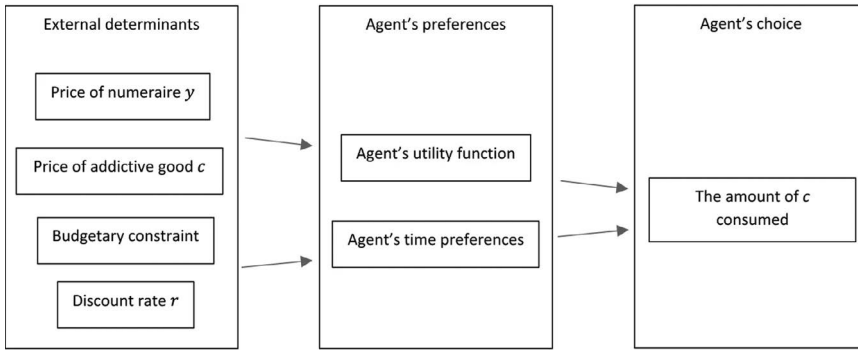

Figure 7.2 The classical rational addiction mechanism complex (Maziarz, 2021).

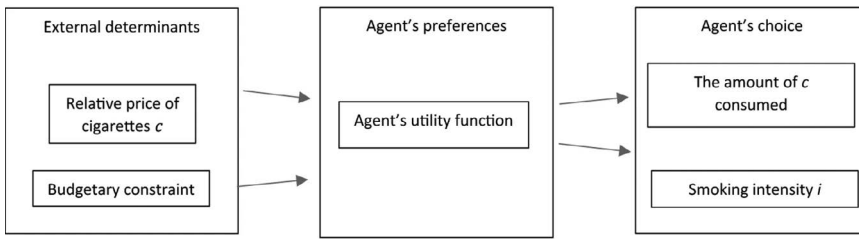

Figure 7.3 The rational addiction mechanism complex (Maziarz, 2021).

addictive good) via the agent's utility function and time preferences (that is, the preference for immediate utility at present over that in the future). See Figure 7.2. The other mechanistic hypothesis (*RAM*2) is that an agent's choice regarding consumption and intensity is caused by external determinants (e.g., the relative price of cigarettes) via her utility function alone (see Figure 7.3). Epidemiological studies (e.g., Patterson et al., 2003; Ashley et al., 2011) show that smokers can modify the intensity of smoking and the amount of nicotine extracted from each cigarette. Evidence from these epidemiological studies undermines *RAM*1 but confirms *RAM*2. Moreover, *RAM*2 is consistent with the study of Adda and Cornaglia (2006): people with lower disposable income are more likely to lower their cigarette consumption and increase smoking intensity. Thus, Maziarz (2021, p. 9973) argues that this evidence of mechanisms 'suggests that smokers choose not only the number of cigarettes smoked but also smoking intensity, and hence the econometric results of Adda and Cornaglia (2006) are additionally supported'. In contrast, Abrevaya and Puzzello's study fails to represent the actual mechanism of the addict's utility maximisation and their correlation claim is 'an artefact produced by the applied statistical techniques' (Maziarz, 2021, p. 9973). It is in this way that Evidential Pluralism can help economists to resolve controversies that arise from association studies.

§26 Mechanisms and Theory in Economics

We addressed the problem of defining 'mechanism' from a general perspective in §21. In this section, we shall focus on the problem as it pertains specifically to economics. 'Mechanism' in economics seems to mean different things in different contexts. Thus, the question arises whether the notion of mechanism encapsulated in Evidential Pluralism captures all the different senses of mechanism in economics.

Reiss (2013) identifies four different senses of mechanism: mechanism as individual causal relation, mechanism as mediating variable, mechanism as underlying structure and mechanism as a piece of theory.

> *Mechanism as individual causal relation*. Econometricians and others who model causal systems as systems of equations sometimes refer to whatever is represented by an individual equation as a "mechanism" . . . Mechanism in this sense contrasts with mere association and does not mean anything beyond causality. For instance, the aggregate relation "money causes prices" could well be described as the "monetary mechanism" in this sense.
>
> *Mechanism as mediating variable*. This notion builds on the idea that, in the social sciences and elsewhere, causes affect outcomes via intermediaries. Military service affects wages via its effect on schooling (among other things). Smoking affects the development of lung cancer via tar deposits in the lungs . . . Importantly, it is not necessary that the "mediating" variable obtains at a lower level than the original cause and effect variables. Mechanisms in this sense can for instance obtain entirely at the social or aggregate level.
>
> *Mechanism as underlying structure or process*. Social or aggregate variables are constituted by entities and processes that lie at a deeper level. A change in money supply, for instance, might be implemented through a variety of instruments, and in a particular case an open-market operation, say, is simply what constitutes the change in money supply on that occasion . . . To provide a mechanism for an aggregate relation, then, is to describe how the entities and processes that underlie the aggregate variables are organized and interact with one another in such a way that the aggregate relation results.
>
> *Mechanism as a piece of theory*. Economists often mean by a mechanism not a thing in the world but rather a theory or part thereof . . . The main difference between this notion and the previous ones, apart from referring to a piece of theory rather than a thing in the world, is that these theoretical mechanisms are often highly idealized descriptions of the interactions among individuals. In the "Hired Gun Mechanism" . . . , for instance, individuals are assumed to be behave perfectly rationally, understand other players' optimal strategies, have stable preferences and so on. While the interactions described certainly obtain, if at all, at the level of individuals and not aggregates, they

> are unlikely to be realistic descriptions of who these individuals are and what they do.
>
> (Reiss, 2013, pp. 104–105)

It is clear that mechanism *as* individual causal relationship is not what we mean by 'mechanism'. Evidential Pluralism talks about the mechanism *responsible for* a causal relationship, and does not conflate the causal relationship with the mechanism responsible for that relationship. Thus, we take talk of the monetary mechanism and the price-specie flow mechanism to refer to the complex-systems mechanisms and mechanistic processes that yield the corresponding causal relationships, not to the relationships themselves.[6] This usage is preferable, because it is only by distinguishing causation and mechanism that one can clearly represent the evidential relationships in causal enquiry (Figure 1.1).

Reiss' two subsequent senses of mechanism in economics, namely, mechanism as mediating variable and mechanism as underlying structure, correspond nicely to the two main senses of mechanism to which Evidential Pluralism appeals: mechanistic process and complex-systems mechanism. In addition, Marchionni (2018) provides an account of mechanism as underlying structure in economics: 'Mechanisms in economics are complexes of rational agents, usually classified into social categories, whose actions and interactions generate causal relationships between aggregate-level variables' (Marchionni, 2018, p. 421). This represents the complex-systems view of mechanisms.

Let us turn to Reiss' fourth sense of mechanism: mechanism as a piece of theory. From the point of view of Evidential Pluralism, a piece of economic theory can encapsulate or confirm a specific mechanism hypothesis. That the theory contains idealisations is no barrier to it describing key features of a mechanism—indeed idealisation is often required precisely in order to isolate the salient features of the mechanism from those that are incidental. For example, the hired gun mechanism is an idealised model that identifies key strategies that feature in mechanisms for ensuring full compliance with certain social actions (Andreoni and Gee, 2011).[7] The life-cycle mechanism (Modigliani and Brumberg, 1954, 1980) is another example of a mechanism hypothesis as a piece of theory. It assumes an idealised scenario of the consumption and saving patterns of individuals in order to identify salient features of the mechanisms responsible for levels of saving.

6 Both the monetary mechanism and the price-specie flow mechanism can be traced back to Hume. The former is the view that prices of goods are caused by the stock of money in relation to the stock of available goods (Hume, 1876a, p. 172), while the latter refers to the idea that a change in prices is caused by redistribution of money across borders (Hume, 1876b, pp. 185–186).

7 The hired gun mechanism appeals to a rule of punishment for non-compliance, according to which only the largest deviation from compliance is punished. Andreoni and Gee (2011, p. 1045) argue that the hired gun mechanism is a low-cost and efficient 'device' that can deter free-riding behaviour and improve public welfare.

Thus Evidential Pluralism can accommodate all four of Reiss' senses of mechanism, albeit with two caveats: a mechanism is responsible for a causal relationship, rather than identifiable with the relationship, and mechanism-as-theory is a mechanism hypothesis, not the underlying mechanism itself.

There is another concern that is related to the role of theory in economics. This stems from the observation that theories often play an important role in causal enquiry in economics (e.g., Cooley and LeRoy, 1985; Cartwright, 2007). If theory provides a distinct source of evidence for causal claims in economics, this poses a problem for Evidential Pluralism, which posits only two sources of evidence, namely evidence of correlation and evidence of mechanisms.

In response, while we agree that theory can play an evidential role in causal enquiry, we maintain that theory can confirm causation only insofar as (i) the theory helps to confirm correlation or mechanism, and (ii) the theory is itself confirmed. This is because (i) establishing correlation and mechanism is precisely what is required to establish causation (§2), and (ii) correlation and mechanism can only be established on the basis of evidence—theories or conjectures do not have confirmatory value on their own.

Consider again the study of unemployment and crime. Burdett et al. (2003, p. 1776) develop a theoretical framework, namely, 'a search equilibrium framework' to analyse the interrelation between crime and unemployment. It is used by Calvó-Armengol and Zenou (2003) to confirm *UCPC*. For Calvó-Armengol and Zenou (2003, p. 184), Burdett et al.'s search equilibrium framework provides 'theoretical motivation' for the causal claim. This is a good example of how a theory provides evidence for a causal claim in economics, but it is compatible with Evidential Pluralism. As we have seen in §24, such theoretical motivation can be understood as providing evidence of mechanisms: Burdett et al.'s search equilibrium framework provides evidence for a specific mechanistic hypothesis proposed by Cantor and Land (1985). Moreover, the authors take this theoretical framework to be well confirmed. This research—as with many other instances of theoretical research in economics—draws conclusions from a set of formal, idealised assumptions. Studies like this provide evidence to the extent that the assumptions do hold and the reasoning from these assumptions is valid. Of course, the authors argue that these assumptions do hold, at least approximately. Insofar as they succeed in establishing that the assumptions hold, the theory can provide evidence of causation. All this can be accounted for by Evidential Pluralism.

Let us consider another example: the study of savings and economic growth. That there is a correlation between savings and economic growth is well established. However, the corresponding causal claim has been widely debated and has become a 'central problem' in macroeconomics (Lewis, 1954, p. 155). The debate has focused on whether savings cause economic growth or vice versa. Modigliani (1970) argues for the causal claim (*GCS*) that greater economic growth rates cause an increase in rates of saving, by appealing to the life-cycle theory of saving. The life-cycle theory of saving (*LCT*) was originally developed by Modigliani and Brumberg (1954, 1980) to provide a theoretical framework of how, why, and how

much people will save. *LCT* assumes that individuals try to maximise the utility deriving from their entire life-cycle consumption (the utility assumption). Thus, a key motivation of savings is the need to provide for retirement. Based on the life-cycle theory and empirical data at the time, Modigliani argues that increases in national rates of saving are caused by increases in economic growth rates.

The basic idea is as follows. Consider an economy in which the population is growing, or in which incomes are growing, so that each generation earns more than their parents. With the growth of the population, there are more young people than older people and more people are saving than are spending their savings. Accordingly, the total expenditure of savings of the old will be less than the total savings of the young, and there will be a net positive saving. If incomes are growing, the young will be saving on a larger scale than the old are spending their savings so that economic growth, like population growth, causes positive saving. Moreover, the faster the growth, the higher the saving rate. Thus, it does not much matter whether it is the population or per capita income that grows; what matters for saving is simply the rate of economic growth.

Modigliani's argument provides a theoretical account to support the causal claim *GCS*. Modigliani's novel contribution is not that the causal claim *GCS* is consistent with various statistical findings. Rather it is that he provides and confirms a mechanism hypothesis based on *LCT*. As Deaton (2010, p. 5) points out, 'Modigliani's insight provided a new mechanism for an old correlation . . . Modigliani's propositions were remarkable for their specificity. It is not simply that growth and saving should be correlated, but the relationship has a specific form, dependent on other observable quantities, and the relationship needs to be consistent with different relationships estimated on both cross-section and time-series data'.

Modigliani's causal claim was very well regarded until the 1980s and 1990s. An objection was put forward by Carroll and Summers (1991): if *LCT* were true, old people in the countries with high economic growth rates would be lifetime poorer than young people. Thus, the consumption of old people would be lower than that of young people. Accordingly, old people in the countries with low economic growth rates would have higher consumption than those in the countries with high economic growth. However, this is not confirmed by empirical studies. For example, in the 1980s, despite the fact that Japan has a higher growth rate than the US, the consumption of old Japanese people was higher than old Americans (Carroll and Summers, 1991, p. 316). The universality of Modigliani's causal claim based on *LCT* was thus undermined.

The rise and fall of Modigliani's causal claim sheds some light on the connection between theory and mechanism hypotheses in the analysis of macroeconomic causes. The direction of the causal arrow in the relationship between savings and economic growth has been widely debated. The initial positive response to Modigliani's causal claim *GCS* was due to the fact that the existing evidence of correlation acted in concert with his theory *LCT*, which provided a credible specific mechanism hypothesis. The causal claim was eventually treated more sceptically, however, on the grounds that his mechanism hypothesis was undermined by new

data. Thus, Modigliani's study of economic growth and saving illustrates our claim that when theory fails to support correlation or mechanism, it cannot confirm a causal claim. An economic theory can provide evidence for a causal claim only if it confirms the relevant correlation or mechanisms, and only if it is itself confirmed by evidence.

§27 Causal and Methodological Pluralism in Economics

We turn now to the problem of causal monism, which we considered in some detail in §20. In this section, we will extend this discussion by considering some views that appeal to economics.

Maziarz (2019) argues that economics is a methodologically pluralist discipline: economists use different methods to establish causal claims. For Maziarz, economists are not only methodological pluralists, but also causal pluralists:

> [E]conomists as a group are conceptual pluralists: they use various methods of causal inferences that allow for formulating causal conclusions understood in line with different notions of this relationship. These different types of evidence support causal claims based on different notions of causality.
>
> (Maziarz, 2020, p. 9)

Maziarz identifies five main concepts of causality in economics: the regularity concept, the probabilistic concept, the counterfactual concept, the mechanistic concept and the interventionist concept. These, according to Maziarz, are presupposed by different groups of economists with different research aims. Maziarz's argument for causal pluralism in economics can be summarised as an inference from methodological pluralism to causal pluralism via a claim about evidential diversity:

> *Methodological.* Economists use different methods to establish causal claims.
>
> *Evidential.* Different methods yield different types of evidence in economics.
>
> *Causal.* Different types of evidence support different concepts of causality in economics.

Indeed, for Maziarz, there are five types of evidence that correspond to his five concepts of cause: 'correlational evidence' provides evidence for regularity causation (Maziarz, 2020, pp. 19, 40), 'probabilistic evidence' provides evidence for probabilistic causation (Maziarz, 2020, pp. 19, 40), 'counterfactual evidence' for counterfactual causation (Maziarz, 2020, p. 8), 'mechanistic evidence' for mechanistic causation (Maziarz, 2020, pp. 8, 108–145) and 'manipulationist evidence' for interventionist causation (Maziarz, 2020, pp. 155, 164, 188).

Maziarz's argument is, in essence, a finer-grained version of Goertz and Mahoney's argument for causal pluralism, which we considered in §20. We argued there that the inference from methodological diversity to causal pluralism

is fallacious and that Evidential Pluralism provides a better explanation of methodological pluralism than does causal pluralism. Exactly the same argument can be used to counter Maziarz's case for causal pluralism. His five types of evidence can be accommodated perfectly well by Evidential Pluralism. Association studies can be thought of as providing correlational and probabilistic evidence, while mechanistic studies clearly provide mechanistic evidence. Both association studies and mechanistic studies can provide evidence pertinent to inferences about counterfactuals and manipulations. There is no need to resort to causal pluralism.

Reiss (2009) argues that there are cases in which methodological pluralism cannot be properly explained by Evidential Pluralism, however. His key example concerns the study of unemployment and inflation. Economists identified a relationship between inflation and unemployment in the 1960s (lower inflation was associated with higher unemployment). However, attempts at controlling unemployment through macroeconomic policy appeared increasingly unsuccessful. Reiss concludes that we cannot always support causal conclusions understood in terms of one sense of cause (e.g., a probabilistic account of causality) from evidence that is relevant under another sense of cause (a manipulationist account). For Reiss, this example tells against Evidential Pluralism and in favour of causal pluralism.

But this conclusion is too hasty. The case of unemployment and inflation does not undermine Evidential Pluralism—at best, it merely shows that policy interventions failed to confirm causation. As Crasnow (2011, p. 44) indicates, '[Reiss] has not established that the failure is due to the need for evidence for a different type of cause. His example of the failure of inference could be read as a situation where the problem is that more evidence is needed, not evidence for a different type of causality'. In order to argue that causal pluralism is true, Reiss needs to show that statistical studies and interventionist studies establish two different types of causal claim. But he does not. As Crasnow (2011, pp. 43–44) observes, '[Reiss'] argument only shows that some causal claims made about the connection between unemployment and inflation were not supported. He does not show that there is an alternative interventionist claim that would support some other conclusion and hence some alternate intervention'.

To conclude, the idea of causal pluralism does not genuinely challenge the application of Evidential Pluralism to economics.

In this chapter, we have defended the application of Evidential Pluralism to economics. Evidential Pluralism validates good examples of causal enquiry in economics, can help us to understand the structure of causal enquiry in economics and can guide research strategies in economics. It can also provide an account of the role of economic theory in causal enquiry. Finally, it can accommodate a variety of research methods without necessitating a move to conceptual pluralism about causation.

8 Political Science

As Kellstedt and Whitten (2018) observe, political scientists routinely need to develop and test theories about the causes and effects of political phenomena. The importance of causal enquiry to political science is witnessed by the fact that 42% of articles published in the *American Journal of Political Science* and 36% in the *American Political Science Review* between 2003 and 2022 mentioned 'cause', 'causal', or 'effect' in the abstracts.[1]

In this chapter, we explore the consequences of Evidential Pluralism for political science. In §28, we argue that Evidential Pluralism explains and justifies the need for methodological diversity in political science. In §29, we illustrate this with a case study of violence and rebellions. In §30, we argue that Evidential Pluralism both avoids causal pluralism and accords well with important research methods in political science.

§28 The Need for Methodological Diversity in Political Science

Political scientists recognise the importance of different research designs and methods in the assessment of causal claims in political science. Research designs include large-*N* cross-case analysis, small-*N* cross-case analysis, and within-case analysis. Large-*N* cross-case analysis uses a large sample. Small-*N* cross-case analysis relies on a small sample—say, two or three cases. Within-case analysis draws on evidence from within a single case.

Gerring (2006) argues for the complementarity of large-*N*, small-*N* and within-case analyses. Gerring distinguishes nomothetic studies from single-outcome studies in political science: nomothetic studies seek general patterns in a large population, while single-outcome studies are concerned with 'what purports to explain only a single case' (Gerring, 2006, p. 707). Examples of single-outcome studies include the study of the causes of World War I, the study of the causes of the terrorist attack on the World Trade Center on 11 September 2001 and the study of the causes of the weak welfare state in the United States. Gerring focuses on single-outcome studies and argues that:

1 These figures were obtained from the Web of Science dataset.

DOI: 10.4324/9781003143000-11

> [A]ll three of these methods [i.e., large-*N* cross-case analysis, small-*N* cross-case analysis, and within-case analysis] ought to be employed, wherever possible. We gain leverage on a causal question by framing the research design in different ways and evaluating the evidence drawn from those separate and independent analyses. To the extent that a particular explanation of an outcome is confirmed by [large-*N* crosscase analysis, small-*N* cross-case analysis, and within-case analysis], one has successfully triangulated.
>
> (Gerring, 2006, p. 726)

Gerring's distinction between nomothetic studies and single-outcome studies can be reframed in terms of the distinction between generic and single-case causal claims. Single-outcome studies can be construed as studies of single-case causal claims, while nomothetic studies can be understood as studies of generic causal claims. Thus, the upshot of Gerring's account is that large-*N* analysis, small-*N* analysis, and within-case analysis are complementary to one another in the assessment of single-case causal claims.

