

42

CAUSAL KNOWLEDGE AND THE
PROCESS OF POLICY MAKING

Toward a Bottom-Up Approach

*Luis Mireles-Flores***Key messages:**

- The widespread intuition that to attain effective policy outcomes all one needs is valid and well-established scientific evidence is misleading.
- Philosophical studies on the policy relevance of causal knowledge tend to leave notions such as “policy,” “policy goal,” or “policy process” almost entirely unanalyzed.
- Policy making is a dynamic process with distinct stages, and each stage involves distinct aims and evidential requirements.
- Once we realize that policy making cannot be reduced to affecting the value of a unidimensional output variable by intervening in its causes, then there is no reason to presuppose that there is one universal evidential recipe that would be the best for policy making.
- By considering policy making as a process with clear stages, one could analyze separately and more precisely how different types of causal evidence and evidential methods play different roles at each distinct stage of the policy process.

Key readings:

- Anjum, R. L., Copeland, S., and Rocca, E. (Eds.) (2020) *Rethinking causality, complexity and evidence for the unique patient*, Cham: Springer. (An interdisciplinary volume that expounds upon a bottom-up approach to the challenges of applying causal knowledge in modern clinical practice.)
- Cairney, P. (2016) *The politics of evidence-based policy*, Basingstoke: Palgrave Macmillan. (An extended treatment of the complexities surrounding and affecting the policy process.)
- Cartwright, N. D., and Hardie, J. (2012) *Evidence-based policy: a practical guide to doing it better*, Oxford: Oxford University Press. (A systematic and critical philosophical treatment of the role of evidence in supporting policy prescriptions.)

- Dunn, W. (2016) *Public policy analysis: an integrated approach*, New York: Routledge. (A classic textbook offering a thorough characterization of the policy-making process.)
- Smith, K. B., and Larimer, C. W. (2013) *The public policy theory primer*, Boulder: Westview Press. (A good and easy-to-follow introduction to the history and research topics of public policy theory.)

42.1 Introduction

A large part of scientific practice consists of assessing and establishing causal knowledge (that is, knowledge about causal relations, causal processes, causal generalizations, causal evidence, and so forth). In this chapter, I focus on the question: what are the roles of scientific causal knowledge in relation to the evidential requirements of policy making? I begin by concisely reviewing the existing epistemological approaches in philosophy of science to the policy relevance of causal knowledge. I identify the topics and questions on which these philosophical discussions have focused in the last three decades and show how, even if such contributions are illuminating and valuable, they only offer a partial account of the relation between causal knowledge and policy making.

I claim that the current philosophical views connecting causation and policy making can be characterized as top-down approaches, which start by investigating philosophical concerns about causal inference and scientific evidence, and then make broad generalizations about the (potential) policy relevance of the posited philosophical conclusions. Drawing on insights from public policy research, I put forward some preliminary ideas toward an alternative bottom-up philosophical approach to the policy relevance of causal knowledge, that is, one that would start by recognizing concrete policy problems and specific causal and evidential requirements involved in dealing with such problems from a policy-making perspective, and only then proceed to examine how different causal notions and evidential methods could contribute to the aims of the policy-making process.

42.2 The standard view on the practical value of causal knowledge

The idea that knowledge about causal relations is valuable because it can be exploited for the attainment of practical goals seems rather straightforward and uncontroversial across disciplines. Knowing that oxygen is a cause of fire can be used to produce or to extinguish fire; knowing that exposure to asbestos is a cause of cancer can be used to prevent cancer; knowing that excess demand is a cause of price increases can help determine whether it is lucrative to trade certain commodities; or knowing that education is a cause of income can be used to earn higher income. Leuridan et al. (2008) have referred to this relatively simple idea as the “standard view on the practical value of causal knowledge” (p. 298), which they define as “the thesis that the practical value of causal knowledge lies in the fact that manipulation of causes is a good way to bring about a desired change in the effect” (Leuridan et al. 2008: 299). The basic intuition contained in the standard view suggests that if one knows that “X causes Y” is true, then one also knows that bringing about X is an effective way to bring about Y.