The view that different methods can be complementary in the assessment of causal claims is shared by many political scientists (e.g., George and Bennett, 2005; Lieberman, 2005; Steinberg, 2007; Haggard and Kaufman, 2016). For example, Tilly (1997, p. 48) argues that causal claims are established on the basis of evidence drawn from both 'large-*N* statistical analyses' and 'close reconstructions of particular historical consequences'. Moreover, some political scientists have followed this practice. By referencing the works of Ostrom (1990), Baumgarther and Jones (1993) and Putnam (1993), Steinberg (2007, p. 185) points out that 'scholars who are quite comfortable with quantitative approaches often find that small-*N* research methods are indispensable for producing credible causal explanations'. On the other hand, within-case analysis can benefit from large-*N* studies. Beach and Pedersen (2013), for example, suggest that process tracing, a within-case method, is complementary to large-*N* studies. That said, neither Gerring nor other political scientists fully articulate why different methods are complementary and should be employed in the assessment of causal claims. Gerring simply suggests that the use of large-*N*, small-*N* and within-case analyses contributes to triangulation, but says little on how triangulation is achieved. Steinberg requires establishing a correlation alongside a logic of association:[2]

> [C]redible theories of political behavior and policy processes must not only demonstrate correlations but must establish a logic of association.
>
> (Steinberg, 2007, p. 185)

This coheres with Evidential Pluralism if we understand a 'logic of association' as a mechanism. Nevertheless, it is still not made clear why establishing a correlation and a logic of association is important. Nor is it made clear how

2 Steinberg uses 'association' and 'correlation' interchangeably.

large-*N* analyses, small-*N* analyses and within-case analyses help to establish the existence of a correlation and a logic of association.

Evidential Pluralism helps to fill this explanatory gap: the demand for methodological diversity is justified by the need to establish different objects of evidence, namely, correlation and mechanisms. That there are two objects of evidence in causal enquiry (object pluralism) motivates the need to pay attention to both association studies and mechanistic studies (study pluralism). Moreover, one needs diverse methods, such as large-*N* analysis, small-*N* analysis and within-case analysis to perform such studies. Association studies in political science are typically produced by large-*N* cross-case analysis. As Toshkov (2016, p. 201) points out, the main goal of large-*N* cross-case analysis is to examine 'the pattern of associations between the variables in our set of observations' by focusing on the relationship between one outcome variable and one main explanatory variable. On the other hand, small-*N* cross-case analysis and within-case analysis are widely employed in mechanistic studies. In particular, within-case analysis typically aims to explore a multitude of features of a single case in order to shed light on mechanistic detail (Toshkov, 2016, p. 285). From the point of view of Evidential Pluralism, we ought to use these methods precisely because we typically need both association studies and mechanistic studies. These methods are complementary because association studies and mechanistic studies play complementary, mutually reinforcing roles.

One might think that Evidential Pluralism is not a new idea in political science. For example, Lieberman (2005) argued for a 'combined approach': large-*N* analysis and small-*N* analysis are complementary to each other in the sense that small-*N* analysis provides additional evidence for causal inferences. Like Evidential Pluralism, Lieberman highlights the significance of evidential diversity in causal analysis. However, there is a crucial difference between Lieberman's combined approach and Evidential Pluralism. Lieberman's 'evidential diversity' refers to the diversity of the types of evidence-gathering methods, while the evidential diversity required by Evidential Pluralism is a diversity of objects of evidence and a diversity of studies. Thus, Lieberman's 'evidential diversity' is basically a call for methodological diversity, whereas Evidential Pluralism requires object pluralism and study pluralism, which are different to methodological diversity. However, we saw earlier that Evidential Pluralism can explain and justify the need for methodological diversity.

In the next section, we will present an example to show how different methods and different objects of evidence are used by political scientists to assess causal claims and how this approach can be explained and justified by Evidential Pluralism.

§29 Case Study: Resource Wealth and Violence in Rebellions

Violence against civilians is one of the fundamental characteristics of civil wars. Between 1945 and 2007, over 16 million people died in civil wars—five times more than those who died in interstate wars. Rebel groups often share responsibility for

violence against non-combatants with governments in civil wars. However, variations in violence have posed a persistent puzzle for political scientists. Different rebel groups in civil wars have different strategies of violence: some rebel groups abuse non-combatant populations, while others exhibit discipline and restraint. Some kill the victims selectively, while others do so indiscriminately. Why do patterns of insurgent violence vary so much?

Weinstein (2007) suggests that the level of resources available to a rebel group is a cause of the kind of violence it exhibits:

> [R]ebel groups that emerge in environments rich in natural resources or with the external support of an outside patron tend to commit high levels of indiscriminate violence; movements that arise in resource-poor contexts perpetrate far fewer abuses and employ violence selectively and strategically.
>
> (Weinstein, 2007, p. 7)

Weinstein hypothesises a mechanism linking a group's initial endowments to its strategies of violence. The initial endowments (*I*) that rebel leaders have access to constrain their tactics of recruitment (*R*), which shape the membership profile of a rebel group (*M*). The membership profile of a rebel group affects its internal organisation (*O*), which in turn affects the governance of the group (*G*) and ultimately its strategies of violence in war (*V*). Let us call this mechanism, depicted in Figure 8.1, the '*IRMOGV*' mechanism. Rebel groups with access to substantial economic resources are more likely to be able to use these endowments to provide incentives in order to motivate individuals to join the rebellion, while those

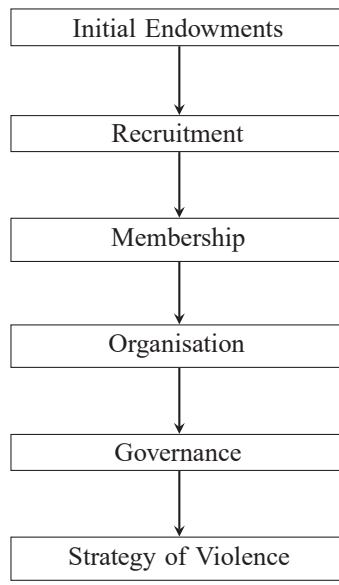

Figure 8.1 The *IRMOGV* mechanism.

groups with limited economic endowments more often attract individuals by promising them future and collective benefits from a rebel victory and reinforce these promises by means of ethnic, religious or ideological ties. Where participation is risky and short-term gains are unlikely, rebel groups tend to attract highly committed individuals. On the other hand, where participation involves lower risk and leads to immediate rewards, groups tend to attract less committed individuals. Weinstein calls the former type of rebellion 'activist rebellions' and the latter 'opportunistic rebellions'. Leaders of activist rebellions tend to use violence strategically, as they have clearer guidelines about how combatants should behave and strong mechanisms for enforcing discipline. Thus, activist insurgencies are able to selectively identify targets, implement attacks, and discipline the use of force. In contrast, opportunistic rebellions find it difficult to identify potential defectors and are prone to making mistakes. A constant demand for short-term rewards drives combatants to loot, destroy property and attack indiscriminately.

Drawing on interviews with about two hundred combatants and civilians, Weinstein uses a small-*N* comparative method to show that the *IRMOGV* mechanism is confirmed by four rebellions—one each in Mozambique and Uganda, and two in Peru. Here we consider two examples: the case of Renamo in Mozambique and that of the NRA in Uganda.

Renamo. The rebellion in Mozambique led by the Resistencia Nacional Mozambicana (Renamo) provides one example.

($I_R \rightarrow R_R$) Renamo originated in 1976 as a radio station in Mozambique, Voz da Africa Livre (Voice of a Free Africa). It was quickly transformed into a fully fledged rebel movement, supported by the Rhodesian Central Intelligence Organisation (CIO) and the Rhodesian army. Starting with a few hundred members, Renamo grew quickly in the first few years. In 1979, it had nearly two thousand members. New recruits were provided with weapons, uniforms, food, comfortable shelters and money. For example, the soldiers of the infantry were rewarded with salaries of RHD 20 a month.

($R_R \rightarrow M_R$) A consequence of these recruitment strategies was that early joiners had few social or ethnic ties with one another. Most of them had no education beyond primary school. The leadership structure also encompassed a broad range of ethnic groups. However, Renamo's strategies of recruitment changed after 1979 when the Rhodesian government collapsed and Renamo was left without an important source of economic and logistical support. Although the South African government quickly stepped forward to fill the gap by providing logistical support, military supplies and related training, Renamo was still struggling financially. Renamo began recruiting heavily by force: children, young men and refugees were taken, or even abducted, to Renamo's bases. It is estimated that about 90% of its members were recruited by force (Minter, 1989). This new strategy of recruitment reinforced the lack of coherent social ties within Renamo. A striking feature of the membership profile of Renamo is that though Renamo had more than 20,000 soldiers under arms by coercive recruitment, the vast majority of the soldiers had little educational background (see Table 8.1).

Table 8.1 The educational profile of demobilised soldiers in Mozambique's civil war (Weinstein, 2007, p. 115).

Level of Education	*Number of demobilised soldiers*
None	26,434
Primary school	25,381
Middle school	23,754
Secondary school	2,197
University	271

($M_R \rightarrow O_R$) Renamo offered its members little political education to 'set in place shared beliefs about the purpose of the war and the way in which it should be conducted' (Weinstein, 2007, p. 145). In contrast, its members were given extensive military training exercises with a focus on weapons handling. This military training shaped a belief among Renamo recruits: they were soldiers of an army rather than members of a political movement. Because abduction was often used for recruitment, Renamo's leaders saw their members as expendable. As a consequence, Renamo's combatants usually avoided direct confrontation with government armies. Instead, they attacked non-combatant populations to gain wealth and power. As Weinstein (2007, p. 147) points out, there was no clear code of conduct in Renamo. Nor was there any regulation of combatant-civilian interaction. What is worse, Renamo's organisational structure reinforced its members' impression that the movement lacked a long-term aim. Thus, Renamo soldiers focused on enriching themselves by attacking civilian populations while trying to avoid direct conflict with government forces. In short, Renamo ended up with a situation in which both commanders and combatants aimed at immediate rewards rather than a long-term political objective. Moreover, little effort was made to counter this internal culture.

($O_R \rightarrow G_R$) Regarding its way of governing, Renamo relied on traditional structures by returning power to the colonial and precolonial *regulos* (traditional local leaders). Although Renamo allegedly aimed at achieving multiparty democracy, fair elections and freedoms for civilians, it behaved differently at the local level. For example, Renamo did not allow shared military-civilian governance of the movement. Its policy was entirely generated within its military organisation. Basically, the governance of civilians for Renamo was a military task. Due to its unilateral military control and coercive governance, Renamo arrived at a system for regularising the collection of food: often, armed Renamo soldiers showed up at people's doors asking for food.

($G_R \rightarrow V_R$) As a result, Renamo used violence without restraint. Victims even included supporters of Renamo and non-combatant civilians. Renamo abused noncombatants extensively by abducting recruits and civilian supporters, destroying buildings in government-controlled areas, looting civilian property and killing civilians. Aggregate data shows that Renamo was responsible for most incidents of violence against civilians (Weinstein, 2007, p. 231). For example, Renamo committed 112 massacres (i.e., the killing of more than 15 non-combatants) in

the period from 1980–1992. This pattern of violence employed by Renamo was consistent both geographically and temporally during the civil war.

NRA. The National Resistance Army (NRA) in Uganda provides a contrasting case.

($I_N \rightarrow R_N$) The NRA emerged with little external financial support. The early development of the NRA mainly relied on social endowments. In its early years, the members of the NRA had a similar social background. For example, newly recruited soldiers were typically educated university students. Many members had similar political experiences: some had worked with Yoweri Museveni (the leader of the NRA) in the 1980 anti-government movement, while others had fought with Museveni in his guerrilla FRONASA in the 1970s. The majority of the members in the early years were ethnically Banyankole from western Uganda, and many were born-again Christians.

($R_N \rightarrow M_N$) The NRA's reliance on social endowments rather than economic endowments led to a selective and disciplined strategy of recruitment when it expanded its influence to the Triangle-Baganda (the central part of Uganda). It tried to tie civilian supporters to the NRA by means of political allegiances. A political agenda that linked the historical experiences of Baganda to those of the Bayankole was articulated and highlighted. Typically, the leaders of the NRA mobilised supporters by explaining their political motivation, emphasising the desire to defeat the corrupt government (the Obote regime) and promising to end ethnic tribalism in Uganda. They also made promises about the future (especially the prospects of the new government, the benefits and compensation). A rigorous military and political training was instituted for the new recruits. Lectures on the history of Uganda, political economy, and strategies of guerrilla warfare were provided. Reading the works of Mao and Fanon was encouraged.

($M_N \rightarrow O_N$) The NRA's political training not only helped its members to share political beliefs but also shaped the expectations and behaviours of the new recruits. For example, NRA members were expected to 'help and support the local population' (Weinstein, 2007, p. 141). The NRA also built a system of political commissars into the army structure, in which combatants who were rooted into the position of commissar were responsible for explaining the NRA's political aims and goals to the civilian population. In addition, the NRA promoted the idea of comradeship, which helped to create an environment of sustained cooperation. Moreover, the NRA introduced formal guidelines as well as punishment mechanisms. For example, its code of conduct, first circulated in 1981, detailed the norms of interaction about how combatants should treat other combatants, their commanders and the civilian populations. It explicitly prohibited abuses of civilians and clarified the organisational structure of the army. It also articulated the punishments for various types of undisciplined behaviour. Membership and promotion in the NRA depended entirely on merit and performance. A democratisation of the NRA and a decentralisation of power were also introduced by Museveni himself. All these efforts, as Weinstein (2007, p. 145) puts it, 'contributed to the NRA's reputation for cooperation, cohesion, and trust'.

($O_N \rightarrow G_N$) For Museveni, the NRA's dependence on civilian support was closely connected with its need to provide security in order to maintain sustained cooperation with non-combatant civilians. In order to reinforce the expectations of mutual exchange between the NRA and the civilian population, a formal democratic local government that shared power with non-combatants was instituted. These new forms of local government changed social hierarchies and thus transformed the traditional relationships between different classes. They broadened civilian participation in governance and resulted in a power-sharing administration that involved both the NRA and the civilian population. As Weinstein (2007, p. 180) points out, 'Fundamentally, the NRA maintained civilian support by providing public goods [i.e., security and health care] in exchange for civilian contributions'.

($G_N \rightarrow V_N$) In the civil war, the behaviour of the NRA combatants was restrained and disciplined. It was responsible for relatively few incidents of violence. Indeed, the NRA committed only 17 percent of the total incidents of violence against civilians recorded in the war (see Figure 8.2). This can be regarded as a consequence of the NRA's disciplined organisation and democratic governance. It had a clear policy for dealing with its opponents: the focus was on the members of the armed forces and the police. It also prohibited behaviours like looting and the destruction of civilians' personal property and punished acts of indiscipline. What is more, the NRA sought to behave with consistency wherever it operated.

Weinstein argues that the cases of the Renamo and NRA provide strong evidence for the *IRMOGV* mechanism hypothesis. In addition, Weinstein uses data

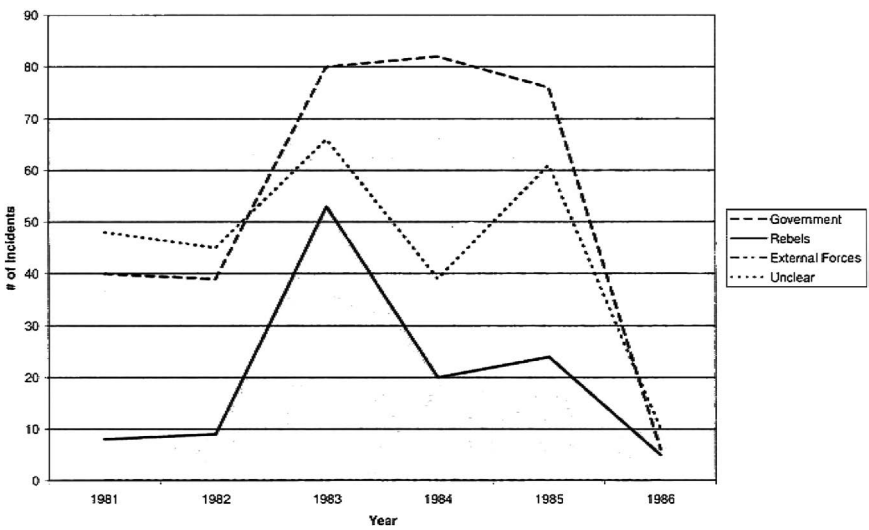

Figure 8.2 Incidents of violence against civilians in Uganda's civil war (Weinstein, 2007, p. 221).

on all civil wars that occurred between 1945 and 2000 to confirm the claim that the initial endowments of rebel groups are causes of strategies of violence (see Figure 8.3). Ordinary least squares (OLS) regression analysis shows that the use of contraband resources to finance an insurgency is associated with the higher levels of violence, controlling for the duration of the war. External support for rebellion also significantly increases the level of civil war violence.

Evidential Pluralism explains and justifies Weinstein's use of diverse evidence. Weinstein in fact looks for two objects of evidence—evidence of mechanisms and evidence of correlation—by employing different methods. By conducting small-*N* analysis of the cases of rebellions in Uganda, Mozambique and Peru, based on qualitative data from interviews, Weinstein confirms a specific mechanistic hypothesis, namely, the *IRMOGV* mechanism hypothesis. By estimating the statistical relationships between key variables including combat-related deaths and measures of the level of violence against non-combatants in civil war (using large-*N* analysis), Weinstein obtains evidence of a correlation between the initial endowments and the strategies of violence. Thus these different methods are used in order to look for evidence of correlation and evidence of mechanisms for the purpose of confirming the causal claim that the initial conditions facing the leaders of rebel groups are causes of the strategies of violence. Weinstein's research on wealth resources and violence accords well with the evidential relationships posited by Evidential Pluralism.

	Dependent Variable: Log of Combat-Related Deaths (Excluding Mass Killing)			
Duration	0.123	0.045	0.120	0.056
	[3.60]***	[2.17]**	[3.20]***	[3.33]***
Contraband	0.961	1.249		
	[2.29]**	[3.19]***		
Pro-Rebel intervention			0.820	1.226
			[1.75]*	[3.60]***
Constant	8.057	8.260	7.913	8.028
	[25.77]***	[31.30]***	[22.07]***	[28.21]***
Observations	42	78	42	78
R-squared	0.33	0.16	0.35	0.20
Sample	Wars fought over state control	All civil wars	Wars fought over state control	All civil wars

Notes: OLS estimates. The dependent variable is combat-related deaths, excluding mass killing. T-statistics (calculated with robust standard errors) in parentheses. *** Significant at 0.01 level; ** Significant at 0.05 level; * Significant at 0.10 level.

Figure 8.3 Resource wealth and violence in civil wars, 1945–2000 (Weinstein, 2007, p. 307).

§30 Understanding Causal Enquiry in Political Science

Not only does Evidential Pluralism explain successful instances of causal enquiry in political science, but it can also help to improve our general understanding of causal enquiry in political science, as we argue in this section. Firstly, in §30.1, we provide an example to undermine the claim that political science research requires causal pluralism—the view that there are multiple concepts of cause in use in political science. In §30.2 and §30.3, we show how two key methods in political science, process tracing and the use of multi-method large-*N* qualitative studies, fit very naturally with Evidential Pluralism.

§30.1. Evidential Pluralism versus Causal Pluralism in Political Science

Causal pluralism is currently popular in political science. Crasnow (2019, p. 41), for example, suggests that 'conceptual pluralism about causality' is required to promote methodological diversity in political science. In §20, we argued that Evidential Pluralism provides a simpler and more unified account of methodological diversity in the social sciences than does causal pluralism. Here we appeal to Htun and Weldon's study on violence against women as a means to illustrate the way in which Evidential Pluralism obviates the need to resort to causal pluralism.

Htun and Weldon (2012) seek to establish a causal relationship between strong, autonomous women's movements and government adoption of social policies to address violence against women. Their research is multi-method and takes place in several stages, beginning with the collection of both qualitative and quantitative data from a variety of sources in 70 countries over 40 years. They find a clear association between the dependent variable (government adoption of social policies) and the independent variable (presence of strong, autonomous women's movements) consistent with the hypothesis that 'strong, autonomous feminist movements will be significant influences on policies on violence against women at all points in time' (Htun and Weldon, 2012, p. 554). They also obtain 'case evidence' (in our terminology, evidence of mechanisms) by using process tracing in order to unpack the links between events and determine their causal order (Htun and Weldon, 2012, p. 560). It is evident that Htun and Weldon take themselves to be working with a single concept of cause: they do not view their different methods to be targeting different kinds of causal claims. In their own description of their approach, they identify their quantitative analysis as providing evidence of correlation but not evidence of the hypothesised mechanism. They provide additional evidence that supports their mechanism hypothesis that strong, autonomous feminist movements affect policies on violence against women through the institutionalisation of feminist ideas in international norms. Thus, Htun and Weldon's methodological pluralism cannot be easily accounted for by causal pluralism. Indeed, it provides a good example of Evidential Pluralism: different methods are used to obtain evidence of correlation and evidence of mechanisms, in order to confirm a single causal claim. This sort of example favours Evidential Pluralism over causal pluralism in political science.

§30.2. The Role of Process Tracing

Another consideration that favours Evidential Pluralism is its ability to elucidate the roles of various methods in political science. As we argued in §28, large-*N* analyses are typically employed in association studies, while small-*N* analysis and process tracing feature in mechanistic studies. Each method plays an important role in confirming causal claims in political science. In this section, we shall explore in some detail how Evidential Pluralism sheds light on the role of process tracing in political science research.

Process tracing is typically construed as a qualitative within-case method, in that it focuses on an individual case rather than multiple cases. It is commonly held that there are three main variants of process tracing: theory-testing process tracing, theory-building process tracing and explaining-outcome process tracing (Beach and Pedersen, 2013).[3]

Theory-testing process tracing typically begins with hypothesising a mechanism linking putative cause *A* to putative effect *B*, including the operations of the mechanism, and ends with evidence collection to test this specific mechanism hypothesis. From the point of view of Evidential Pluralism, theory-testing process tracing is a method that mechanistic studies in political science can employ. Where accompanied by studies of the correlation between *A* and *B*, this kind of process tracing can be very helpful to causal enquiry.

Theory-building process tracing starts with evidence collection to determine observable events (or in Beach and Pedersen's term, 'observable manifestations'), and then infers a mechanism from these observable events. Beach and Pedersen (2013) identify two research situations in which theory-building process tracing is utilised:

> (1) when we know that a correlation exists between *X* and *Y* but we are in the dark regarding potential mechanisms linking the two (*X*-*Y*-centric theory building) as we have no theory to guide us: or (2) when we know an outcome (*Y*) but are unsure about the causes (*Y*-centric theory building). In the second instance, the analysis first traces backward from *Y* to undercover a plausible *X*, turning the study into an *X*-*Y*-centric analysis.
>
> (Beach and Pedersen, 2013, p. 16)

As Beach and Pedersen suggest, in the first, *X*-*Y*-centric situation, one can only begin tracing the process from *X* to *Y* if there is an established correlation between *X* and *Y*. In the second, *Y*-centric situation, one needs to turn a *Y*-centric analysis into an *X*-*Y*-centric analysis, which assumes a correlation between *X* and *Y*. These two scenarios accord perfectly with Evidential Pluralism: theory-building process

3 Beach and Pedersen (2019) add another variant of process tracing, namely, theory-revision process tracing. We will not discuss this variant here, as it does not directly concern the process of establishing or assessing a causal claim.

tracing is used to provide evidence that is complementary to evidence of correlation, as correlation has already been established.[4]

Explaining-outcome process tracing is distinct from theory-testing and theory-building process tracing because it seeks to explain a particular historical outcome (Beach and Pedersen, 2013, p. 63). There are two kinds of explanation that might be sought. One is what might be called a 'narrative explanation', which links a sequence of events by means of a narrative that can be thought of as elucidating a mechanistic process that connects these events. The second kind of explanation is causal explanation: here one needs not only the mechanism that links the start of the chain of events to the last, but also a correlation, so that one can infer that the first event made a difference to the last event. Consider, for example, explanations of the start of World War 1 that appeal to the assassination of Archduke Ferdinand. A narrative explanation would link the assassination to the outbreak of war by a chain of events, each of which led to the next. A causal explanation would go further by asserting that the assassination was a cause of the war: this would require evidence to support the claim that the assassination significantly increased the probability of the war, i.e., if there had been no assassination, the chance of war would have been significantly lower.[5]

Insofar as explaining-outcome process tracing is used to provide narrative explanation, it falls outside the remit of Evidential Pluralism, which focuses on establishing and assessing causation. For example, Wood (2003) is cited by Beach and Pedersen (2013, p. 63) as an example of explaining-outcome process-tracing. Wood provides an explanation of insurgent collective action in El Salvador by identifying 'three reasons that participants supported the mobilisation and insurgency' (Wood, 2003, p. 231). This study stops short of asserting causal claims. Such cases pose no challenge to Evidential Pluralism because they are not examples of causal enquiry.

On the other hand, where explaining-outcome process tracing is used for the purposes of causal explanation, Evidential Pluralism does apply, and evidence of correlation needs to be provided alongside the evidence of mechanism for the following reason. As Beach and Pedersen (2013, p. 18) observe, where explaining-outcome process-tracing studies seek causes, they aim 'to craft a minimally sufficient explanation of a particular outcome, with sufficiency defined as an explanation that accounts for all of the important aspects of an outcome with no redundant parts being present'. In order to establish X as a minimally sufficient explanation of Y by process tracing, there needs to be a correlation between X and Y, for otherwise X would give no reason to expect Y.

4 Bennett (2010, pp. 208–209) also observes that process tracing is mainly used to provide further evidence for causation, having already obtained evidence of correlation.

5 Although a narrative explanation can often be represented as a chain of causal relationships, causation is not necessarily transitive, i.e., the first event in the chain is not necessarily a cause of the last, because it need not raise the probability of the last. A narrative explanation only becomes a causal explanation where there is transitivity—where the event invoked as the explainer is correlated as well as mechanistically connected to the event to be explained.

Here the causal claim is single-case (i.e., X and Y are single-case outcomes), so the correlation required is a single-case correlation: the chance of Y given X differs from the chance of Y in the absence of X, conditional on potential confounders. Evidence for this single-case correlation may take the form of association studies that support a generic correlation in a population to which this particular case belongs. But the evidence might also include mechanistic studies that elucidate the mechanisms in operation in the presence of X and those in operation in the absence of X. A comparison of these two cases can confirm correlation via routes μ_1 and μ_3 of Figure 1.1. In the case of the causes of World War 1, the relevant evidence is unlikely to come from association studies: there is no large statistical study that finds an association between assassinations of archdukes and subsequent wars. Rather, evidence of (single-case) correlation is likely to be provided by careful counterfactual reasoning from established mechanisms together with detailed facts about the particular context.[6] Note that Evidential Pluralism does not require that all this evidence be new—the correlation may have been established previously, for example.

Skocpol's study on social revolutions includes examples of this kind of process tracing (Goldstone, 1997; George and Bennett, 2005; Mahoney, 2012).[7] Skocpol (1979) argues that international pressure and peasant rebellion are two causes of social revolutions in France, Russia and China. She uses process tracing to show how each of the two independent variables (i.e., international pressure and peasant rebellion) set into motion a complex sequence of events that culminate in revolutionary social transformation in each country. Overall, as George and Bennett (2005, p. 227) argue, Skocpol's use of process tracing shows 'how these two variables were causally related to the revolutionary social transformation in

6 This sort of counterfactual analysis is widely used to establish a correlation between historical events (e.g., Fearon, 1991; Mahoney and Barrenechea, 2019). For example, Stepan (1978) used a careful counterfactual analysis to study the 1964 military takeover in Brazil. Another well-known example is Frank Harvey's study of the 2003 invasion of Iraq (Harvey, 2011). Gillies (2019, §10.2) provides an example of this kind of counterfactual reasoning in the context of medicine. Nolan (2013) offers a general discussion of the use of counterfactual reasoning in historical explanation.

Other methods used by social scientists include time-varying parameter models (Issac and Griffin, 1989; Griffin and Isaac, 1992) and event-structure analysis (Corsaro and Heise, 1990; Griffin, 1993). For example, Isaac et al. (1994) employ time-varying parameter models and event-structure analysis to establish the single-case causal claim that the death of Martin Luther King Jr led to the expansion of race-based poor relief at the expense of more progressive programmes of class-based economic reform.

7 Skocpol can be thought of as providing examples of explaining-outcome process tracing for the following reason. Beach and Pedersen (2013, p. 19) construe explaining-outcome process tracing as an iterative research strategy that aims to trace 'case-specific mechanisms' that 'cannot be detached from the particular case'. However, 'explaining-outcome studies often have theoretical ambitions that reach beyond the single case' (Beach and Pedersen, 2013, p. 19). These descriptions fit Skocpol's work well, as she identifies three case-specific mechanisms and connects them to two causal variables. Thus her study has ambitions to connect the three cases, but while she does make single-case causal claims, she stops short of making any generic causal claim.

each of these countries'. Skocpol does this by providing evidence of correlation in addition to evidence of mechanism: by unpacking the detail of the mechanisms and carrying out a contrastive analysis, Skocpol argues that international pressure and peasant rebellion are associated with outbreaks of social revolutions. Skocpol's contrastive analysis involves comparing the abortive Russian Revolution of 1905 with the successful Russian Revolution of 1917 and using some aspects of English, Japanese and German history as contrasts to those of French, Russian and Chinese history.[8]

In sum, then, Evidential Pluralism suggests that the primary role of process tracing is its use in mechanistic studies. These mechanistic studies can provide evidence for specific mechanism hypotheses that describe the process that is traced. Thus, process tracing studies clearly provide evidence of mechanisms. However, evidence of mechanisms needs to be complemented by evidence of correlation for the purposes of causal enquiry: process tracing can play a role in providing this evidence of correlation via techniques such as counterfactual analysis.

§30.3. Multi-method Large-N Qualitative Analysis

Evidential Pluralism also accords well with multi-method large-*N* qualitative analysis (LNQA), which is a new method in political science. Multi-method LNQA builds upon LNQA, which is depicted in Figure 8.4. LNQA seeks 'to multiply the number of qualitative case studies in order to strengthen causal inference' (Goertz and Haggard, 2023, p. 282):

> The approach starts with a theory and hypotheses along with a proposed causal mechanism concerning the relationship between *X* and *Y*. The stipulation of a causal mechanism proves important because it structures the within-case analysis. The next step establishes the scope conditions of the claim by defining a relevant population. Once the empirical population is defined, LNQA studies report *X* or *Y* regularities. These regularities take the form of "if $X = 1$ then $Y = 1$," an *X* regularity, or "if $Y = 1$ then $X = 1$," a *Y* regularity. A regularity is simply the share of cases that appear to conform with the generalization.
>
> (Goertz and Haggard, 2022, p. 4)

Multi-method LNQA is similar to LNQA, but adds a statistical component. Typically, it begins with a significant average treatment effect (ATE) of *X* on *Y*, ascertained by quantitative methods from an association study. Then standard LNQA is applied:

8 A question arises as to whether this contrastive analysis, which can be thought of as an instance of Mill's method of difference, is a process-tracing technique or something that goes beyond process tracing (Goldstone, 1997; George and Bennett, 2005). Either way, Skocpol's study clearly provides both evidence of correlation and evidence of mechanisms for her within-case causal inferences.

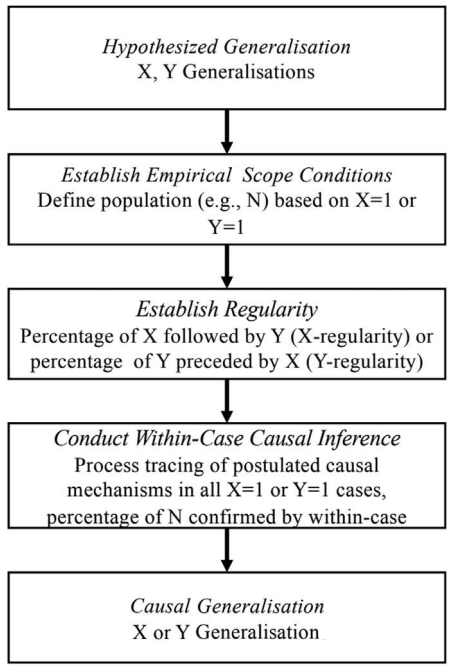

Figure 8.4 LNQA (Goertz and Haggard, 2022, p. 5).

> Following standard qualitative methods approaches, however, multi-method LNQA typically focuses on cases in which *X* and *Y* are both present, which we call the (1,1) cases. It is these cases which are expected to exhibit the presence and operation of the causal mechanism. If within-case causal inference justifies the causal relationship in the (1,1) cases then LNQA strengthens confidence that the ATE is causal.
>
> (Goertz and Haggard, 2022, p. 4)

Thus, multi-method LNQA is an integration of LNQA and statistical analysis (see Figure 8.5). Goertz and Haggard (2022, p. 7) argue that the procedure rests on 'the presumption that a diversity of very different types of evidence—for example, statistical and with respect to mechanisms—can nonetheless be exploited to increase our confidence in causal claims'. In particular, they maintain that multi-method LNQA is 'a commitment to the belief that the confidence in statistical inference can be strengthened through within-case causal inference' (Goertz and Haggard, 2022, p. 8). These views clearly accord well with Evidential Pluralism.

Recently, however, this concordance between multi-method LNQA and Evidential Pluralism has been challenged by Runhardt (2022). Runhardt's argument can be sketched as follows:

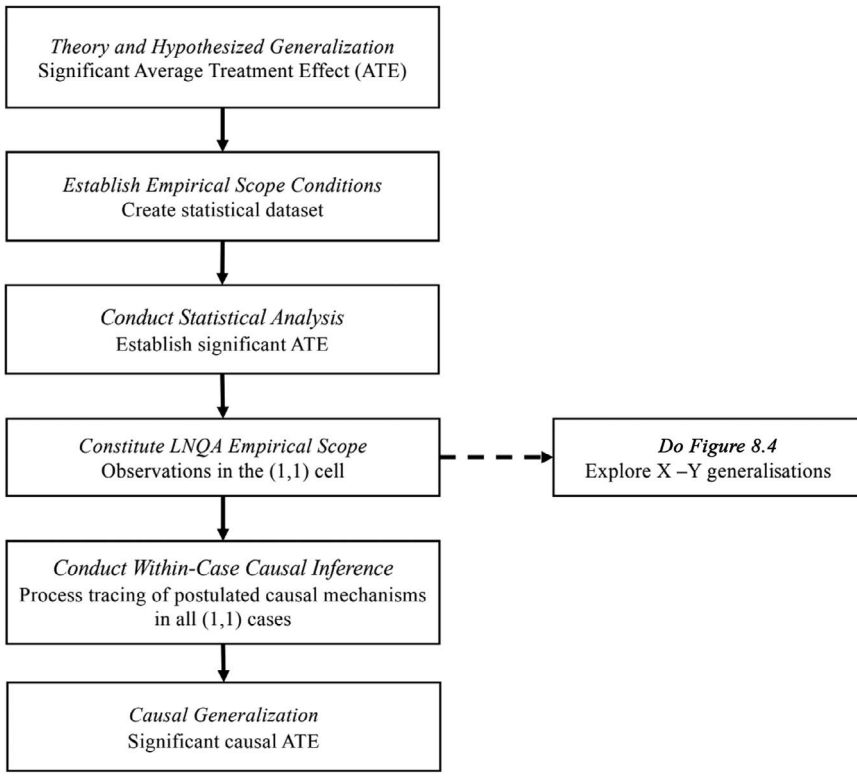

Figure 8.5 Multi-method LNQA (Goertz and Haggard, 2022, p. 6).

LNQA1. There are two types of generic causal claim: 'mechanistic generality' and 'average treatment effect'.

LNQA2. What multi-method LNQA researchers really aim at are mechanistic generality claims, rather than average treatment effect claims.

LNQA3. Mechanistic generality claims can be established by LNQA alone.

LNQA. The statistical step in multi-method LNQA is redundant and the approach conflicts with Evidential Pluralism.

Runhardt's argument rests on a distinction between mechanistic generality and average treatment effect, and on viewing both of these as kinds of causal claims. Runhardt follows Hausman (2010) in defining an ATE claim as follows: '*X* is a cause of *Y* iff, when one holds fixed the frequencies of all the other background factors relevant to *Y* (apart from *X* and its effects) at their frequency in population *P*, there is a significant difference in average outcomes *Y* between cases where

X =0 and cases where X = 1' (Runhardt, 2022, p. 16). On the other hand, for Runhardt, a mechanistic generality claim is the claim that 'the same causal mechanism is behind the causal relation between X and Y in all the cases in the population where X = 1 and Y = 1' (Runhardt, 2022, p. 16). We are sympathetic to the distinction between mechanistic generality claims and ATE claims. That said, we do not agree that mechanistic generality and ATE claims are two types of generic causal claim. Rather, it is more appropriate to call them generic mechanistic claims and generic correlation claims respectively. In particular, to identify causation with an ATE is to conflate causation with a kind of correlation.

Given this observation, Runhardt's argument can be reformulated as follows.

*LNQA1**. There are two types of generic claim: 'mechanistic generality' and 'average treatment effect'.

LNQA2. What multi-method LNQA researchers really aim at are mechanistic generality claims, rather than average treatment effect claims.

LNQA3. Mechanistic generality claims can be established by LNQA alone.

LNQA. The statistical step in multi-method LNQA is redundant and the approach conflicts with Evidential Pluralism.

If *LNQA*2 is true, only the first part of the conclusion follows. If average treatment effect claims do not matter for multi-method LNQA researchers, the statistical step in multi-method LNQA is redundant. But this does not undermine Evidential Pluralism, because Evidential Pluralism concerns causation, not mechanistic generality. Thus Runhardt provides an argument against multi-method LNQA rather than Evidential Pluralism.

On the other hand, if *LNQA*2 is not true, neither multi-method LNQA nor Evidential Pluralism is undermined.

Either way, then, Evidential Pluralism remains intact. But we would go further, by rejecting *LNQA*2. As Goertz and Haggard (2022, p. 1) note, multi-method LNQA aims at establishing causal claims, rather than merely mechanistic generality claims or ATE claims: 'Multi-method LNQA . . . systematically uses within-case causal inference to confirm that the statistically significant treatment effects are causal'. Thus, *LNQA*2 is false. We conclude that multi-method LNQA does accord well with Evidential Pluralism.

In this chapter, we have argued that Evidential Pluralism can account for methodological diversity in causal enquiry in political science, that it undermines an inference from methodological diversity to causal pluralism, and that it validates two important methods in political science, namely process tracing and multi-method LNQA. We conclude that Evidential Pluralism accords well with the practice of causal enquiry in political science.

9 Law

Causation is a key concept in the law. Legal theorists and lawyers are concerned with the causes of harms in order to assess liability. Judges, meanwhile, often need to explain to a jury what is required to establish causation in the context of an offence before the court. This chapter examines the application of Evidential Pluralism to the law. In §31, we set out and examine what we call the 'bifurcation' approach to causation in the law: this identifies two conditions for establishing causation. In §32 and §33, we show that there are two ways in which Evidential Pluralism can fruitfully be applied to causal enquiry in the law. We develop a new concept, namely that of *liability-tracing mechanism*, in §33 to facilitate this application. Finally, in §34 we argue that the concept of causation in the law should not be treated as entirely autonomous from the concept of causation as studied in philosophy.

§31 The Bifurcation Approach to Causation in the Law

There is a widely held view that the law is, or should be, autonomous from philosophy in its use of the concept of causation:

> The lawyer cannot afford to adventure himself with philosophers in the logical and metaphysical controversies that beset the idea of cause. (Pollack, 1901, p. 36)

> The word "cause" has in philosophy given rise to embarrassments which in this connection should not affect the judge. (Caswell v Powell Duffryn Associated Collieries Ltd, 1940, p. 164)

> Each science—natural, social and humanistic—and philosophy have their own meaning of causation. The proper meaning of causation in law is also sui generis. (Ryu, 1958, p. 508)

> When the lawyer uses the concept of causation, he is not bound to use it in the same way as a philosopher, or a scientist, or an ordinary man. (Williams, 1961, pp. 75–76)

> Traditionally, lawyers disdained philosophical enquiries into "causation" as being too abstract or vague. (Stapleton, 2008, p. 447)

DOI: 10.4324/9781003143000-12

> [M]ost lawyers follow Hume in thoroughly disparaging the project of marrying law and metaphysics. (Stapleton, 2015, p. 702)

Many legal theorists maintain that philosophers are primarily interested in the metaphysical nature of causation. As Stapleton (2015, p. 702) argues, 'debates as to what is metaphysically relevant about the existence of a phenomenon do not need to trouble lawyers'. In contrast, causation in the law is thought of as a conceptual tool to assess liability. As Hart and Honoré put it,

> Causal questions . . . appear in every branch of the law and there is a variety of ways, even in a single branch, in which legal rules make causal connection an element in responsibility. . . . The most frequent type of causal questions which courts face is whether a human action or omission caused some specific harm but even this one form of question may be relevant to legal responsibility in different ways. In criminal law this question usually has to be answered because criminal offences are often defined in simple terms as acts causing specific harms: in such cases a causal connection between some action of the accused and the specified harm must be shown in order to establish the existence of liability, i.e. that a particular offence has been committed and that the accused is liable to punishment. . . . In other types of case it is necessary to determine the extent of liability; and in others still it will be necessary, in order to determine both the existence and the extent of liability. (Hart and Honoré, 1985, p. 84)

A standard approach to causation in the law is what we shall call the bifurcation approach.[1] As Witjens (2014, p. 164) summarises, 'According to the dominant view, causation in criminal law is a bifurcated test'. The basic idea is to reduce a causal question to two further questions as follows:

> The single question typically confronted by courts: 'Was this harm (*Y*) the consequence of this act or omission (*X*)?' is divided into two questions. First: 'Would *Y* have occurred if *X* had not occurred?' Second: 'Is there any principle which precludes the treatment of *Y* as the consequence of *X* for legal purposes?' (Hart and Honoré, 1985, p. 104)

Often, these two questions are thought to correspond to two components of causation in the law.