As Leuridan et al. (2008: 298–299) point out, the philosophical research on causation has typically taken the standard view as a given. For instance, Nancy Cartwright’s (1979) probabilistic account of causality is explicitly motivated by a formalized version of the standard view. As in other probabilistic accounts (e.g., Suppes 1970; Eells 1991), conditional

probabilities, subject to the right set of causal assumptions, can be used to establish whether causal claims like “X causes G” are true. Cartwright then points out that there is “a natural connection between causes and strategies that should be maintained: if one wants to obtain a goal, it is a good (in the pre-utility sense of good) strategy to introduce a cause for that goal” (Cartwright 1979: 431). And thus, the way causal knowledge connects to effective strategies is put as follows: if X causes G is true, then bringing about X “will be an effective strategy for G in any situation” (p. 432).

Relatedly, the standard view can also be found as one of the key drivers behind the scientific aim of distinguishing *genuine* causes from *spurious* ones, since claims about genuine causation “are needed to ground the distinction between effective strategies and ineffective ones” (Cartwright 1979: 420). Consequently, this view is (implicitly) embedded as a foundational motivation in the literature on causal inference (see, e.g., Simon 1954; Spirtes et al. 1993; Scheines 1997; Glymour 1997; Pearl 2000; Hoover 2001; Hitchcock 2001; Shadish et al. 2002; Morgan and Winship 2007; Angrist and Pischke 2015; Huntington-Klein 2022).

Overall, the standard view (namely, the idea that knowledge of true causal claims has practical value because introducing a cause is a good strategy to bring about a desired effect) has traditionally served philosophers as a good justification for why it is important to investigate the details of genuine causation and the epistemic merits of methods of causal inference employed in scientific practice. In real life, however, the ways science informs policy are of course much more complicated than the standard view can convey. If one is interested in a more comprehensive philosophical understanding of how scientific causal knowledge can inform, relate to, or take part in policy making, then various aspects of the standard view must be unpacked. Is there only one type of causal knowledge? If not, do distinct types of causal knowledge have different kinds of policy relevance? What does it really mean to have policy relevance (or practical value) in the first place? How exactly do we assess it? Is it really the case that all policy goals are accurately represented by the idea of *bringing about a desired change* in a (measurable and neatly defined) effect variable? If there are distinct kinds of policy goals, would it not make sense that the causal requirements for achieving those goals also be of different kinds? The fact that causal knowledge can tell us that bringing about a cause is a good strategy for bringing about a desired change in the effect seems like a promising starting insight, but it does not tell us anything significant about how to answer the kinds of questions raised above, or (as this chapter will illustrate) many other relevant questions concerning the interaction of scientific practice and policy making.

As important as the issue may be, the investigation of how science informs policy making has become a more noticeable topic of interest in philosophy of science only during the last 30 years or so. The existing research, of what can be called the “first wave” of philosophical elaboration on the standard view, has for the most part been devoted to the thorough investigation of the varieties and epistemological intricacies of *scientific evidence*. The main aim of this research has been to offer a more substantial account of the evidential standards required to establish reliable causal knowledge that could be used to inform policy making. For exposition purposes, the explorations in philosophy of science of how causal knowledge relates to policy making can be summarized along three broad lines of research:

- Discussions on evidence diversity and the epistemic import of the different types of evidence that support scientific causal claims.
- Philosophical assessment of the evidence-based policy (EBP) movement, first in medical practice, and later more generally in various policy-oriented social sciences.
- Philosophical analysis of the so-called problem of external validity.

I will briefly describe each of these lines of research in the following three sections. See also Jukola (this volume), Ghiara (this volume), Ilardo and Reiss (this volume), and Carusi (this volume), on different facets and difficulties of using causal evidence in policy regulation.

42.3 Evidence diversity and policy making

There is a long tradition in philosophy of science of reflecting on how evidence can be best used to attain well-established knowledge (see Haack 1993; Mayo 1996; Achinstein 2003; Howson and Urbach 2005; Reiss 2011). More recently, part of this research has focused on analyzing the epistemic contribution of different types of evidence (*evidence diversity*) for the assessment of causal claims. The literature on evidence diversity often refers to typologies of evidence along several distinct dimensions, e.g., probabilistic versus mechanistic, or quantitative versus qualitative, or theory-based versus data-driven. Hence, the answer to the question “what is the best kind of evidence for supporting scientific causal claims?” varies depending on which of the alternative typologies of evidence one has in mind (see also Kelly, this volume). In actual cases of scientific practice, however, these typologies may overlap.