> [T]he central and most common form of causal relation has two different aspects which correspond roughly with the two halves of the bifurcated question. (Hart and Honoré, 1985, p. 104)

1 Note that there are other approaches, such as the NESS (necessary element of a sufficient set) approach (Wright, 1985, 2011). In this chapter, we focus on the bifurcation approach.

> It is widely accepted that the causation requirement in law actually consists of two components. (Witjens, 2014, p. 165)
>
> [L]egal causation is constituted by two distinct components. (Moore, 2019)

These two components of causation in the law are typically called *actual* (or *factual*) causation and *proximate* (or *legal*) causation. Actual causation can be thought of as a descriptive notion, while proximate causation is evaluative. An actual cause, or 'cause-in-fact', is what produces a certain harm in a physical sense, while a proximate cause is what is judged to be legally liable for a certain harm. Accordingly, a standard view is that in order to determine the liability of a defendant, one typically needs to establish two claims: the defendant's action is an actual case of a particular harm, and the defendant's action is a proximate cause of the harm. 'But-for' tests and their variants are widely employed to assess actual causes and proximate causes; for a summary of but-for tests, see Table 9.1 and Table 9.2.[2]

Table 9.1 Two types of but-for tests and their variants for actual causes (Moore, 2011, 2019).

1.	Explicitly defined counterfactual test: the defendant's action must be necessary to the occurrence of the harm
2.	Modified counterfactual tests, where the defendant's act must be:
2.a.	A necessary element of a set of factors that are together sufficient for the harm, where the total set of factors is itself unnecessary for that harm to have occurred (the "INUS" and "NESS" tests)
2.b.	Necessary to every detail in the time, place, and manner of an effect's occurrence
2.c.	Necessary to accelerations (but not retarding) of the effect
2.d.	Not necessary in the sense that the existence of the act is necessary, so long as that aspect of the defendant's action as made that act culpable, is necessary
2.e.	A "substantial factor" in the production of the harm, where the necessity of that act is an always sufficient criterion of a causal factor being "substantial" while not being a necessary criterion
2.f.	Necessary to an increase in the chance of an effect occurring (rather than being necessary to the effect actually occurring)

2 Although it is widely accepted that there are two aspects of causation in the law, some legal theorists suggest that a causal claim can be established in a simpler way. For example, Stapleton (2015) argues that there are two dimensions of legal causation: how it would have been prevented and how it was produced. Accordingly, she proposes:

> a specified factor is a cause of the existence of a particular phenomenon (as that phenomenon is individuated by the law) only if, but for that factor alone, (i) the phenomenon would not exist or (ii) an actual contribution to an element of the positive requirements for the existence of the phenomenon would not exist.
>
> (Stapleton, 2015, p. 725)

Stapleton further argues that such a causal claim can be straightforwardly established by employing her 'extended but-for' test:

> a breach [of obligation] is a cause of an injury only if, but for it alone, (i) the injury would not exist or (ii) an actual contribution to an element of the positive requirements for the occurrence of the injury would not exist.
>
> (Stapleton, 2015, p. 726)

Table 9.2 Three types of but-for tests and their variants for proximate causes (Moore, 2011, 2019).

1.	Tests regarding proximate "causation" to be a balance of policies
1.a.	Tests based on a wide range of policies
1.1.i.	Ad hoc policy balancing in each case where the resulting policy balance is honoured as a conclusion of "proximate causation" in that case
1.a.ii.	Rules justified by policy balancing, such as: 1 year and 1 day rule, first house rule, last wrongdoer rule.
1.b.	Tests based on the single policy of gauging culpability (mental state) of the actor
1.b.i.	Foreseeability test: was the harm foreseeable to the defendant as he acted?
1.b.ii.	Harm-within-the-risk test: was the harm that occurred an instance of the type of harm the risk of which (or the intent or foresight of which by the defendant) made the defendant's action negligent (or otherwise culpable)?
2.	Tests regarding proximate causation to be a matter of fact (about real causal relations) rather than a matter of policy
2.a.	Space-time proximity tests and the petering out of causation as it runs through a large number of events in the chain between cause and effect
2.b.	Direct cause test: sudden breaks in the causal chain formed by the existence of "intervening" or "superseding" causes that literally break causal chains that would otherwise exist
3.	Tests regarding proximate causation to be partly causal and partly policy: direct cause combined with the requirement that the intervening cause be unforeseeable to the defendant at the time she acted

Note that the bifurcation approach does not require a commitment to causal pluralism in the law. It is not generally assumed that there are two concepts of cause employed in the law: actual causation and proximate causation. Rather, the bifurcation approach appeals to two components or aspects of causation—two conditions that need to be met in order for *A* to be considered a cause of *B* in the law. Thus causal monism in the law remains tenable. On the other hand, one can appeal to the bifurcation approach to argue that causation in the law, which requires proximate causation, differs from causation in other contexts—in particular, from causation as used in the other social sciences—which does not hinge on questions of liability and which aligns more closely to actual causation.

This leads to a different kind of causal pluralism: one concept of cause in the law and another concept everywhere else. The idea that causation in the law is different to our usual concept of cause motivates the view that causation in the law should be treated autonomously from causation in philosophy, which tends to focus on the notion of cause prevalent outside the law. We will argue in §34 that, while there are important differences between causation in the law and causation elsewhere, the two notions are far from autonomous. Indeed, Evidential Pluralism can shed light on both notions of cause.

To understand the bifurcation approach, it is helpful to consider three possible conclusions about causation when the bifurcation approach is applied to legal cases:

1 A causal claim is established when both actual causation and proximate causation are identified by but-for tests.
2 A causal claim is not established when actual causation is identified but proximate causation is not identified.
3 A causal claim is not established when no actual cause is identified.

R v Williams (2010) is a typical case of conclusion 1.[3] The defendant, Mr Williams, was driving on a dual carriageway when Mr Loosemore was crossing the road. Mr Loosemore was hit and died the next day. Evidence from two witnesses showed that Mr Williams was not exceeding the speed limit and Mr Loosemore stepped out in front of Mr Williams's vehicle too suddenly for Mr Williams to be able to do anything to prevent the accident. However, Mr Williams had no driving licence or insurance at the time. He was then convicted of causing Mr Loosemore's death by driving without a licence and without insurance, and he was sentenced to nine months' imprisonment. Later, Mr Williams appealed by focusing on the following two points:

i) The offence could not be committed without some fault or other blameworthy conduct on the part of [Mr Williams]. "Cause" as used in the sub-section must be construed as importing some fault or other blameworthy conduct. There was no blameworthy conduct; his sole fault was a failure to have a licence and insurance, which was unrelated to the cause of the accident and the ensuing death.
ii) If that construction of "cause" was not correct, the word should be construed so that the Crown did not merely have to prove the appellant's driving was "a cause" which was not minimal but was a substantial or major cause of the death of the deceased. The facts clearly established that the substantial or major cause of death was due to the actions of Mr Loosemore and not those of [Mr Williams].

(R v Williams, 2010, ¶7)

Accordingly, two issues were debated in the court. Firstly, was another blameworthy act (in addition to driving without licence and insurance) required? Second, was it sufficient that Mr Williams' driving was a cause of Mr Loosemore's death? The judge argued that the offence of causing death is established because Mr Williams' driving whilst without a driving licence and insurance is a case of careless or inconsiderate driving under §3ZB of the Road Traffic Act 1988.[4] In

3 *R v Williams* [2010] EWCA Crim 2552, [2011] 1 WLR 588.
4 §3ZB of the Road Traffic Act 1988 is phrased as follows:

> A person is guilty of an offence under this section if he causes the death of another person by driving a motor vehicle on a road and, at the time when he is driving, the circumstances are such that he is committing an offence under-
>
> (a) Section 87(1) of this Act (driving otherwise than in accordance with a licence);
> (b) Section 103(1) of this Act (driving whilst disqualified), or
> (c) Section 143 of this Act (using a motor vehicle while uninsured or unsecured against third party risks). (Road Safety Act, 2006, §21(i))

addition, the judge argued that because Mr Williams was driving unlicensed and uninsured, he was liable for the death of Mr Loosemore by his driving, no matter how Mr Loosemore might be at fault. In other words, it is sufficient that Mr Williams' driving unlicensed and uninsured caused the death of Mr Loosemore.

From the perspective of the bifurcation approach, it is clear that Williams' driving was an actual cause of the death of the victim. As the judge concluded, Williams' driving was 'a contributing cause other than a merely minute or negligible one' (R v Williams, 2010, ¶5). Moreover, for the judge, Williams' driving without licence and insurance was a proximate cause under §3ZB of the Road Traffic Act 1988. Thus, the claim that Williams' driving caused the death of the victim was established in the context of the law, given that both actual and proximate causation can be clearly identified.

R v Hughes (2013) is an example of conclusion 2.[5] Mr Hughes was driving a van when he was confronted by a car driven by Mr Dickinson. The two vehicles collided and Mr Dickinson was killed. The collision was apparently the fault of Mr Dickinson, who was driving in the wrong direction, overtired and under the influence of heroin. In contrast, Mr Hughes was driving in a careful way and could do little to avoid the car accident. That said, he was driving without insurance. The court ruled that 'Mr Hughes had—in law—caused the death' under §3ZB (R v Hughes, 2013, ¶5). Mr Hughes appealed. The following question was then examined:

> Is an offence contrary to section 3ZB of the Road Traffic Act 1988, as amended by section 21(1) of the Road Safety Act 2006, committed by an unlicensed, disqualified or uninsured driver when the circumstances are that the manner of his or her driving is faultless and the deceased was (in terms of civil law) 100% responsible for causing the fatal accident or collision?
>
> (R v Hughes, 2013, ¶35)

The judge argued that in order to establish the claim that Mr Hughes caused Mr's Dickinson's death by his driving, it is required by §3ZB that there is some act or omission in the control of the car, which involves some element of fault or blame and which contributes to the death in some significant way (but not necessarily as a primary cause). However, the manner of Mr Hughes' driving did not contribute to the death of Mr Dickinson. Thus, the court allowed the appeal.

According to the bifurcation approach, Mr Hughes' driving was an actual cause of Mr Dickinson's death. However, Mr Hughes' driving was not identified as a proximate cause. Thus, the claim that Hughes' driving caused the death of the victim was not established in the context of the law.

Fraser v 301–52 Townhouse Corp. 2006 is a good example of conclusion 3.[6] In this case, the Frasers, former residents of a unit in a New York building owned by 301–52 Townhouse Corp., brought an action against the defendants for personal

5 *R v Hughes* [2013] UKSC 56, [2014] Crim LR 234.
6 Fraser v 301–52 Townhouse Corp. [2006] NY Slip Op 51855(U).

injuries (including respiratory problems, rash and fatigue) arising from dampness in the building and the mould that resulted from such dampness. A ten-day hearing was held to examine whether the causal claim that the existence of mould caused health problems is 'generally accepted in the relevant scientific community' and whether the methods used by the Frasers to measure the mould were generally accepted scientific methods (Fraser v 301–52 Townhouse Corp., 2006, §I(C)). On 27 September 2006, it was ordered that 'plaintiffs are precluded from introducing testimony demonstrating that mold caused their health complaints and plaintiffs' causes of action based upon personal injury are dismissed with prejudice' (Fraser v 301–52 Townhouse Corp., 2006, §II). This was largely because the plaintiffs had not demonstrated that a crucial generic causal claim is generally accepted in the relevant scientific community (i.e., specialists in occupational and environmental medicine, allergies and immunology): this is the claim that dampness can cause the kinds of harms that the Frasers allegedly suffered. Moreover, it was also noted that 'even were there a showing of causation here, this case could not go forward' (Fraser v 301–52 Townhouse Corp., 2006, §II), because there was no measure of moisture, bacteria, endotoxins, microbial volatile organic compounds and beta-glucans which were claimed to cause the health problems in the Frasers' apartment. In short, the Frasers' personal injury claims were dismissed by the court.

Later, the Frasers appealed and submitted more than 20 peer-reviewed publications that supported the generic causal claim that building dampness and mould can cause the type of irritative symptoms that they described. However, the appeal was dismissed. While admitting that 'plaintiffs established the reliability of their experts' opinions', the court highlighted that it is key 'not only to consider the general question of whether the link between building dampness and illness is generally accepted, but also that a scientific foundation existed for plaintiffs' experts' conclusion that plaintiffs were sickened by the conditions in their apartment' (Fraser v 301–52 Townhouse Corp., 2008).

According to the bifurcation approach, this case can be understood as follows. In the trial in 2006, the causal claim that mould in the Frasers' apartment caused their health problems was not taken to be established because mould was not identified as an actual cause. According to the court, the generic causal claim that indoor dampness and mould cause health problems was not generally accepted within the scientific community. In the trial in 2008, though the plaintiffs provided extra evidence to support the generic causal claim that building dampness and mould cause illness, the single-case causal claim that 'the [Frasers] were sickened by the conditions in their apartment' was not established (Fraser v 301–52 Townhouse Corp., 2008). In other words, no actual causation was identified in either trial. Thus, the causal claim that the existence of the mould in the apartment caused the Frasers' health problems was not established in the context of the law.

§32 The Bifurcation Approach and Evidential Pluralism

At first glance, there are some similarities between the bifurcation approach and Evidential Pluralism. Both approaches admit two components to establishing

causal claims and admit the use of a variety of methods. However, there are also some crucial differences. According to the bifurcation approach, one needs to establish both actual causation and proximate causation in order to establish a causal claim. According to Evidential Pluralism, on the other hand, one needs to establish both correlation and mechanism in order to establish a causal claim. The bifurcation approach employs various but-for tests to assess actual and proximate causation, while Evidential Pluralism employs a combination of association studies and mechanistic studies.

Moreover, the bifurcation approach and Evidential Pluralism are motivated by different considerations. Evidential Pluralism develops the view that establishing a correlation between *A* and *B* is not sufficient to establish that *A* causes *B*. It maintains that establishing the existence of an appropriate mechanism linking *A* to *B* is also required. The bifurcation approach, in contrast, stems from the idea that identifying an actual cause (say, *X*'s action causes a particular harm *H*) is not sufficient to determine the liability of *X*. It maintains that identifying the proximate cause is also required.

Given these differences, one might question the applicability of Evidential Pluralism to the law. In particular, it might seem that evidence of mechanisms is not important for assessing causal claims in the law, contra Evidential Pluralism. This is because the main tests used in the law are but-for tests, and these tests assess counterfactual dependence. They can be construed as providing evidence of single-case correlation—evidence of whether the chance of harm would have been substantially reduced if the defendant had done certain things differently. In none of the three cases discussed here does there seem to be much explicit discussion of mechanisms.

In response to this concern, we would argue that although the use of evidence of mechanisms is not always obvious, establishing mechanism is indeed crucial in legal cases. The need to establish mechanism is often obscured by the fact that the relevant generic mechanisms are often already well established, in which case no new evidence of mechanisms needs to be provided for them. For example, it is well established that in road traffic accidents there are mechanisms by which a collision can cause injury or death—the existence of a generic mechanism is simply not contentious here. What is in contention is whether the relevant generic mechanism is instantiated in the single case that is the subject of the legal action. Evidence that the mechanism was operational in a single case may include tyre markings on the road surface formed when braking, witness statements and photographic evidence of the collision scene, for example. In conjunction with established generic mechanisms, this evidence helps to establish the existence of a mechanism in the single case in question.

In some situations, however, the generic mechanism is not well established. Consider again the case of Fraser v 301–52 Townhouse Corp. A key issue was whether the claim that building dampness causes certain illnesses is 'generally accepted in the relevant scientific community' (Fraser v 301–52 Townhouse Corp., 2006, §I(C)). This is a generic causal claim and, as we have argued in previous chapters, establishing mechanism is required in order to establish a causal

claim. Thus Evidential Pluralism applies at this stage. Note that whether there is sufficient evidence to establish the generic causal claim is a question for the relevant scientific community, not for a particular court hearing; the legal question is whether the relevant scientific community takes the generic causal claim to be established. Thus, there is a role for Evidential Pluralism at this stage, but it is at a step removed from the court. There is also a role for Evidential Pluralism at the next stage: determining whether the corresponding single-case causal claim is established. The question at this next stage is whether dampness caused the Frasers harm. Again, this is a matter of assessing correlation and mechanism. That the court decided that single-case causation was not established might be attributable to the lack of an established single-case mechanism, or to the lack of an established single-case correlation, or both.

In other situations, the importance of establishing mechanism can be much more obvious. Let us consider the case of R v Pagett 1983.[7] On 11 June 1980, Mr Pagett took Ms Gail Kinchen as a hostage during a confrontation with armed police. Mr Pagett fired shots at police whilst using Ms Kinchen as a human shield. The police fired back, and one of their shots killed Ms Kinchen. Mr Pagett was convicted of Ms Kinchen's manslaughter, but he appealed on the grounds of lack of causation. Mr Pagett argued that the police officers' shooting at the victim constituted a *novus actus interveniens* (i.e., an intervening act or event that breaks the causal chain between a crime committed by the defendant and subsequent harms and therefore relieves the defendant from liability for these harms).

Mr Pagett's conviction for manslaughter was upheld, however. The judge argued that the police officers acted in self-defence when they returned fire: their actions were not free but involuntary. Rather, it is Mr Pagett who caused Ms Kinchen's death by using her as a human shield while shooting at the police officers. The police officers would not have fired the shots which killed Ms Kinchen *but for* Mr Pagett taking her hostage and using her as a human shield. In other words, Mr Pagett remained liable for Ms Kinchen's death.

In this case, establishing the existence of an appropriate mechanism is key to establishing the causal claim that Mr Pagett's firing caused the death of Ms Kinchen As the judge indicated,

> [I]f you were satisfied that [Mr Pagett] did those 2 unlawful and deliberate acts [i.e., (1) the firing of the gun at the police officers (which he explained could constitute an assault), and (2) the physical force applied to Gail so that her body could be used as a shield], the question now becomes whether by those acts he caused or was a cause of Gail's death. It sometimes happens that difficult questions arise when a jury has to decide whether something is a cause of the death of the victim. This is just such a case. In those circumstances it is for me to decide as a question of law whether by his unlawful and deliberate acts the defendant caused or was a cause of Gail's death, but the

7 *R v Pagett* [1983] EWCA Crim 1, [1983] Crim LR 394.

> answer to that question of law depends upon findings of fact which you alone can decide, and accordingly I have to direct you that if you find the facts I am about to mention proved beyond all reasonable doubt, then the defendant would have caused or been a cause of Gail's death. It might help you if I explain that the act of an accused person on the charge of murder need not be the sole cause or even the main cause of the death of the victim. It is not necessary to prove actual physical violence on the victim by the accused, but let me illustrate that point by an old case which came before the courts years and years ago. A man quarrelled with the woman he was living with. He ran towards her in order to hit her but he did not succeed in doing so. In fear or in retreat from the attack she jumped or fell from a window and was killed. The courts held that he, although he had not laid a finger on her, was a cause of her death.
>
> I turn now to the question whether the defendant caused or was a cause of Gail's death. I am going to mention the facts. If you are sure that the following facts have been proved beyond all reasonable doubt, then the defendant would have caused or would have been a cause of Gail's death. First of all, that he fired the shot-gun deliberately at the police officers before any shot was fired by them. In other words, sure that he fired first. Secondly that his act in firing at the police officers caused them to fire back with the result that bullets from their weapons shot Gail and caused her death. Next, that in firing back for that reason the police acted reasonably either by way of self-defence or in the performance of their duty as police officers, or both. I will explain that in a little more detail in a moment. Lastly, that from the beginning to the end of the firing Gail was being used against her will and by force by the defendant as a shield to protect him from any shots fired by the police. If you are not sure about any of those matters, acquit him, and you will acquit him of course because the chain which links his deliberate and unlawful acts with Gail's death will have been broken.
>
> (R v Pagett, 1983)

The existence of an appropriate mechanism clearly needs to be established in order to establish the causal claim in this case. According to the judge, the jurors needed to establish 'both that [Mr Pagett] fired at the police officers and thereby caused them to fire back, and that he used Gail Kinchen as a shield by force and against her will' in order to establish causation (R v Pagett, 1983). The first component (i.e., that Mr Pagett fired at the police officers and thereby caused them to fire back) refers to a mechanistic process from Pagett's firing to the death of the victim that proceeds via the police officers' firing. In order to decide whether the causal claim is established, the jury needs to decide whether this specific mechanism hypothesis is established on the basis of the evidence presented to them. This provides an illustration of how evidence of mechanisms can be relevant to establishing a causal claim in the legal context.

Certain problematic cases remain, however. In some legal cases, both correlation and mechanism are established, and thus causation—in the sense used elsewhere in the book—is established, but there is no proximate causation and thus legal

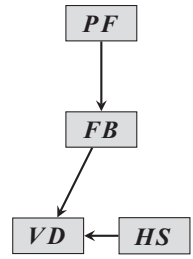

Figure 9.1 Liability-tracing mechanism of R v Pagett 1973. *PF*: Pagett's firing; *FB*: the policemen's firing back; *HS*: Pagett's taking Kinchen as a human shield; *VD*: the death of Kinchen.

causation is not established. Consider a classic case: Ryan v New York C.R. Co.[8] On 15 July 1854, New York Central Railroad Company's woodshed was set on fire by its careless management of a rail engine and the fire spread to and destroyed Mr Ryan's house, which was 130 feet away from the woodshed. Mr Ryan sued the New York Central Railroad Company, but the judge terminated Mr Ryan's action. This was affirmed following the subsequent appeal.

It is clear in this case that both correlation and mechanism are established. There is good evidence that the fire started by the New York Central Rail Road Company raised the probability of the burning of Mr Ryan's house. There is also evidence that there is a mechanistic process from the fire set by the rail engine to the burning of Mr Ryan's house. However, by referring to the general principle that 'every person is liable for the consequences of his own acts', the judge argued that one 'is liable in damages for the proximate results of his own acts, but not for remote damages' (Ryan v New York Central Railroad, 1866, p. 211). Thus, it was decided that the fire set by the New York Central Rail Road Company was not a proximate cause, and thus not a cause of the destruction of Mr Ryan's house in the context of the law.

As noted earlier, there is a distinction between causation as used in the law, which usually needs to take rules of liability into account, and our general notion of causation, which does not. These two kinds of causation come apart in the earlier example. The question remains whether Evidential Pluralism can say anything meaningful about causation in the law where these two kinds of causation come apart. This is a question that we will address in the next section.

§33 Liability-Tracing Mechanisms

In this section, we suggest that Evidential Pluralism can shed light on causation in the law, even in situations where causation in the law comes apart from our usual notion of causation. In order to do this, Evidential Pluralism needs to be

8 Ryan v New York Central Railroad [1866] 35 N.Y. 210.

modified slightly in order to accommodate the difference between causation in the law and our usual notion of causation. According to this modified account, in order to establish a causal claim in the the law one normally needs to establish correlation and mechanism, as usual, but, in the context of the law, the correlation and mechanism will typically be single-case, and the mechanism will be a *liability-tracing mechanism*.

A *liability-tracing mechanism* from *A* to *B* is a mechanism from *A* to *B* that satisfies the legal rules needed to hold *A* (or the agent responsible for *A*) to be liable for *B*. Thus, a liability-tracing mechanism is an ordinary mechanism along which one can successfully trace liability for the putative effect.

In the case of R v Pagett 1973, we have the following liability-tracing mechanism. Pagett's firing (*PF*) led to the police firing back (*FB*). This firing back, together with Pagett taking the victim as a human shield (*HS*), led to the death of the victim (*VD*) (see Figure 9.1). Such a liability-tracing mechanism was implicit in the argument that Pagett's firing was a proximate cause of the death of the victim.

In the case of R v Williams 2011, we also have a liability-tracing mechanism. In this case, liability can be traced back from the victim's death (*VD*) to Williams' driving, in virtue of the fact that Williams' driving was an instance of careless or inconsiderate driving (*WCD*), on account of his failure to obtain a driving licence (*WFDL*) (see Figure 9.2). This mechanism is a liability-tracing mechanism because it satisfies a legal condition: 'A person is guilty of an offence under this section if he causes the death of another person by driving a motor vehicle on a road and, at the time when he is driving, the circumstances are such that he is committing an offence under . . . (c) Section 143 of this Act (using a motor vehicle while uninsured or unsecured against third party risks)' (Road Safety Act, 2006, §21(i)).

In each of these two cases, we have a single-case correlation established by means of but-for tests alongside a liability-tracing mechanism, so the corresponding causal claims are established.

In contrast, in those cases where causal claims are not established due to no proximate cause being identified, this is often because no liability-tracing mechanism has been established. Consider the case of R v Hughes 2013. In the original decision, Mr Hughes was convicted based on the decision of R v Williams 2011. Thus, a similar liability-tracing mechanism was posited: Hughes' failure to obtain insurance (*HFI*) led to Hughes' careless or inconsiderate driving (*HCD*), which in turn significantly contributed to the death of the victim (*VD*) (see Figure 9.3). However, in the Court of Appeal in 2013, this liability-tracing mechanism was questioned. In particular, doubt was cast on the link from *HCD* to *VD*: 'there is no suggestion that there was anything which the defendant either did or omitted to do in the driving of the car which contributed to the least extent to the fatality' (R v Hughes, 2013, p. 13). Although the driving was an actual cause of the victim's death, the idea is that liability for the death cannot be traced back to the driving because the omission to obtain insurance was not a part of the mechanistic process actually responsible for the death. Thus, the causal claim was not taken to be established.

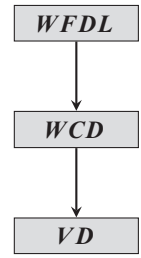

Figure 9.2 Liability-tracing mechanism for R v Williams 2011. *WFDL*: Williams' failure to obtain a driving licence; *WCD*: Williams' careless or inconsiderate driving; *VD*: the death of Loosemore.

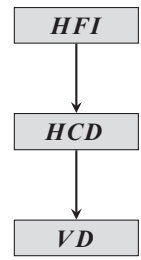

Figure 9.3 Liability-tracing mechanism for R v Hughes 2011. *HFI*: Hughes' failure to obtain insurance; *HCD*: Hughes' careless or inconsiderate driving; *VD*: the death of Dickinson.

§34 Against Causal Autonomy in the Law

The thesis of causal autonomy in the law can be spelt out as follows:

> *Causal Autonomy.* The law ought to be autonomous from philosophy in its use of the concept of causation.

We shall argue in this section that this thesis does not hold.

Let us examine Jane Stapleton's arguments for causal autonomy in the law. There are four main arguments: one argument from complexity, another from contextuality, another from contrastivity and a fourth argument from relevance. All these arguments are rooted in apparent differences between philosophical and legal uses of causation.

The argument from complexity arises from two observations. First, causes in the law refer to 'a wide range of phenomena' (Stapleton, 2015, p. 699). Second, causal enquiry in the law is 'heterogeneous' (Stapleton, 2015, p. 699). As Stapleton elaborates,

> [Law] seamlessly accommodates as 'causes' factors variously characterised as specific events, facts, states of affairs, aspects of conduct events or things,

> absences, omissions, information and reasons, including entirely abstract phenomena such as a specific marriage, divorce, forfeiture and breach of obligation. (Stapleton, 2015, p. 699)

In contrast, it seems to many legal theorists that philosophers think 'there is only one relation that is "causal"' (Stapleton, 2015, p. 701). In short, causal enquiry in the law is much more complicated than causal analysis in philosophy.

> [Law] is interested not only in the states of affairs on which philosophers focus (such as a cannonball's state of rest upon a cushion) but also in abstract, often law-designated, states such as the legal validity of a certain document or the state of marriage. Similarly, law is interested not only in the transitions on which philosophers focus (such as the making of a statue, the collision of billiard balls or the occurrence of a physical injury) but also in abstract transitions such as entry into a commercial transaction, the revocation of a will or the forfeiture of a lease. (Stapleton, 2015, p. 699)

Thus, it seems to some legal theorists that the complexity of causal phenomena in the law provides a good reason to advocate conceptual autonomy in the law (Stapleton, 2008, 2015). The argument for causal autonomy from complexity can be formulated as follows:

CAP1. Causes, effects and causal enquiry in the law are heterogeneous.

CAP2. Philosophical accounts of causation accommodate only a narrow range of phenomena.

CAP3. Philosophical accounts of causation cannot accommodate the complexity of causal enquiry in the law.

CA. Therefore, the law ought to be autonomous from philosophy in its use of the concept of causation.

There are various problems with such an argument, however. Firstly, CAP2 is arguably false. Philosophers are not merely interested in analysing the causation between cannonballs, even if they may use the collision of cannonballs as an illustrative example. Philosophers seek to analyse all sorts of causal phenomena. Moreover, there are philosophical theories of causation that do not impose restrictions on the kinds of things that can be causes and effects. For example, the epistemic theory of causality of §5 imposes no such restrictions. Second, CAP3 is controversial. For example, Schaffer (2010) argues that his contrastive account of causation provides a description of causal enquiry in legal practice. Even some legal theorists challenge CAP3. For example, Wright (1985) defends causal monism in the law and argues that his NESS [necessary element of a sufficient set] test 'captures the essential meaning of the concept of causation' (Wright, 1985, p. 1789). Thus the argument from complexity fails to justify the thesis of causal autonomy in the law.

The argument from contextuality stems from the view that causation in the law is context-sensitive, while causation in philosophy is context-free. As Stapleton elaborates,

> [I]n contrast to the context-free project of philosophers . . ., the conceptual framework and methodology of the legal project provide contextualizing devices that render finite the number of factors whose possible involvement is subject to investigation; that individuate (at a level of modal fragility to serve law's purposes) the specified factor whose influence is being examined and therefore the hypothetical worlds with which comparison is being made; and that individuate the phenomenon (again, at a level of specificity to serve law's purposes) in relation to which that influence is judged, an individuation of outcome that often excludes any problem of pre-emption.
>
> (Stapleton, 2009, p. 750)

The argument for causal autonomy from contextuality can be formulated as follows.

CAC1. Causation in the law is contextual.

CAC2. Philosophical accounts of causation are non-contextual.

CA. Therefore, the law ought to be autonomous from philosophy in its use of the concept of causation.

Again, CAC2 is based on an oversimplified understanding of the philosophical literature on causation. Certain philosophical accounts of causation maintain that causation is contextual in nature. For example, Reiss (2009, p. 34) maintains that '[t]he value of investigating the truth of causal hypotheses lies in the degree to which these claims help in realizing scientists' purposes and in the value of realizing these purposes'. Accordingly, Reiss suggests that causal claims are contextual in nature: they depend upon the purposes and methods of the context. Menzies (2007) also develops a philosophical account of causation that renders causation contextual.

A related argument is the argument from contrastivity, which is identified as another crucial difference between causation in philosophy and in the law.

> So in addressing the existence of a phenomenon in the actual world (a baby's death) in which the obliged party omitted to fulfil an obligation (to feed the baby), the law is especially interested in the contrastive information that, in the hypothetical 'counterfactual' world where that specific omission was reversed (and the obligation to feed was fulfilled), the phenomenon would not have existed.
>
> (Stapleton, 2015, p. 702)

The argument for causal autonomy from contrastivity can be formulated as follows.

CACon1. Causation in the law is contrastive.

CACon2. Philosophical accounts of causation are non-contrastive.

CA. Therefore, the law ought to be autonomous from philosophy in its use of the concept of causation.

CACon2 is arguably false. As we noted, Schaffer (2005, 2010) develops a contrastive account of causation and applies it to the law. Moreover, the question arises as to whether 'contrastive' is the right word here. The feature of causation in the law that Stapleton highlights might be classed as counterfactual rather than merely contrastive. By contrasting the actual event with what would have happened in a hypothetical world, legal theorists and lawyers employ counterfactual reasoning: had *X* not existed, would *Y* have been the same? As Stapleton (2008, p. 435) observes, 'Often we [legal theorists and lawyers] are interested to compare the actual world of the particular phenomenon (which, of course, includes our specified factor) with a hypothetical world (which we construct by notionally omitting the specified factor and sometimes other factors). By doing this we can then determine, in the context of that comparison, the role the specified factor played, if any, in the existence of the actual phenomenon'. If 'contrastive' is used in the sense of 'counterfactually contrastive', then CACon2 is clearly false. Philosophers have appealed to counterfactual accounts of causation since, arguably, the time of Hume—Lewis (1973) champions such an account, for example.

Another related argument appeals to the idea of relevance. It has been argued that unlike philosophers, legal theorists and lawyers are free to take any relevant factors into consideration to establish causal claims:

> Because the law has no need to resolve metaphysical disputes about the concept of 'causation' it is free to choose what character relations must have before it will describe them as 'causal' relations. (Stapleton, 2015, p. 702)

> Given the range of investigations modern Law needs to make, lawyers should choose an interrogation underlying causal usage in the Law that captures all ways in which the factor might be involved (identified by our knowledge of the physical laws of nature, evidence of behaviour and so on) in the existence of the particular phenomenon in issue. (Stapleton, 2008, p. 441)

> In the important context of the Law, lawyers must designate the underlying interrogation the results of which we are reporting when we use causal language for legal purposes. Here lawyers are at a considerable advantage relative to philosophers because the legal project is always focused and specified: was it *A* who stabbed *B*? Did the lie that *C* told *A* prompt *A* to stab *B*? And so on. In this sense there is a significant affinity between the Law and the precise focus that characterises a scientific experiment, a focus that is achieved by the application of scientific method (which requires explicit specification of the particular phenomenon under investigation, the specified factor whose role is in issue and so on). (Stapleton, 2008, p. 440)

The argument for causal autonomy from relevance can be formulated as follows:

CAR1. Relevance is used in a distinctive way in causal inference in the law.

CAR2. This usage cannot be properly explained by philosophical accounts of causation.

CA. Therefore, the law ought to be autonomous from philosophy in its use of the concept of causation.

We are sympathetic to the view that there is something special or distinctive about causation in the law. As Moore (2019, §2.1) indicates, whether a causal claim is established or not needs to be 'resolved by arguments of policy'. That being said, this does not imply that philosophical work on causality has little of use to say about causation in the law. Nor does it suggest that causation in the law is incommensurable with, or independent of, causation as used elsewhere. Indeed, many legal theorists' works on causation (e.g., Wright, 1985; Hart and Honoré, 1985; Moore, 2011, 2019) have been influenced by philosophical analyses of causation. Moreover, as we have argued in this chapter, one particular philosophical treatment of causation, Evidential Pluralism, can provide an account of causal enquiry in the law that mirrors the account it gives of causal enquiry elsewhere.

10 The Scope of Evidential Pluralism in the Social Sciences

§35 Across the Social Sciences

In Part I, we developed a general account of Evidential Pluralism, no longer tied to the biomedical sciences, which was the context in which the theory originated. To be sure, there are limits to the extent to which Evidential Pluralism can be applied across the sciences. For example, Evidential Pluralism is less applicable to the mathematical sciences, where causal enquiry is apparently not such a central activity.[1] The main message of this book is that Evidential Pluralism can be fruitfully applied across the range of the social sciences.

In Part II, we argued that Evidential Pluralism leads to a new approach to evidence-based policy, and that it can motivate mixed methods research. Policy making is a task that cuts across the social sciences, and mixed methods research is a methodology that can be applied across the social sciences, so these methodological implications of Evidential Pluralism help to demonstrate its broad scope.

In Part III, we considered some particular social sciences in more detail and argued that Evidential Pluralism can be fruitfully applied to sociology, economics, political science and law. These disciplines vary in the extent to which current practice conforms to the norms of Evidential Pluralism. Where existing practice already conforms, Evidential Pluralism can shed light on the roles of particular methods and research designs in causal enquiry. Elsewhere, Evidential Pluralism can provide new strategies for causal enquiry. Either way, thinking in terms of Evidential Pluralism brings benefits.

These conclusions do not hinge on the particular selection of disciplines in Part III: Evidential Pluralism can be fruitfully applied wherever a social science is engaged in causal enquiry. In this section, we shall sketch how Evidential Pluralism might be applied to other social sciences. We will briefly consider anthropology, psychology, demography, geography, management science and education research here, to give a flavour of the broad scope of Evidential Pluralism.

Different social sciences engage in causal enquiry to a different extent. In anthropology, for example, much research is devoted to describing and conceptualising cultural phenomena, rather than to causal enquiry. Nevertheless, causal

1 See Williamson (2021b) for one potential application, however.

DOI: 10.4324/9781003143000-13

enquiry remains key to some research. For example, Matthews et al. (2022) looked at the relationship between parents' beliefs and their hesitancy in getting their children vaccinated. They found that beliefs are better predictors of vaccine hesitancy than demographics or social network effects. While this may seem at first sight to be a study of association rather than causation, they use their results to support certain pro-vaccination messaging strategies. Thus their research does have a causal component, because causal claims are required to support interventions such as pro-vaccine policies. From the point of view of Evidential Pluralism, their study can be thought of as a mechanistic study, shedding light on the mechanisms that are intervened upon by these pro-vaccination strategies. From this point of view, their results support the claim that parental beliefs are a key factor in mechanisms for vaccine hesitancy. This provides an example of the way in which Evidential Pluralism can help us understand the role of a particular study in causal enquiry in anthropology.

In psychology, causal enquiry is more central. There, association studies are often combined with the mechanistic studies in order to confirm causation. For example, Felig et al. (2022) were interested in whether self-objectification prevents women from feeling cold. Their study found an association, and they identified a range of evidence of mechanisms (pp. 456–457), but they viewed the evidence base to be insufficient to rule out confounding (p. 465) and so to fall short of establishing causation. Lee et al. (2022a), meanwhile, confirmed the claim that reminders of COVID-19 social distancing can intensify physical pain, by appealing to prior mechanistic work that shows that social and physical pains share a common neural alarm system and by performing two experimental association studies to test for correlation. These examples indicate that causal enquiry in psychology can indeed be cast in terms of Evidential Pluralism.

Demography provides an interesting example, because one might think it appeals only to observational association studies, such as large cohort studies, in order to establish causal claims. This is not the case. Angelini et al. (2022), for example, use observational studies to establish an association between a child growing up in a 'golden nest' (i.e., in a family with high socioeconomic status) and the age at which that child eventually leaves home, conditional on potential confounders. But in order to causally explain why some children leave home at a greater age, they do not only appeal to this association. They confirm three specific mechanism hypotheses: that education mediates on one pathway between family socioeconomic status and age at leaving home, but not on every pathway; that children leave home later to maintain standards of living; and that children leave home later to improve access to economic resources (parental income and their own lifetime income processes). They appeal to a life-cycle model in order to articulate this last mechanism hypothesis in more detail.

Causal enquiry also plays a role in geography—and where causal enquiry is to be found, Evidential Pluralism can offer advice. For example, Wu (2022) finds an association between dry days during heatwaves and road traffic collisions in Alabama, conditional on a range of potential confounders. He does not infer causality, however, as bias has not been ruled out. On the other hand,

Wu does intend that the results should influence law enforcement policy, which would only be warranted if the results have causal import. Thus the question arises as to what to do in order to better confirm causality. Wu (2022, p. 1324) notes, 'The possible reasons behind the positive associations found for driving in nonprecipitation conditions include its higher traffic volume or more vehicles running on the roads than in the precipitation condition'. These can be viewed as two specific mechanism hypotheses, and these hypotheses would be fairly straightforward to test. By confirming one or both of these hypotheses, the causal import of Wu's study would be significantly strengthened, and his policy suggestions would be more compelling. But without supporting evidence, the mechanism hypotheses have no confirmatory value, according to Evidential Pluralism.