In an influential contribution, Russo and Williamson (2007) argue that there are at least two types of evidence commonly available to empirical scientists for establishing causal claims: difference-making and mechanistic evidence. The most usual example of difference-making evidence is probabilistic evidence, which broadly refers to (quantitative) results obtained using statistical analysis to estimate the influence that a change in one variable has over another variable on average, at the level of some properly defined population. By contrast, mechanistic evidence, also broadly put, refers to (qualitative or quantitative) information about the existence of an underlying mechanism or a causal chain that links the posited cause to (and modulates its influence over) the posited effect, typically at a lower level of analysis than the general or population level.

Debates about difference-making versus mechanistic evidence had previously emerged in connection to epistemological concerns about how to get more reliable causal explanations (e.g., Elster 1989; Little 1991; Bunge 1997; Machamer et al. 2000; Hedström 2005; Glennan 2017). Using more recent terminology, the same explanatory concern can be expressed as how best to establish what Woodward and Hitchcock (2003) have referred to as *explanatory generalizations*, which are invariant causal generalizations that can be used to answer relatively specific what-if-things-had-been-different questions. See Fagan (this volume) for a critical survey of traditional accounts of causal explanation.

Russo and Williamson (2007) famously suggest—in what has come to be known as the Russo-Williamson thesis—that to properly establish a causal claim, scientists require both probabilistic (difference-making) evidence and also evidence of a mechanism connecting the cause to the effect. This thesis has typically been supported and exemplified with cases from the health sciences (Parkkinen et al. 2018), such as the connection between smoking and lung cancer (Gillies 2011) or research on the effects of medical drugs (Aronson et al. 2018), whereas attempts to proclaim its significance for establishing causal relations in the social realm remain quite controversial (see, e.g., Steel 2004; Reiss 2007; Claveau 2012; Moneta and Russo 2014; Marchionni 2017; Ylikoski 2017; Shan and Williamson 2021; Beach 2021; Runhardt 2022; Ghiara 2022). Unsurprisingly, the specifics of what counts as a mechanism or as mechanistic evidence vary greatly from context to context, and from discipline to discipline, which has contributed to making the epistemic role of mechanisms in science an extremely prolific topic of philosophical analysis, speculation, and debate during the last two decades

(see Illari 2011; Glennan and Illari 2017; Marchionni and Reijula 2019; Jiménez-Buedo and Squitieri 2019; Weber, this volume).

Most of the research on the epistemology of causality does not typically make any substantial distinction between assessing the reliability of causal claims and assessing that of policy prescriptions. The latter are taken as special cases of the former and, consequently, there is a tendency to assume that policy prescriptions can be straightforwardly evaluated with the same evidential standards and methods used to test causal hypotheses in scientific practice. Accordingly, philosophical reflections on what type of evidence is the best to inform policy making emerged as a natural extension of the general discussion on the epistemic import of evidence for establishing scientific causal claims.

For instance, following a similar argumentation to those used in defending the Russo-Williamson thesis, mechanistic evidence has been put forward as *indispensable* for the policy relevance of causal claims (e.g., Weber 2007; Grüne-Yanoff 2016), or as an epistemically crucial complement to other types of difference-making evidence to support policy-oriented causal conclusions in general (e.g., Clarke et al. 2014; Marchionni and Reijula 2019; Shan and Williamson 2021). There are also some proposals of evidential pluralism in which no particular type of evidence is said to be indispensable or having priority in the assessment of policy-oriented causal claims (e.g., Claveau 2011, 2012; Stegenga 2022).

It should be noted that all these philosophical studies are valuable and insightful contributions in relation to a specific aim: to assess the epistemic merits of different types of evidence in relation to the scientific goal of establishing causal claims which, in turn, could *potentially* be useful for informing policy decisions. But is this literature on the epistemic import of scientific evidence also suitable for answering the evidential questions that *actually* emerge in the process of policy making?

42.4 Philosophy of science and evidence-based policy

During the 1980s, researchers in medical science started advocating an allegedly more evidence-based approach to scientific practice. This standpoint emerged as a reaction against previous research practices based on what were often considered low-quality and not-so-reliable forms of evidence (e.g., expert opinions, common sense, informal traditional conventions, and the like). The main idea motivating the evidence-based medicine (EBM) approach was that researchers should use more “scientific” and systematic evidence-evaluating standards and methods to establish causal knowledge (Cochrane 1972; Guyatt et al. 1992; Sackett et al. 1996; Petty 2006). In this context, “scientific” means conducive to high and measurable levels of accuracy, combined with a minimum amount of subjective influence on the research outcomes. Accordingly, EBM supporters considered randomized controlled trials (RCTs) as the most scientific of the available evidential methods and promoted the use of rankings of evidential methods from those with allegedly more to those with less scientific validity. See Kelly (this volume) for a more detailed description and assessment of EBM, specifically in the UK.