An example from management science shows this strategy in action. Jacob et al. (2022) argue that tax uncertainty causes firms to delay large capital investments. The authors carried out an association study which estimated an average delay of about 4.5 months, controlling for firm size. Crucially,

> We also provide evidence for the mechanism underlying the effect of tax uncertainty on capital expenditures. We argue that investment decisions of managers facing financial constraints (firms with more costly outside financing) are more likely to be affected by tax uncertainty. Hence, managers of such firms likely delay large investments and sideline cash in response to tax uncertainty. Consistent with this prediction, we find that the effect of tax uncertainty on investment timing is concentrated among financially constrained firms and that these firms sidelined cash in response to [tax uncertainty]. We also show that [tax uncertainty] reduces the sensitivity of investment to growth opportunities.
>
> (Jacob et al., 2022, p. 4067)

This strategy of appealing to mechanistic studies alongside association studies, which can be justified by appeal to Evidential Pluralism, allows the authors to draw causal conclusions with some confidence.

In education research, meta-analysis and systematic review are key tools for causal enquiry. A meta-analysis usually acts as a kind of association study, and can be very informative in certain situations—for example, where there are several large, well conducted RCTs among the primary studies. Where a meta-analysis is inconclusive on its own, it can be helpful to augment it with mechanistic studies. For example, Lee et al. (2022b) used a meta-analysis to argue that dyslexia negatively affects the learning of statistical patterns of complex environmental input. In this scenario, one clearly cannot randomly assign the cause, dyslexia, to individuals. The question thus arises as to how one can shore up causal conclusions drawn from this sort of association study. An appeal to mechanisms can help. Lee et al. hypothesised several mediating and moderating variables—including working memory, rapid automatised naming, phonological skills and reading ability—on the mechanism from dyslexia to the learning of complex statistical patterns,

and cited some evidence that these variables are associated with dyslexia. They also developed a 'statistical learning and reading' model, which articulates a set of specific mechanism hypotheses. Unfortunately, they did not assess these specific mechanism hypotheses in a systematic way, so this model does not play a confirmatory role in their analysis. They hence concluded that 'the relationships obtained in the meta-regression cannot be used to prove causality' (Lee et al., 2022b, p. 681). A more systematic assessment of the specific mechanism hypotheses could have led to a more conclusive study.

These examples give a flavour of the broad scope of Evidential Pluralism in the social sciences. Evidential Pluralism can be relevant in various ways, many of which we have encountered in previous chapters. In linguistics, for example, Busse and Walter (2013) carry out a longitudinal mixed methods study to argue that intrinsic motivation and self-efficacy beliefs affect students' engagement with foreign language learning. The relevance of Evidential Pluralism to mixed methods research has been developed in detail in Chapter 4. In archaeology and history, single-case causal claims and single-case causal explanation feature prominently. The single case has been discussed in §30.2, for example. Our general thesis is that, although Evidential Pluralism may say different things about different kinds of causal enquiry, wherever there is causal enquiry in the social sciences, Evidential Pluralism has something relevant to say.

§36 Where We Stand

In this book, we have argued that Evidential Pluralism has several merits. It provides a general framework for causal enquiry that can offer guidance with respect to both establishing and assessing causal claims. It can clarify the role of specific methods for causal enquiry. It can motivate methodological diversity. It is a theory of the epistemology of causation, not tied to any particular metaphysical view or conceptual analysis of causation—many of which are highly questionable if not strictly indefensible. It validates good examples of causal enquiry in the social sciences and can suggest strategies for improving one's own causal enquiries.

Most of all, Evidential Pluralism offers a new way of thinking about causal enquiry. Thinking about things differently can open up new possibilities and provide new insights. As Herbert Butterfield points out when considering the origins of modern science,

> in both celestial and terrestrial physics—which hold the strategic place in the whole movement—change is brought about, not by new observations or additional evidence in the first instance, but by transpositions that were taking place inside the minds of the scientists themselves. In this connection it is not irrelevant to note that, of all forms of mental activity, the most difficult to induce even in the minds of the young, who may be presumed not to have lost their flexibility, is the art of handling the same bundle of data as before, but placing them in a new system of relations with one another by giving them a

> different framework, all of which virtually means putting on a different kind of thinking-cap for the moment.
>
> (Butterfield, 1949, p. 13)

Evidential Pluralism offers this different kind of thinking-cap for understanding causal enquiry.

Although Evidential Pluralism can motivate certain strategies for evaluating social policies, it is not a recipe that eradicates the need for expert judgement. Domain expertise is required to select appropriate standards for establishing a causal claim, and more generally for determining the status of the various claims—causal, correlational, mechanistic—to which Evidential Pluralism appeals. Domain expertise is also crucial in order to determine potential confounders for an assessment of correlation, and to assess confirmation along the various evidential channels of Figures 1.1 and 1.3. Evidential Pluralism is an aid, but not a substitute, for expertise.

This is not to suggest that no more can be said about standards for establishing, what counts as a potential confounder and the other aspects of causal enquiry that demand expert judgement. Indeed, this is an important area for further research.

Another crucial task for further research is to say more about how Evidential Pluralism bears on single-case causal enquiry in the social sciences. In particular, we have only briefly touched on the role of specific mechanism hypotheses in establishing single-case correlation (channel μ_3 of Figure 1.1), and there is plenty of scope for providing further detail here.

While this book has adopted an epistemology-driven approach and has largely sought to avoid questions of the metaphysics and conceptual analysis of causality, these metaphysical and conceptual questions remain. It would be interesting to assess in more depth which theories of causality are compatible with Evidential Pluralism. We have suggested that the epistemic theory of causality is compatible with Evidential Pluralism and that causal pluralism often does not sit well with Evidential Pluralism. But much more remains to be said here.

Bibliography

Abdin, Y., Auker-Howlett, D. J., Landes, J., Mulla, G., Jacob, C., and Osimani, B. (2019). Reviewing the mechanistic evidence assessors E-Synthesis and EBM+: A case study of amoxicillin and drug reaction with eosinophilia and systemic symptoms (DRESS). *Current Pharmaceutical Design*, 25(16):1866–1880.

Abrevaya, J. and Puzzello, L. (2012). Taxes, cigarette consumption, and smoking intensity: Comment. *The American Economic Review*, 102(4):1751–1763.

Adams, P., Hurd, M. D., McFadden, D., Merrill, A., and Ribeiro, T. (2003). Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics*, 112(1):3–56.

Adda, J., Chandola, T., and Marmot, M. (2003). Socio-economic status and health: Causality and pathways. *Journal of Econometrics*, 112(1):57–63.

Adda, J. and Cornaglia, F. (2006). Taxes, cigarette consumption, and smoking intensity. *The American Economic Review*, 96(4):1013–1028.

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., and Syme, S. L. (1994). Socioeconomic status and health. The challenge of the gradient. *The American Psychologist*, 49(1):15–24.

Adler, N. E. and Newman, K. (2002). Socioeconomic disparities in health: Pathways and policies. *Health Affairs*, 21:60–76.

Agerfalk, P. J. (2013). Embracing diversity through mixed methods research. *European Journal of Information Systems*, 22:251–256.

Albertini, M. and Dronkers, J. (2009). Effects of divorce on children's educational attainment in a Mediterranean and Catholic society. *European Societies*, 11(1):137–159.

Andreoni, J. and Gee, L. K. (2011). The hired gun mechanism. *Working Paper 17032*, National Bureau of Economic Research.

Angelini, V., Bertoni, M., and Weber, G. (2022). The long-term consequences of a golden nest: Socioeconomic status in childhood and the age at leaving home. *Demography*, 59(3):857–875.

Anscombe, G. (1971). Causality and determination. In Sosa, E. and Tooley, M., editors, *Causation*, pages 88–104. Oxford University Press, Oxford, 1993 edition.

Aronson, J. K., Auker-Howlett, D., Ghiara, V., Kelly, M. P., and Williamson, J. (2021). The use of mechanistic reasoning in assessing coronavirus interventions. *Journal of Evaluation in Clinical Practice*, 27(3):684–693.

Aronson, J. K., La Caze, A., Kelly, M. P., Parkkinen, V.-P., and Williamson, J. (2018). The use of mechanistic evidence in drug approval. *Journal of Evaluation in Clinical Practice*, 24(5):1166–1176.

Ashley, M., Saunders, P., Mullard, G., Prasad, K., Mariner, D., Williamson, J., and Richter, A. (2011). Smoking intensity before and after introduction of the public place smoking ban in Scotland. *Regulatory Toxicology and Pharmacology*, 61(3 Suppl.):S60–S65.

Atkins, D., Best, D., Briss, P. A., Eccles, M., Falck-Ytter, Y., Flottorp, S., Guyatt, G. H., Harbour, R. T., Haugh, M. C., Henry, D., Hill, S., Jaeschke, R., Leng, G., Liberati, A., Magrini, N., Mason, J., Middleton, P., Mrukowicz, J., O'Connell, D., Oxman, A. D., Phillips, B., Schünemann, H. J., Edejer, T. T.-T., Varonen, H., Vist, G. E., Williams, J. W. J., and Zaza, S. (2004). Grading quality of evidence and strength of recommendations. *BMJ (Clinical Research Ed.)*, 328:1490.

Auker-Howlett, D. (2020). *Evidence Evaluation and the Epistemology of Causality in Medicine*. PhD thesis, University of Kent. https://kar.kent.ac.uk/82097/1/287PhD_Thesis_ Daniel_Auker-Howlett.pdf.

Auker-Howlett, D. and Wilde, M. (2020). Reinforced reasoning in medicine. *Journal of Evaluation in Clinical Practice*, 26(2):385–681.

Baron, J. (2018). A brief history of evidence-based policy. *The ANNALS of the American Academy of Political and Social Science*, 678(1):40–50.

Baumgarther, F. and Jones, B. D. (1993). *Agendas and Instability in American Politics*. The University of Chicago Press, Chicago.

Beach, D. (2021). Evidential pluralism and evidence of mechanisms in the social sciences. *Synthese*, 199:8899–8919.

Beach, D. and Pedersen, R. B. (2013). *Process-Tracing Methods*. The University of Michigan Press, Ann Arbor, MI.

Beach, D. and Pedersen, R. B. (2019). *Process-Tracing Methods*. The University of Michigan Press, Ann Arbor, MI, 2 edition.

Becker, G. S. and Murphy, K. M. (1988). A theory of rational addiction: The case of cigarettes. *Journal of Political Economy*, 96(4):675–700.

Befani, B. and Mayne, J. (2014). Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation. *IDS Bulletin*, 45(6):17–36.

Bennett, A. (2010). Process tracing and causal inference. In Brady, H. E. and Collier, D., editors, *Rethinking Social Inquiry*, pages 207–220. Rowman and Littlefield, Lanham, MD, 2 edition.

Bernard, C. (1865). *An Introduction to the Study of Experimental Medicine*. Dover, New York, 1957 edition. Trans. Henry Copley Greene.

Bernardi, F. and Boertien, D. (2017). Non-intact families and diverging educational destinies: A decomposition analysis for Germany, Italy, the United Kingdom and the United States. *Social Science Research*, 63:181–191.

Bhaskar, R. (1975). *A Realist Theory of Science*. Verso, London.

Bhaskar, R. (2011). *Critical Realism: A Brief Introduction*. Routledge, London.

Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 95–118. Sage Publications, Inc., Thousand Oaks, CA, 2 edition.

Black, F. (1970). Banking and interest rates in a world without money. In *Business Cycles and Equilibrium*. Basil Blackwell, Oxford.

Black, T. R. (1999). *Doing Quantitative Research in the Social Sciences*. Sage, London.

Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure*. Wiley, New York.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., and Moore, L. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine*, 75(12):2299–2306. Special Issue: Place, migration & health.

Bosworth, B. (2018). Increasing disparities in mortality by socioeconomic status. *Annual Review of Public Health*, 39(1):237–251.

Boudon, R. (1998). Social mechanisms without black boxes. In Hedström, P. and Swedberg, R., editors, *Social Mechanisms: An Analytic Approach to Social Theory*, pages 172–203. Cambridge University Press, Cambridge.

Bourdieu, P. (1973). Cultural reproduction and social reproduction. In Brown, R., editor, *Knowledge, Education, and Cultural Change: Papers in the Sociology of Education*, pages 71–112. Tavistock, London.

Boustani, M. M., Frazier, S. L., Becker, K. D., Bechor, M., Dinizulu, S. M., Hedemann, E. R., Ogle, R. R., and Pasalich, D. S. (2015). Common elements of adolescent prevention programs: Minimizing burden while maximizing reach. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(2):209–219.

Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, 31:223–243.

Burdett, K., Lagos, R., and Wright, R. (2003). Crime, inequality, and unemployment. *The American Economic Review*, 93(5):1764–1777.

Busse, V. and Walter, C. (2013). Foreign language learning motivation in higher education: A longitudinal study of motivational changes and their causes. *The Modern Language Journal*, 97(2):435–456.

Butterfield, H. (1949). *The Origins of Modern Science 1300–1800*. The Free Press, New York, 2 (1957) edition.

Cabinet Office (2018). *What Works Network: Membership Criteria*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/747965/WWmembership-paper_ Oct2018.pdf.

Calvó-Armengol, A. and Zenou, Y. (2003). Does crime affect unemployment? The role of social networks. *Annales d'Économie et de Statistique*, 71/72:173–188.

Campbell Collaboration (2020). Campbell systematic reviews: policies and guidelines. *Campbell Policies and Guidelines*. https://doi.org/10.4073/cpg.2016.1.

Campbell, D. T. (1974). Evolutionary epistemology. In Schilpp, P. A., editor, *The Philosophy of Karl Popper*, pages 413–463. Open Court, La Salle, IL.

Cantor, D. and Land, K. C. (1985). Unemployment and crime rates in the post-world war II United States: A theoretical and empirical analysis. *American Sociological Review*, 50(3):317–332.

Carroll, C. D. and Summers, L. H. (1991). Consumption growth parallels income growth: Some new evidence. In Bernheim, D. and Shoven, J. B., editors, *National Saving and Economic Performance*, pages 305–348. University of Chicago Press, Chicago, IL.

Cartwright, N. (2007). *Hunting Causes and Using Them*. Cambridge University Press, Cambridge.

Cartwright, N. (2020). Middle-range theory: Without it what could anyone do? *Theoria*, 35(3):269–323.

Cartwright, N. (2021). What counts as evidence for a mechanism? And why? *The Reasoner*, 15(6):53–54.

Cartwright, N. and Hardie, J. (2012). *Evidence-based Policy: A Practical Guide to Doing it Better*. Oxford University Press, Oxford.

Caswell v Powell Duffryn Associated Collieries Ltd (1940). *AC 152. House of Lords*. https://www.iclr.co.uk/document/1941000095/casereport_59740/html.

Chamlin, M. B., Myer, A. J., Sanders, B. A., and Cochran, J. K. (2008). Abortion as crime control: A cautionary tale. *Criminal Justice Policy Review*, 19:135–152.

Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. *Educational Researcher*, 21(6):13–17.

Chilisa, B. (2012). *Indigenous Research Methodologies*. Sage Publications, Inc., Thousand Oaks, CA.

Clarke, B. (2011). *Causality in Medicine with Particular Reference to the Viral Causation of Cancers*. PhD thesis, Department of Science and Technology Studies, University College London, London. https://discovery.ucl.ac.uk/id/eprint/1440583/.

Clarke, B., Gillies, D., Illari, P., Russo, F., and Williamson, J. (2014). Mechanisms and the evidence hierarchy. *Topoi*, 33(2):339–360.

Claveau, F. (2012). The Russo-Williamson theses in the social sciences: Causal inference drawing on two types of evidence. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(4):806–813.

Comanor, W. S. and Phillips, L. (2002). The impact on income and family structure on delinquency. *Journal of Applied Economics*, 5(2):209–232.

Comte, A. (1830). *Cours de Philosophie Positive*. Baillière, Paris.

Cook, T. D. and Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Rand McNally, Chicago, IL.

Cooley, T. F. and LeRoy, S. F. (1985). Atheoretical macroeconometrics. *Journal of Monetary Economics*, 16:283–308.

Copeland, D. C. (2015). *Economic Interdependence and War*. Princeton University Press, Princeton, NJ.

Corsaro, W. and Heise, D. (1990). Event structure models from ethnographic data. *Sociological Methodology*, 20:1–57.

Cox, D. R. (1992). Causality: Some statistical aspects. *Journal of the Royal Statistical Society, Series A*, 155:291–301.

Crasnow, S. (2011). Evidence for use: Causal pluralism and the role of case studies in political science research. *Philosophy of the Social Sciences*, 41(1):26–49.

Crasnow, S. (2019). Political science methodology: A plea for pluralism. *Studies in History and Philosophy of Science*, 78:40–47.

Creswell, J. W. (2012). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage, Thousand Oaks, CA.

Creswell, J. W. and Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research*. Sage, London.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., and Hanson, W. E. (2003). Advanced mixed methods research designs. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 209–240. Sage Publications, Inc., Thousand Oaks, CA, 1 edition.

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., and Stetler, C. (2012). Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical Care*, 50:217–226.

Cutler, D. M., Lleras-Muney, A., and Vogl, T. (2011). Socioeconomic status and health: Dimensions and mechanisms. In Glied, S. and Smith, P. C., editors, *The Oxford Handbook of Health Economics*, pages 124–163. Oxford University Press, Oxford.

Deaton, A. (2010). Understanding the mechanisms of economic development. *Journal of Economic Perspectives*, 24(3):3–16.

Deaton, A. and Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine (1982)*, 210:2–21.

Dills, A. K., Miron, J. A., and Summers, G. (2010). *What Do Economists Know about Crime?*, pages 269–302. University of Chicago Press, Chicago, IL.

Dilthey, W. (1883). *Einleitung in die Geisteswissenschaften*. Duncker and Humblot, Leipzig.

Dolton, P. and O'Neill, D. (1996). Unemployment and the restart effect: Some experimental evidence. *Economic Journal*, 106(435):387–400.

Donohue, J. J. and Levitt, S. D. (2001). The impact of legalized abortion on crime. *The Quarterly Journal of Economics*, 116(2):379–420.

Donohue, J. J. and Levitt, S. D. (2019). The impact of legalized abortion on crime over the last two decades. *Working Paper 25863*, National Bureau of Economic Research.

Dutton, D. B. (1986). Social class, health, and illness. In Aiken, L. and Mechanic, D., editors, *Applications of Social Science to Clinical Medicine and Health Policy*, pages 31–62. Rutgers University Press, New Brunswick, NJ.

Elster, J. (1998). A plea for mechanisms. In Hedström, P. and Swedberg, R., editors, *Social Mechanisms: An Analytic Approach to Social Theory*, pages 45–73. Cambridge University Press, Cambridge.

Engle, R. E. and Hendry, D. F. (1993). Testing super exogeneity and invariance in regression models. *Journal of Econometrics*, 56:119–139.

Erman, J. and Härkönen, J. (2017). Parental separation and school performance among children of immigrant mothers in Sweden. *European Journal of Population*, 33:267–292.

Evans, G. W. and Saegert, S. (2000). Residential crowding in the context of inner city poverty. In Wapner, S., Demick, J., Yamamoto, T., and Minami, H., editors, *Theoretical Perspectives in Environment-Behavior Research: Underlying Assumptions, Research Problems, and Methodologies*, pages 247–267. Springer US, Boston, MA.

Fama, E. (1980). Banking in the theory of finance. *Journal of Monetary Economics*, 6(1):39–57.

Fay, B. (1975). *Social Theory and Political Practice*. Routledge, London.

Fearon, J. D. (1991). Counterfactuals and hypothesis testing in political science. *World Politics*, 43(2):169–195.

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4(1):6–16.

Felig, R. N., Jordan, J. A., Shepard, S. L., Courtney, E. P., Goldenberg, J. L., and Roberts, T.-A. (2022). When looking 'hot'means not feeling cold: Evidence that self-objectification inhibits feelings of being cold. *British Journal of Social Psychology*, 61(2): 455–470.

Fisher, A. M. and Nicholetti, G. (1993). Regression direction and weak exogeneity. *Journal of Monetary Economics*, 32:213–235.

Foote, C. L. and Goetz, C. F. (2008). The impact of legalized abortion on crime: Comment. *The Quarterly Journal of Economics*, 123:407–423.

Fraser v 301–52 Townhouse Corp. (2006). *NY Slip Op 51855(U)*. New York State Law Reporting Bureau. https://www.nycourts.gov/Reporter/3dseries/2006/2006_51855.htm.

Fraser v 301–52 Townhouse Corp. (2008). *NY Slip Op 10208*. New York State Law Reporting Bureau. https://www.nycourts.gov/Reporter/3dseries/2008/2008_10208.htm.

Freedman, D. A. (1991). Statistical analysis and shoe leather. *Sociological Methodology*, 21:291–313.