The tenets of EBM have quickly spread to other policy-oriented disciplines such as political science, development economics, criminology, sociology, to name a few, in what can be seen as a broader evidence-based policy (EBP) movement. As a notable example, EBP has had an obvious influence on economic policy evaluation in the form of what has been called the “empirical” or “credibility turn” in economics (Angrist and Pischke 2010). Advocates of this approach promote the use of design-based empirical techniques as the most scientifically reliable methods to assess the effectiveness of policy programs (see, e.g., Cohen and Easterly 2009; Banerjee and Duflo 2011). Design-based observational studies are evidential methods founded

on the potential-outcomes framework of causal inference, such as instrumental variables, difference-in-differences, regression discontinuity, and the like (see Holland 1986; Morgan and Winship 2007; Angrist and Pischke 2015; Huntington-Klein 2022; Runhardt, this volume).

By the turn of the 21st century, philosophers of science began scrutinizing the guidelines and rankings of the evidential methods promoted by the EBM (e.g., Worrall 2002, 2007; Cartwright 2010; Howick 2011), and subsequently also the methodological procedures defended by the more general EBP movement (e.g., Cartwright and Stegenga 2011; Cartwright and Hardie 2012; Sampson et al. 2013; Gondolf 2015; Cairney 2016; Stafford and Mears 2015). Acknowledging that philosophy of science has a long tradition of studying the epistemology of evidence, Worrall (2002) urged philosophers to direct their intellectual efforts to scrutinizing EBM, with the aim of developing “a fully coherent, articulated and detailed account of the correct relationship between the evidence and various therapeutic and causal claims” (p. 317). In Worrall’s view, EBM clearly was “an area where philosophers of science can, for once, be of real practical value” (Worrall 2002: 317). During the last two decades, many philosophers of science have answered to this call.

Content-wise, philosophical analyses of EBM and EBP have focused on the assessment of the merits and problems of the evidential methods used to test policy-oriented causal claims. In particular, the view that RCTs are the best type of scientific evidence to support causal claims (together with meta-analyses and systematic reviews of RCTs) has been heavily contested. The main criticism has been that it is not at all obvious why RCTs would necessarily have higher epistemic weight, in comparison to other types of evidential methods, and in all contexts and situations (see, e.g., Ashcroft 2004; Borgerson 2009; Worrall 2002, 2007; Cartwright 2006, 2009, 2010; Howick 2011; Stegenga 2011).

Debunking RCTs as the gold standard of empirical evidence has triggered renewed reflections on the epistemic virtues of alternative types of evidential methods, such as observational studies (e.g., Black 1996; Benson and Hartz 2000), expert knowledge (e.g., Selinger and Crease 2006; Collins and Evans 2007; Martini and Boumans 2014), case studies (e.g., Gerring 2007; Ruzzene 2012; Crasnow 2012), process-tracing methods (Beach and Pedersen 2013; Bennett and Checkel 2014; Ruzzene 2014; Crasnow 2017; Zahle, this volume), and other forms of qualitative evidential techniques (e.g., Silverman 2001 [1993]; Swann 2006). Therefore, the ongoing debate on EBP has also fostered increasing philosophical interest in another form of *evidential diversity*, this time not about types of evidence but about different methods of causal inference.

One final issue to point out in relation to EBP is that, regardless of all the careful and thorough research that has been produced by philosophers on the epistemic limitations of the evidential methods promoted by EBP (such as experimental, quasi-experimental, and observational designs), these empirical techniques continue growing in popularity among social scientists at the expense of other alternative non-quantitative or non-experimental methods (Sampson et al. 2013; Kvangraven 2020a; Khosrowi 2021). Moreover, the hype has spread among and influenced the policy makers’ circles as well (Bédécarrats et al. 2017; Boswell 2018).

42.5 The so-called problem of external validity

As described in the previous section, the EBP approach is based on the idea that improving the evidential standards of science by using the best methods of causal inference will result in scientific knowledge that is more reliable for informing policy prescriptions. And ideally, as happens with some well-established medical treatments, well-established policy prescriptions

are expected to be *generalizable* to other populations and contexts in addition to those in which the policies have been tested. In relation to this expectation, another philosophical criticism raised against the EBP movement has been that it often ignores, or takes lightly, what has been labeled “the problem of external validity.” In philosophy of science, the problem of external validity is understood as the difficulty of knowing with certainty the requirements that could epistemically justify the extrapolation of causal results from one setting to another (Cartwright and Hardie 2012).

It should be noted that, in philosophical discussions of the policy relevance of causal knowledge, the notion of “external validity” has commonly been taken out of its original experimental context and meaning (e.g., Cook and Campbell 1979; Shadish et al. 2002), which has sometimes led to confusion and perhaps unnecessary controversies (see Jiménez-Buedo 2011; Jiménez-Buedo and Russo 2021; Blanco-Sequeiros and Mireles-Flores 2024). In philosophical treatments, external validity is typically equated to the necessary and sufficient evidential conditions for extrapolating causal effects from one population to another (Cartwright 2009, 2012; Muller 2014, 2015; Deaton and Cartwright 2018; Reiss 2019). However, on other occasions it has also been taken to mean the abstract generalizability of theoretical claims (Rol 2008), the transportability of causal information from one type of study design to another type (Pearl and Bareinboim 2014), or the requirements for expecting causal results to keep obtaining in the (long-term) future (Weber 2007).

In Cartwright’s view, EBP has failed to provide a good basis for effective policy because their preferred evidential methods do not deal properly with her version of the external validity problem (see Cartwright 2006, 2009). The EBP movement claims that RCTs are good evidence to support claims about implementing “T in order to bring about outcome O,” but Cartwright argues that RCTs only support “claims of one particular form, essentially, ‘T causes O in particular circumstances X in particular population Φ ’” (2009: 129). But if the goal is to bring about outcome O in a different population, how can one be sure that T will be effective there as well? And “for what other populations can we expect these same conclusions to hold?” (2006: 986). Cartwright and other critics of EBP often discuss this problem with a focus on the results of RCTs, but the same concerns about the reliability of extrapolating causal results would affect most, if not all, types of causal-inference methods.

The philosophical study of the policy relevance of causal knowledge has taken the problem of external validity as a major concern. Cartwright has produced a substantial amount of research defining and exposing the external validity problem as well as searching for solutions to it (e.g., Cartwright 2012, 2013). Cartwright and Efstathiou (2011) elaborate an account of evidence for use that recommends taking into consideration all the contextual causal conditions (enabling and disturbing factors) that are relevant for the causal effect to obtain. Then Cartwright and Hardie (2012) develop a conceptual framework for dealing with the problem of external validity, which involves a characterization of the *conditions* under which it would be epistemically justified to predict, from the truth of a causal claim in one situation, the truth of the same claim in a different situation.

Many philosophical studies on the problem of external validity for policy-oriented research have essentially followed or challenged Cartwright’s pioneering studies. The general consensus has been to keep trying to *conceptually or formally* characterize the necessary and sufficient evidential conditions for successful extrapolation of causal effects (Steel 2010; Bareinboim and Pearl 2013; Muller 2014, 2015, 2020, 2021; Khosrowi 2019a, 2022). Whereas a few more pragmatically driven studies have instead focused on exploring *how and to what extent* the outcomes of different types of empirical research could contribute to making inferences

about particular, localized, and contextual policy concerns (e.g., Deaton 2009; Olsen et al. 2013; Claveau and Mireles-Flores 2014, 2016; Ruzzene 2015; Schork 2015; Anjum 2016; Peters et al. 2017; Anjum and Mumford 2017; Khosrowi 2019b; Favereau and Nagatsu 2020; Mireles-Flores 2022). While the extrapolation of causal results is certainly a topic of importance in the context of policy making, it definitely does not exhaust the spectrum of philosophical and practical problems related to causation and causal inference that are relevant to the process of policy making.

42.6 Top-down versus bottom-up philosophical approaches

The lines of research summarized in the previous sections have largely followed what can be called a *top-down approach*: start with a given causal question of philosophical interest, develop a “philosophical account,” and then move, only secondarily, to check whether there are any potential and broadly characterized policy implications that may (at least conceptually) follow. The studies produced have indeed contributed to begin unpacking the standard view, and to that extent they have been relevant for understanding the role of causal knowledge in the process of policy making. Nevertheless, these contributions still adopt a rather idealized, linear, and unidimensional depiction of the relation between scientific knowledge and policy making. And, consequently, they are all carried out in abstraction of the real practice and dynamics of the policy-making process.