Garrison, J. W. (1986). Some principles of postpositivistic philosophy of science. *Educational Researcher*, 15:12–18.

George, A. L. and Bennett, A. (2005). *Case Studies and Theory Development in the Social Sciences*. The MIT Press, Cambridge, MA.

Gerring, J. (2005). A unified framework for the social sciences. *Journal of Theoretical Politics*, 17(2):163–198.

Gerring, J. (2006). Single-outcome studies: A methodological primer. *International Sociology*, 5(21):707–734.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. (2011). *Impact Evaluation in Practice*. The World Bank Group, Washington, DC, 2 (2016) edition.

Ghiara, V. (2019). *Inferring Causation from Big Data in the Social Sciences*. PhD thesis, Department of Philosophy, University of Kent. https://kar.kent.ac.uk/75419/.

Gillies, D. (2000). *Philosophical Theories of Probability*. Routledge, London and New York.

Gillies, D. (2011). The Russo-Williamson thesis and the question of whether smoking causes heart disease. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the Sciences*, pages 110–125. Oxford University Press, Oxford.

Gillies, D. (2019). *Causality, Probability, and Medicine*. Routledge, Abingdon.

Glennan, S. (2017). *The New Mechanical Philosophy*. Oxford University Press, Oxford.

Glynn, L. (2011). A probabilistic analysis of causation. *British Journal for the Philosophy of Science*, 62:343–392.

Goertz, G. and Haggard, S. (2022). Large-*N* qualitative analysis (LNQA): External validity and generalization in case study and multi-method research. *Under Review*.

Goertz, G. and Haggard, S. (2023). Generalization, case studies, and within-case causal inferences: Large-*N* qualitative analysis (LNQA). In Kincaid, H. and van Bouwel, J., editors, *The Oxford Handbook of Philosophy of Political Science, pages 281-304.* Oxford University Press, Oxford.

Goertz, G. and Mahoney, J. (2012). *A Tale of Two Cultures*. Princeton University Press, Princeton and Oxford.

Goldstone, J. A. (1997). Methodological issues in comparative macrosociology. *Comparative Social Research*, 16:107–120.

Goldthorpe, J. H. (1996). The quantitative analysis of large-scale data-sets and rational action theory: For a sociological alliance. *European Sociological Review*, 12(2):109–126.

Goldthorpe, J. H. (1998). Rational action theory for sociology. *The British Journal of Sociology*, 49(2):167–192.

Goldthorpe, J. H. (2001). Causation, statistics, and sociology. *European Sociological Review*, 17(1):1–20.

Greenberg, D. F. (2001). Time series analysis of crime rates. *Journal of Quantitative Criminology*, 17(4):291–327.

Greene, J. C. (2006). Towards a methodology of mixed methods social inquiry. *Research in the Schools*, 13(1):93–98.

Greene, J. C. (2007). *Mixed Methods in Social Inquiry*. Jossey-Bass, San Francisco, CA.

Greene, J. C., Caracelli, V. J., and Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3):255–274.

Greene, J. C. and Hall, J. N. (2010). Dialectics and pragmatism: Being of consequences. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 119–144. Sage Publications, Inc., Thousand Oaks, CA, 2 edition.

Greenhalgh, T., Fisman, D., Cane, D. J., Oliver, M., and Macintyre, C. R. (2022). Adapt or die: How the pandemic made the shift from EBM to EBM+ more urgent. *BMJ Evidence Based Medicine*, 27(5):253–260.

Griffin, L. (1993). Narrative, event-structure analysis, and causal interpretation in historical sociology. *American Journal of Sociology*, 98:1094–1133.

Griffin, L. and Isaac, L. W. (1992). Recursive regression and the historical use of 'time' in time-series analysis of historical process. *Historical Methods*, 25:166–179.

Grüne-Yanoff, T. (2016). Why behavioural policy needs mechanistic evidence. *Economics and Philosophy*, 32(3):463–483.

Guba, E. G. (1990). The alternative paradigm dialog. In Guba, E. G., editor, *The Paradigm Dialog*, pages 17–29. Sage, Newbury Park, CA.

Guba, E. G. and Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Sage, Newbury Park, CA.

Guba, E. G. and Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In Denzin, N. K. and Lincoln, Y. S., editors, *Handbook of Qualitative Research*, pages 191–215. Sage, Thousand Oaks, CA.

Guyatt, G., Cairns, J., Churchill, D., Cook, D., Haynes, B., Hirsh, J., Irvine, J., Levine, M., Levine, M., Nishikawa, J., Sackett, D., Brill-Edwards, P., Gerstein, H., Gibson, J., Jaeschke, R., Kerigan, A., Neville, A., Panju, A., Detsky, A., Enkin, M., Frid, P., Gerrity, M., Laupacis, A., Lawrence, V., Menard, J., Moyer, V., Mulrow, C., Links, P., Oxman, A., Sinclair, J., and Tugwell, P. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA*, 268(17):2420–2425.

Guyatt, G., Rennie, D., Meade, M. O., and Cook, D. J., editors (2015). *Users' Guides to the Medical Literature: Essentials of Evidence-based Clinical Practice*. McGraw Hill, New York, 3 edition.

Haggard, S. and Kaufman, R. R. (2016). *Dictators and Democrats*. Princeton University Press, Princeton and Oxford.

Hall, N. (2004). Two concepts of causation. In Collins, J., Hall, N., and Paul, L., editors, *Causation and Counterfactuals*, pages 225–276. MIT Press, Cambridge MA and London.

Hampden-Thompson, G. (2013). Family policy, family structure, and children's educational achievement. *Social Science Research*, 42(3):804–817.

Hart, H. L. A. and Honoré, T. (1985). *Causation in the Law*. Oxford University Press, Oxford, 2 edition.

Harvey, F. P. (2011). *Explaining the Iraq War*. Cambridge University Press, Cambridge.

Hausman, D. M. (2010). Probabilistic causality and causal generalizations. In Eells, E. and Fetzer, J. H., editors, *The Place of Probability in Science*, pages 47–63. Springer, Dordrecht.

Haveman, R. and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 33(4):1829–1878.

Hedström, P. and Bearman, P. (2011). What is analytic sociology all about? An introductory essay. In Bearman, P. and Hedström, P., editors, *The Oxford Handbook of Analytical Sociology*, pages 3–24. Oxford University Press, Oxford.

Hedström, P. and Swedberg, R. (1998). Social mechanism: An introductory essay. In Hedström, P. and Swedberg, R., editors, *Social Mechanisms: An Analytic Approach to Social Theory*, pages 1–31. Cambridge University Press, Cambridge.

Hedström, P. and Ylikoski, P. (2014). Analytic Sociology. In Jarvie, I. C. and Zamora-Bonilla, J., editors, *The Sage Handbook of the Philosophy of Social Sciences*, pages 386–398. Sage Publications Ltd, London.

Hesse-Biber, S. N. (2010). *Mixed Methods Research: Merging Theory with Practice*. The Guilford Press, New York and London.

Hesse-Biber, S. N. (2015). Feminist approaches to mixed methods research: Linking theory and praxis. In Tashakkori, A. and Teddlie, C., editors, *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, pages 169–192. Sage, Thousand Oaks, CA, 2 edition.

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., and Welch, V., editors (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons, Chichester, 2 edition.

Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *The Journal of Philosophy*, 98(6):273–299.

HM Treasury (2011). *Magenta Book: Central Government Guidance on Evaluation*. HM Treasury and Evaluation Task Force, London, 2020 edition. https://www.gov.uk/government/publications/the-magenta-book.

Hojman, D. E. (2004). Inequality, unemployment and crime in Latin American cities. *Crime, Law and Social Change*, 41:33–51.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

Holmlund, H., Lindahl, M., and Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, 49(3):615–651.

Hoover, K. D. (1991). The causal direction between money and prices: An alternative approach. *Journal of Monetary Economics*, 27(3):381–423.

House, J. S., Lepkowski, J. M., Kinney, A. M., Mero, R. P., Kessler, R. C., and Herzog, A. R. (1994). The social stratification of aging and health. *Journal of Health and Social Behavior*, 35(3):213–234.

Howick, J. (2011a). Exposing the vanities—and a qualified defence—of mechanistic evidence in clinical decision-making. *Philosophy of Science*, 78(5):926–940. Proceedings of the Biennial PSA 2010.

Howick, J. (2011b). *The Philosophy of Evidence-based Medicine*. Wiley-Blackwell, Chichester.

Htun, M. and Weldon, S. L. (2012). The civic origins of progressive policy change: Combating violence against women in global perspective, 1975–2005. *American Political Science Review*, 106(3):548–569.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Oxford Philosophical Texts. Oxford University Press, Oxford, 1999 edition.

Hume, D. (1876a). Of money. In *Essays, Literary, Moral, and Political*, pages 167–175. Ward, Lock, and Tyler, Warwick House, London.

Hume, D. (1876b). Of the balance of trade. In *Essays, Literary, Moral, and Political*, pages 184–195. Ward, Lock, and Tyler, Warwick House, London.

Illari, P. M. (2011). Mechanistic evidence: Disambiguating the Russo-Williamson thesis. *International Studies in the Philosophy of Science*, 25(2):139–157.

Illari, P. M. and Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2:119–135.

Illsley, R. and Mullen, K. (1985). The health needs of disadvantaged client groups. In Holland, W. W., Detels, R., and Knox, G., editors, *Oxford Textbook of Public Health*, pages 389–402. Oxford University Press, Oxford.

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309–334.

Imrohoroğlu, A., Merlo, A., and Rupert, P. (2004). What accounts for the decline in crime? *International Economic Review*, 45(3):707–729.

Issac, L. W. and Griffin, L. (1989). Ahistoricism in time-series analysis of historical process: Critique, redirection and illustrations from U.S. labor history. *American Sociological Review*, 54:873–890.

Isaac, L. W., Street, D. A., and Knapp, S. J. (1994). Analyzing historical contingency with formal methods. *Sociological Methods and Research Research*, 23(1):114–141.

Ivankova, N. V. and Stick, S. L. (2007). Students' persistence in a distributed doctoral program in educational leadership in higher education: A mixed methods study. *Research in Higher Education*, 48(1):93–135.

Jacob, M., Wentland, K., and Wentland, S. A. (2022). Real effects of tax uncertainty: Evidence from firm capital investments. *Management Science*, 68(6):4065–4089.

Jagosh, J., Tilley, N., and Stern, E. (2016). Realist evaluation at 25: Cumulating knowledge, advancing debates and innovating methods. *Evaluation*, 22(3):267–269.

Jiménez-Buedo, M. and Squitieri, J. C. (2019). What can mechanisms do for you? Mechanisms and the problem of confounders in the social sciences. *Philosophy of the Social Sciences*, 49(3):210–231.

John, P. (2010). Quantitative methods. In Marsh, D. and Stoker, G., editors, *Theory and Methods in Political Science*, pages 267–284. Palgrave Macmillan, Basingstoke, 3 edition.

Johnson, R. B. (2017). Dialectical pluralism: A metaparadigm whose time has come. *Journal of Mixed Methods Research*, 11(2):156–173.

Johnson, R. B. and Christensen, L. (2014). *Educational Research*. Sage Publications, Inc., Thousand Oaks, CA, 5 edition.

Johnson, R. B., de Waal, C., Stefurak, T., and Hildebrand, D. L. (2017). Understanding the philosophical positions of classical and neopragmatists for mixed methods research. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69:63–86.

Johnson, R. B. and Gray, R. (2010). A history of philosophical and theoretical issues for mixed methods research. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 69–94. Sage Publications, Inc., Thousand Oaks, CA, 2 edition.

Johnson, R. B. and Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Research*, 33(7):14–26.

Johnson, R. B., Onwuegbuzie, A. J., and Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2):112–133.

Johnson, S. D., Tilley, N., and Bowers, K. J. (2015). Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology*, 11(3):459–473.

Johnson, T. and Klepinger, D. (1994). Experimental evidence on unemployment insurance work-search policies. *Journal of Human Resources*, 29(3):695–717.

Jonsson, J. O. and Gähler, M. (1997). Family dissolution, family reconstitution, and children's educational careers: Recent evidence for Sweden. *Demography*, 34(2):277–293.

Joyce, T. (2004). Further tests of abortion and crime. *Working Paper 10564*, National Bureau of Economic Research.

Kelle, U. (2015). Mixed methods and the problem of theory building and theory testing in the social sciences. In Hesse-Biber, S. N. and Johnson, R. B., editors, *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, pages 594–605. Oxford University Press, Oxford.

Kellstedt, P. M. and Whitten, G. D. (2018). *The Fundamentals of Political Science Research*. Cambridge University Press, Cambridge, 3 edition.

Kiser, E. and Hechter, M. (1991). The role of general theory in comparative-historical sociology. *American Journal of Sociology*, 97:1–30.

Klein, D. M. (1987). Causation in sociology today: A revised view. *Sociological Theory*, 5(1):19–26.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press, Chicago, IL.

Lalive, R., Zweimüller, J., and van Ours, J. C. (2005). The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association*, 3(6):1386–1417.

Landes, S. J., McBain, S. A., and Curran, G. M. (2019). An introduction to effectiveness-simplementation hybrid designs. *Psychiatry Research*, 280:112513.

Lantz, P. M., House, J. S., Lepkowski, J. M., Williams, D. R., Mero, R. P., and Chen, J. (1998). Socioeconomic factors, health behaviors, and mortality: Results from a nationally representative prospective study of US adults. *Journal of the American Medical Association*, 279(21):1703–1708.

Lee, C.-C., Wu, H.-J., and Chiou, W.-B. (2022a). Reminders of COVID-19 social distancing can intensify physical pain. *British Journal of Social Psychology*, 61(2):587–598.

Lee, S. M.-K., Cui, Y., and Tong, S. X. (2022b). Toward a model of statistical learning and reading: Evidence from a meta-analysis. *Review of Educational Research*, 92(4):651–691.

Levack-Payne, W. (2022a). *The Application of Evidence-Based Medicine Methodologies in Sports Science: Problems and Solutions*. Phd thesis, Philosophy Department, University of Kent. https://kar.kent.ac.uk/93560/.

Levack-Payne, W. (2022b) . Mechanistic evidence and exercise interventions: Causal claims, extrapolation, and implementation. *Journal of Evaluation in Clinical Practice*, 28(5):745–751.

Levine, P. B., Staiger, D., Kane, T. J., and Zimmerman, D. J. (1999). Roe v. Wade and American Fertility. *American Journal of Public Health*, 89(2):199–203.

Lewis, A. (1954). Economic development with unlimited supplies of labor. *Manchester School of Economics and Social Studies*, 33(2):160–217.

Lewis, D. K. (1973). Causation. *Journal of Philosophy*, 70:556–567.

Lieberman, E. S. (2005). Nested analysis as a mixed-method strategy for comparative research. *American Political Science Review*, 99(3):435–452.

Lincoln, Y. S. and Guba, E. G. (1985). *Naturalistic Inquiry*. Sage, Thousand Oaks, CA.

Lindsay-Smith, G., O'Sullivan, G., Eime, R., Harvey, J., and Van Uffelen, J. G. (2018). A mixed methods case study exploring the impact of membership of a multiactivity, multicentre community group on social wellbeing of older adults. *BMC Geriatrics*, 18(1):1–14.

Link, B. G. and Phelan, J. C. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior* (Extra issue):80–94.

Little, D. (1991). *Varieties of Social Explanations*. Westview Press, Boulder, CO.

Longworth, F. (2006). Causation, pluralism and responsibility. *Philosophica*, 77:45–68.

Lott, J. R. and Whitley, J. (2007). Abortion and crime: Unwanted children and out-of-wedlock births. *Economic Inquiry*, 45:304–324.

Lutfey, K. and Freese, J. (2005). Toward some fundamentals of fundamental causality: Socioeconomic status and health in the routine clinic visit for diabetes. *American Journal of Sociology*, 110(5):1326–1372.

Machamer, P., Darden, L., and Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67:1–25.

MacIntyre, A. (1976). Causality and history. In Manninen, J. and Tuomela, R., editors, *Essays on Explanation and Understanding: Studies in the Foundations of Humanities and Social Sciences*, pages 137–158. D. Reidel, Dordrecht.

Mahoney, J. (2001). Review: Beyond correlational analysis: Recent innovations in theory and method. Reviewed work(s): Social mechanisms: An analytical approach to social theory by Peter Hedström and Richard Swedberg; Fuzzy-set social science by Charles C. Ragin. *Sociological Forum*, 16(3):575–593.

Mahoney, J. (2012). The logic of process tracing tests in the social sciences. *Sociological Methods & Research*, 41(4):570–597.

Mahoney, J. and Barrenechea, R. (2019). The logic of counterfactual analysis in case-study explanation. *The British Journal of Sociology*, 70(1):306–338.

Marchionni, C. (2018). Mechanisms in economics. In Glennan, S. and Illari, P., editors, *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, pages 423–434. Routledge, London and New York.

Marchionni, C. and Reijula, S. (2019). What is mechanistic evidence, and why do we need it for evidence-based policy? *Studies in History and Philosophy of Science Part A*, 73:54–63.

Matthews, L. J., Nowak, S. A., Gidengil, C. C., Chen, C., Stubbersfield, J. M., Tehrani, J. J., and Parker, A. M. (2022). Belief correlations with parental vaccine hesitancy: Results from a national survey. *American Anthropologist*, 124(2):291–306.

Maxwell, J. A. (2011). Paradigms or Toolkits? Philosophical and methodological positions as heuristics for mixed methods research. *Mid-Western Educational Researcher*, 24(2):27–30.

Maxwell, J. A. and Mittapalli, K. (2010). Realism as a stance for mixed methods research. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 145–168. Sage Publications, Inc., Thousand Oaks, CA, 2 edition.

Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *The Canadian Journal of Program Evaluation*, 16(1):1–24.

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, 18(3):270–280.

Mayne, J. (2019). Revisiting contribution analysis. *The Canadian Journal of Program Analysis*, 34(2):171–191.

Maziarz, M. (2019). Methodological pluralism in economics: The 'why' and 'how' of causal inferences. *Filozofia Nauki*, 27(4):43–59.

Maziarz, M. (2020). *The Philosophy of Causality in Economics*. Routledge, London and New York.

Maziarz, M. (2021). Resolving empirical controversies with mechanistic evidence. *Synthese*, 199:9957–9978.

Menzies, P. (2007). Causation in context. In Price, H. and Corry, R., editors, *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, pages 191–223. Oxford University Press, Oxford.

Menzies, P. and Price, H. (1993). Causation as a secondary quality. *British Journal for the Philosophy of Science*, 44:187–203.

Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 135–164. Sage, Thousand Oaks, CA, 1 edition.

Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research*, 1(3):212–225.

Mertens, D. M. (2010). Transformative mixed methods research. *Qualitative Inquiry*, 16(6):469–474.

Mertens, D. M., Bledsoe, K. L., Sullivan, M., and Wilson, A. (2010). Utilization of mixed methods for transformative purposes. In Tashakkori, A. and Teddlie, C., editors, *Sage Handbook of Mixed Methods in Social and Behavioral Research*, pages 193–214. Sage Publications, Inc., Thousand Oaks, CA, 2 edition.

Merton, R. K. (1968). *Social Theory and Social Structure*. The Free Press, New York.

Michie, S., Fixsen, D., Grimshaw, J. M., and Eccles, M. P. (2009). Specifying and reporting complex behaviour change interventions: The need for a scientific method. *Implementation Science*, 4:40.

Minter, W. (1989). *The Mozambican National Resistance (Renamo) as Described by Exparticipants*. Technical report, Ford Foundation and Swedish International Development Agency, Washington, DC.

Modigliani, F. (1970). The life-cycle hypothesis and intercountry differences in the saving ratio. In Elltis, W. A., Scott, M. F., and Wolfe, J. N., editors, *Induction, Growth, and Trade: Essays in Honour of Sir Roy Harrod*, pages 197–225. Oxford University Press, Oxford.

Modigliani, F. and Brumberg, R. H. (1954). Utility analysis and the consumption function: An interpretation of the cross-section data. In Kurihara, K. K., editor, *Post-Keynesian Economics*, pages 388–436. Rutgers University Press, New Brunswick.

Modigliani, F. and Brumberg, R. H. (1980). Utility analysis and aggregate consumption functions: An attempt at integration. In Abel, A., editor, *The Collected Papers of Franco Modiliani, Volume 2: The Life Cycle Hypothesis of Saving*, pages 128–197. MIT Press, Cambridge, MA.

Moneta, A. and Russo, F. (2014). Causal models and evidential pluralism in econometrics. *Journal of Economic Methodology*, 21(1):54–76.

Moore, M. (2011). Causation in the criminal law. In Deigh, J. and Dolinko, D., editors, *The Oxford Handbook of Philosophy of Criminal Law*, pages 168–193. Oxford University Press, Oxford.

Moore, M. (2019). Causation in the law. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy (Winter 2019 Edition)*, https://plato.stanford.edu/archives/win2019/entries/causation-law/.

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry*, 20(8):1045–1053.

Morgan, S. L., Leenman, T. S., Todd, J. J., and Weeden, K. A. (2013). Occupational plans, beliefs about educational requirements, and patterns of college entry. *Sociology of Education*, 86(3):197–217.

Morgan, S. L. and Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, Cambridge.

Morgan, S. L. and Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, Cambridge, 2 edition.

Murnane, R. J. and Willett, J. B. (2011). *Methods Mattter*. Oxford University Press, Oxford and New York.

Nolan, D. (2013). Why historians (and everyone else) should care about counterfactuals. *Philosophical Studies*, 163(2):317–335.

Noyes, J., Booth, A., Cargo, M., Flemming, K., Harden, A., Harris, J., Garside, R., Hannes, K., Pantoja, T., and Thomas, J. (2019). Qualitative evidence. In Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., and Welch, V., editors, *Cochrane Handbook for Systematic Reviews of Interventions*, chapter 21, pages 525–546. John Wiley & Sons, Chichester, 2 edition.

OCEBM Levels of Evidence Working Group (2011). *The Oxford 2011 Levels of Evidence*. Oxford Centre for Evidence-Based Medicine. https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Instuituions for Collective Action*. Cambridge University Press, New York.

Pampel, F. C., Krueger, P. M., and Denney, J. T. (2010). Socioeconomic disparities in health behaviors. *Annual Review of Sociology*, 36:349–370.

Parkkinen, V.-P. (2016). Robustness and evidence of mechanisms in early experimental atherosclerosis research. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 60:44–55.

Parkkinen, V.-P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B., and Williamson, J. (2018). *Evaluating Evidence of Mechanisms in Medicine: Principles and Procedures*. Springer, Cham.

Parkkinen, V.-P. and Williamson, J. (2020). Extrapolating from model organisms in pharmacology. In La Caze, A. and Osimani, B., editors, *Uncertainty in Pharmacology: Epistemology, Methods, and Decisions*, pages 59–78. Springer, Dordrecht.

Patterson, F., Benowitz, N., Shields, P., Kaufmann, V., Jepson, C., Wileyto, P., Kucharski, S., and Lerman, C. (2003). Individual differences in nicotine intake per cigarette. *Cancer Epidemiology Biomarkers and Prevention*, 12(5):468–471.

Pawson, R. and Tilley, N. (1997). *Realistic Evaluation*. Sage, London.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Pearson, K. (1898). Mathematical contributions to the theory of evolution. On the law of ancestral heredity. *Proceedings of the Royal Society of London*, 62:386–412.

Pearson, K. (1900). *The Grammar of Science*. Adam and Charles Black, London, 2 edition.

Pence, C. H. (2011). "Describing our whole experience": The statistical philosophies of W. F. R. Weldon and Karl Pearson. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42(4):475–485.

Pérez-González, S. and Rocca, E. (2022). Evidence of biological mechanisms and health predictions: An insight into clinical reasoning. *Perspectives in Biology and Medicine*, 65(1):89–105.

Phelan, J. C., Link, B. G., Diez-Roux, A., Kawachi, I., and Levin, B. (2004). "Fundamental causes" of social inequalities in mortality: A test of the theory. *Journal of Health and Social Behavior*, 45:265–285.

Phelan, J. C., Link, B. G., and Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: Theory, evidence, and policy implications. *Journal of Health and Social Behavior*, 51(Suppl):S28–S40.

Phillips, D. C. (1990). Postpositivist science: Myths and realities. In Guba, E. G., editor, *The Paradigm Dialog*, pages 31–45. Sage, Thousand Oaks, CA.

Phillips, D. C. and Burbules, N. C. (2000). *Postpositivism and Educational Research*. Rowman and Littlefield, Lanham, MD.

Pole, C. J. and Lampard, R. (2002). *Practical Social Investigation*. Prentice Hall, Harlow.

Pollack, F. (1901). *The Law of Torts*. Stevens and Sons, London.

Putnam, R. D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ.

Quetelet, A. (1835). *Sur L'Homme Et Le Developpement de Ses Facultes, Ou Essai de Physique Sociale*, volume 1. Bachelier, Paris.

Raphael, S. and Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. *The Journal of Law and Economics*, 44(1):259–283.

Reichenbach, H. (1956). *The Direction of Time*. University of California Press, Berkeley and Los Angeles, 1971 edition.

Reiss, J. (2009). Causation in the social sciences: Evidence, inference, and purpose. *Philosophy of the Social Sciences*, 39:20–40.

Reiss, J. (2011). Third time's a charm: Causation, science and Wittgensteinian pluralism. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the Sciences*, pages 907–927. Oxford University Press, Oxford.

Reiss, J. (2012). Causation in the sciences: An inferentialist account. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(4):769–777. Causality in the Biomedical and Social Sciences.

Reiss, J. (2013). *Philosophy of Economics: A Contemporary Introduction*. Routledge, New York and London.

Road Safety Act (2006). https://www.legislation.gov.uk/ukpga/2006/49/contents.

Rohlfing, I. and Zuber, C. I. (2021). Check your truth conditions! clarifying the relationship between theories of causation and social science methods for causal inference. *Sociological Methods and Research*, 50(4):1623–1659.

Rothman, K. (1986). *Modern Epidemiology*. Little, Brown, and Company, Boston, MA.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701.

Runhardt, R. W. (2022). Limits to evidential pluralism: Multi-method large-N qualitative analysis and the primacy of mechanistic studies. *Synthese*, 200:171.

Russo, F. and Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2):157–170.

Russo, F. and Williamson, J. (2011). Generic versus single-case causality: The case of autopsy. *European Journal for Philosophy of Science*, 1:47–69.

Russo, F., Wunsch, G., and Mouchart, M. (2019). Causality in the social sciences: A structural modelling framework. *Quality & Quantity*, 53(5):2575–2588.

R v Hughes (2013). *UKSC 56*. United Kingdom Supreme Court. https://www.supremecourt.uk/cases/docs/uksc-2011-0240-judgment.pdf.

R v Pagett (1983). *76 Cr App R 279*. England and Wales Court of Appeal (Criminal Division). https://www.bailii.org/ew/cases/EWCA/Crim/1983/1.html.

R v Williams (2010). *EWCA Crim 2552*. England and Wales Court of Appeal (Criminal Division). https://www.bailii.org/ew/cases/EWCA/Crim/2010/2552.html.

Ryan v New York Central Railroad (1866). *35 N.Y. 210, 1866 N.Y. 86*. Court of Appeals of the State of New York. https://casetext.com/case/ryan-v-new-york-central-railroad.

Ryu, P. K. (1958). Causation in criminal law. *University of Pennsylvania Law Review*, 106(6):773–805.

Salmon, W. C. (1998). *Causality and Explanation*. Oxford University Press, Oxford.

Samet, J. M., Chiu, W. A., Cogliano, V., Jinot, J., Kriebel, D., Lunn, R. M., Beland, F. A., Bero, L., Browne, P., Fritschi, L., Kanno, J., Lachenmeier, D. W., Lan, Q., Lasfargues, G., Curieux, F. L., Peters, S., Shubat, P., Sone, H., White, M. C., Williamson, J., Yakubovskaya, M., Siemiatycki, J., White, P. A., Guyton, K. Z., Schubauer-Berigan, M. K., Hall, A. L., Grosse, Y., Bouvard, V., Benbrahim-Tallaa, L., Ghissassi, F. E., Lauby-Secretan, B., Armstrong, B., Saracci, R., Zavadil, J., Straif, K., and Wild, C. P. (2020). The IARC monographs: Updated procedures for modern and transparent evidence synthesis in cancer hazard identification. *JNCI: Journal of the National Cancer Institute*, 112(1):1–8.

Sandefur, G. D. and Wells, T. (1999). Does family structure really influence educational attainment? *Social Science Research*, 28:331–357.

Schaffer, J. (2005). Contrastive causation. *The Philosophical Review*, 114(3):327–358.

Schaffer, J. (2010). Contrastive causation in the law. *Legal Theory*, 16:259–297.

Schelling, T. C. (1998). Social mechanisms and social dynamics. In Hedström, P. and Swedberg, R., editors, *Social Mechanisms: An Analytic Approach to Social Theory*, pages 32–44. Cambridge University Press, Cambridge.

Schmidt, M. B. (2003). Money and prices: Evidence from the G7 countries. *Applied Economics*, 35:1799–1809.

Schnall, P. L., Pieper, C., Schwartz, J. E., Karasek, R. A., Schlussel, Y., Devereux, R. B., Ganau, A., Alderman, M., Warren, K., and Pickering, T. G. (1990). The relationship between 'job strain,' workplace diastolic blood pressure, and left ventricular mass index. Results of a case-control study. *Journal of the American Medical Association*, 263(14): 1929–1935.

Schoonenboom, J. (2019). A performative paradigm for mixed methods research. *Journal of Mixed Methods Research*, 13(3):284–300.

Schünemann, H., Brożek, J., Guyatt, G., and Oxman, A., editors (2013). *GRADE Handbook*. Grading of Recommendations, Assessment, Development and Evaluation Working Group. https://gdt.gradepro.org/app/handbook/handbook.html.

Sewell, W. H., Haller, A. O., and Ohlendorf, G. W. (1970). The educational and early occupational status attainment process: Replication and revision. *American Sociological Review*, 35(6):1014–1027.

Sewell, W. H., Haller, A. O., and Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 34(1):82–92.

Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and QuasiExperimental Designs for Generalized Causal Inference*. Houghton, Mifflin and Company, Boston, MA.

Shan, Y. (2020). *Doing Integrated History and Philosophy of Science: A Case Study of the Origin of Genetics*. Springer, Cham.

Shan, Y. and Williamson, J. (2021). Applying evidential pluralism to the social sciences. *European Journal for the Philosophy of Science*, 11(4):96.

Simon, H. A. and Iwasaki, Y. (1988). Causal ordering, comparative statistics, and near decomposability. *Journal of Econometrics*, 39:149–173.

Skocpol, T. (1979). *States and Social Revolutions: A Comparative Analysis of France, Russia and China*. Cambridge University Press, Cambridge.

Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher*, 12(3):6–13.

Sørensen, A. B. (1998). Theoretical mechanisms and the empirical study of social processes. In Hedström, P. and Swedberg, R., editors, *Social Mechanisms: An Analytic Approach to Social Theory*, pages 238–239. Cambridge University Press, Cambridge.

Stapleton, J. (2008). Choosing what we mean by causation in the law. *Missouri Law Review*, 73(2):433–480.

Stapleton, J. (2009). Causation in the law. In Beebee, H., Hitchcock, C., and Menzies, P., editors, *The Oxford Handbook of Causation*, pages 744–769. Oxford University Press, Oxford.

Stapleton, J. (2015). An "extended but-for" test for the causal relation in the law of obligations. *Oxford Journal of Legal Studies*, 35(4):697–726.

Steel, D. (2004). Social mechanisms and causal inference. *Philosophy of the Social Sciences*, 34(1):55–78.

Steel, D. (2008). *Across the Boundaries, Extrapolation in Biology and Social Science*. Oxford University Press, Oxford.

Stegenga, J. (2022). Evidence of effectiveness. *Studies in History and Philosophy of Science*, 91:288–295.

Steinberg, P. F. (2007). Causal assessment in small-n policy studies. *Policy Studies Journal*, 35:181–204.

Stepan, A. (1978). Political leadership and regime breakdown: Brazil. In Linz, J. J. and Stepan, A., editors, *The Breakdown of Democratic Regimes: Latin America*, pages 110–137. The John Hopkins University Press, Baltimore and London.

Stinchcombe, A. L. (1991). The conditions of fruitfulness of theorizing about mechanisms in social science. *Philosophy of the Social Sciences*, 21(3):367–388.

SUNY (2004). *SUNY Downstate Medical Center Evidence Based Medicine Tutorial*. The Medical Research Library of Brooklyn. https://guides.downstate.edu/c.php?g=856794&p=6137636.

Tashakkori, A. and Teddlie, C. (1998). *Mixed Methodology: Combining Qualitative and Quantitative Appraoches*. Sage, Thousand Oaks, CA.

Tebes, J. K. (2012). Philosophical foundations of mixed methods research: Implications for research practice. In Jason, L. A. and Glenwick, D. S., editors, *Methodological Approaches to Community-based Research*, pages 13–31. American Psychological Association, Washington, DC.

Teddlie, C. and Tashakkori, A. (2009). *Foundations of Mixed Methods Research*. Sage, Thousand Oaks, CA.

Teddlie, C. and Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. In Tashakkori, A. and Teddlie, C., editors, *SAGE Handbook of Mixed Methods in Social and Behavioral Research*, pages 1–42. Sage Publications, Inc., Thousand Oaks, CA, 2 edition.

Thomas, J., Kneale, D., McKenzie, J. E., Brennan, S. E., and Bhaumik, S. (2019). Determining the scope of the review and the questions it will address. In Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., and Welch, V., editors, *Cochrane Handbook for Systematic Reviews of Interventions*, chapter 2, pages 13–32. John Wiley & Sons, Chichester, 2 edition.

Thornton, A., Sidebottom, A., Belur, J., Tompson, L., and Bowers, K. (2019). On the development and application of EMMIE: Insights from the what works centre for crime reduction. *Policing and Society*, 29(3):266–282.

Tilley, N. (2016). EMMIE and engineering: What works as evidence to improve decisions? *Evaluation*, 22(3):304–322.

Tilly, C. (1997). Means and ends of comparison in macrosociology. *Comparative Social Research*, 16:43–53.

Timans, R., Wouters, P., and Heilbron, J. (2019). Mixed methods research: What it is and what it could be. *Theory and Society*, 48:193–216.

Tonelli, M. and Williamson, J. (2020). Mechanisms in clinical practice: Use and justification. *Medicine, Health Care and Philosophy*, 23(1):115–124.

Toshkov, D. (2016). *Research Design in Political Science*. Palgrave, London.

Turunen, J. (2014). Adolescent educational outcomes in blended families: Evidence from Swedish register data. *Journal of Divorce & Remarriage*, 55(7):568–589.

Twisk, J., Bosman, L., Hoekstra, T., Rijnhart, J., Welten, M., and Heymans, M. (2018). Different ways to estimate treatment effects in randomised controlled trials. *Contemporary Clinical Trials Communications*, 10:80–85.

United Nations (2013). Effective use of statistics in evidence-based policymaking: Conceptual framework. *Technical Report E/ESCWA/SD/2013/Technical Paper.1*. New York. https://archive.unescwa.org/sites/www.unescwa.org/files/publications/files/e_escwa_sd_13_tp-1_e.pdf.

Vaidyanathan, B., Strand, M., Choi-Fitzpatrick, A., Buschman, T., Davis, M., and Varela, A. (2016). Causality in contemporary American sociology: An empirical assessment and critique. *Journal for the Theory of Social Behaviour*, 46(1):3–26.

van den Berg, G. J., van der Klaauw, B., and van Ours, J. C. (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, 22(1):211–241.

van Grootel, L., Nair, L. B., Klugkist, I., and van Wesel, F. (2020). Quantitizing findings from qualitative studies for integration in mixed methods reviewing. *Research Synthesis Methods*, 11(3):413–425.

Weber, E. (2007). Social mechanisms, causal inference, and the policy relevance of social science. *Philosophy of the Social Sciences*, 30(3):348–359.

Weber, E. (2009). How probabilistic causation can account for the use of mechanistic evidence. *International Studies in the Philosophy of Science*, 23(3):277–295.

Weber, M. (1904). Die "Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für Sozialwissenschaft und Sozialpolitik*, 19(1):22–87.

Weinstein, J. M. (2007). *Inside Rebellion: The Politics of Insurgent Violence*. Cambridge University Press, Cambridge.

Weldon, W. F. R. (1896). The utility of specific characters. *Nature*, 54(1396):294–295.

Weldon, W. F. R. (1904–1905). *Theory of Inheritance*. UCL Library, London.

Weldon, W. F. R., Pearson, K., and Davenport, C. B. (1901). Editorial: The spirit of Biometrika. *Biometrika*, 1(1):3–6.

Westland, J. C. (2015). *Structural Equation Models: From Paths to Networks*. Springer, Cham, 2 (2019) edition.

What Works (2018a). *The Rise of Experimental Government: Cross-government Trial Advice Panel Update Report*. Produced by the What Works Network and the UK Economic and Social Research Council. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/753468/RiseExperimentalGovernment_CrossGovTrialAdvicePanelUpdateReport.pdf.

What Works (2018b). *The What Works Network: Five Years On*. Produced by the What Works Network and the UK Economic and Social Research Council. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/677478/6.4154_What_works_report_Final.pdf.

Wieten, S. (2018). *What Counts as 'What Works': Expertise, Mechanisms and Values in Evidence-based Medicine*. PhD thesis, University of Durham. http://etheses.dur.ac.uk/12606/.

Wilde, M. (2021). Mechanistic reasoning and the problem of masking. *Synthese*, 199(3):6103–6118.

Wilde, M. and Parkkinen, V.-P. (2019). Extrapolation and the Russo–Williamson thesis. *Synthese*, 196(8):3251–3262.

Williams, G. (1961). Causation in the law. *The Cambridge Law Journal*, 19(1):62–85.

Williamson, J. (2005). *Bayesian Nets and Causality: Philosophical and Computational Foun dations*. Oxford University Press, Oxford.

Williamson, J. (2009). Probabilistic theories. In Beebee, H., Hitchcock, C., and Menzies, P., editors, *The Oxford Handbook of Causation*, pages 185–212. Oxford University Press, Oxford.

Williamson, J. (2011). Mechanistic theories of causality. *Philosophy Compass*, 6(6):421–447.

Williamson, J. (2013). How can causal explanations explain? *Erkenntnis*, 78:257–275.

Williamson, J. (2015). Deliberation, judgement and the nature of evidence. *Economics and Philosophy*, 31(1):27–65.

Williamson, J. (2017). Models in systems medicine. *Disputatio*, 9(47):429–469.

Williamson, J. (2019a). Establishing causal claims in medicine. *International Studies in the Philosophy of Science*, 32(2):33–61.

Williamson, J. (2019b). Evidential proximity, independence, and the evaluation of carcinogenicity. *Journal of Evaluation in Clinical Practice*, 25(6):955–961.

Williamson, J. (2021a). Establishing the teratogenicity of Zika and evaluating causal criteria. *Synthese*, 198(10):2505–2518.

Williamson, J. (2021b). Evidential Pluralism and explainable AI. *The Reasoner*, 15(6): 55–56.

Williamson, J. (2021c). The feasibility and malleability of EBM+. *THEORIA. An International Journal for Theory, History and Foundations of Science*, 36(2):191–209.

Williamson, J. (2022). A Bayesian account of establishing. *British Journal for the Philosophy of Science*, 73(4): 903-925.

Witjens, E. (2014). Considering causation in criminal law. *The Journal of Criminal Law*, 78(2):164–183.

Wood, E. J. (2003). *Insurgent Collective Action and Civil War in El Salvador*. Cambridge University Press, Cambridge.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.

Wright, R. W. (1985). Causation in tort law. *California Law Review*, 73:1735–1828.

Wright, R. W. (2011). The NESS account of natural causation: A response to criticisms. In Kahmen, B. and Stepanians, M., editors, *Critical Essays on "Causation and Responsibility"*, pages 285–322. de Gruyter, Berlin and Boston.

Wu, C. Y. H. (2022). Heat waves and road traffic collisions in Alabama, United States. *Annals of the American Association of Geographers*, 112(5):1313–1327.

Yamey, G. and Feachem, R. (2011). Evidence-based policymaking in global health—the payoffs and pitfalls. *Evidence Based Medicine*, 16:97–99.

Ye, C., Beyene, J., Browne, G., and Thabane, L. (2014). Estimating treatment effects in randomised controlled trials with non-compliance: A simulation study. *BMJ Open*, 4(6).

Zeratsion, H., Bjertness, C. B., Bjertness, E., Dalsklev, M., Haavet, O. R., Halvorsen, J. A., Lien, L., and Claussen, B. (2015). The influence of parental divorce on educational ambitions of 18/19 year-old adolescents from Oslo, Norway. *Journal of Child and Family Studies*, 24:2865–2873.

Index