For instance, virtually all accounts referred to in the previous section treat “policy” as a vaguely accounted for notion and focus almost exclusively on analyzing, clarifying, and assessing causal notions (e.g., scientific evidence, mechanisms, external validity) and evidential methods (e.g., RCTs, field experiments, process tracing, simulations, expert assessment). The main aim in these studies is to assess the epistemic import and reliability of the scientific practices of knowledge production. But then, it is typically assumed that whatever policy relevance the notions and methods produced may have follows exclusively from those epistemological considerations.

Knowing what is the best scientific evidence for establishing causal relations falls very short of knowing the evidential requirements for achieving effective policy making. Ignoring this difference could generate hazardous unintended gaps between what scientific researchers produce as causal evidence, and the actual evidential needs of policy making (see Mireles-Flores 2022). Confusing causal efficacy with policy effectiveness could lead to what Cartwright calls “the narrow perspective of the experimenter,” that is, “The experimenter asks, ‘To what is my experiment relevant?’ The policy maker instead needs to ask—‘What is relevant to my policy hypothesis?’” (Cartwright 2009: 133). I believe that a similar “narrow perspective” can be observed in the way philosophers of science have tended to approach the connection between causal knowledge and policy making. The philosopher of science asks, “To what policy is my philosophical account relevant?” while what the policy maker needs to know is, “What insights are relevant to my policy problem?”

To be clear, top-down approaches have shed significant light on the epistemic details of causal knowledge. And investigating, characterizing, and assessing the epistemic aspects of causal concepts, causal evidence, evidential methods, and their potential practical value is part of understanding the interaction between scientific knowledge and policy making. My point, however, is that top-down studies are just not sufficient, since fundamental notions such as “policy,” “policy goal,” or “policy process” are left almost entirely unanalyzed, mostly taken for granted, understood superficially or in general terms, and their broad assessment often

taken as a self-evident consequence of getting the epistemology of causal knowledge right. Moreover, top-down analyses, at best, tend to be meant to uncover relevant epistemological features of the evidence or of the causal methods used to establish policy-related causal claims. At worst, however, they are merely meant to defend an abstract or conceptual philosophical position (ultimately detached from any concrete policy goal or problem).

I propose that one could alternatively follow a *bottom-up approach*: start by engaging directly with the way in which concrete real-life policy problems and goals are construed by policy makers, trying to thoroughly understand the local and diverse aims and evidential requirements for dealing with the problems in the varying contexts of the policy process. Then at the next step, in the light of the specifics of the issue, one can look for the relevant causal theories, accounts of evidence, and causal methods that could better suit the different requirements of the policy-making process in question. See Rocca, Anjun, and Andersen (this volume), Jukola (this volume), Ghiara (this volume), Ilardo and Reiss (this volume), and Carusi (this volume), for studies very much in line with the spirit of a bottom-up approach.

To put my proposal in perspective, let me offer a concise summary of the story so far. For most of the 20th century, the standard view about how causal knowledge has practical value was taken for granted and mostly referred to as a basic motivating intuition behind studying causation and causal inference. But what exactly such “practical value” amounts to was never systematically analyzed. Then, in recent decades, philosophers concerned with how scientific knowledge interacts with social and policy issues began investigating more in depth the varieties, and epistemic particularities of different causal and evidential notions employed in scientific practice. This has led to an extremely valuable collection of epistemological insights and accounts of scientific evidence, evidential diversity, causal inference, extrapolation, and other causality-related topics, together with some tentative elaborations on how these insights could have implications for policy making. Understandably, this research has focused on unpacking and assessing the roles of causality and evidence mainly on the scientific side of the interaction, leaving the details of the policy-making side relatively unexplored. Nevertheless, it seems to me that it is now high time that philosophers interested in the policy relevance of science begin engaging more deeply with the process of policy making in its own right. In the next section, I will put forward some concrete ideas, although mainly programmatic, about how this could be done.

42.7 Toward a bottom-up approach to the policy relevance of causal knowledge

There is a large and very rich literature on public policy which, in contrast to philosophical accounts, takes the “policy” side of the evidence-policy interaction as the main subject of analysis. Public policy research is a relatively young discipline, with a history that dates back to the 1950s, and is considered “a field that is at least as much a science as it is an art or a craft” (Smith and Larimer 2013: 2). There are at least two insights from public policy studies that could be useful to guide a more promising philosophical investigation of the relation between scientific causal knowledge and policy making.

First, in public policy the notions of “policy” and “policy making” are understood as a complex dynamic process composed of several stages, including *agenda setting*, *policy formulation*, *legitimation*, *implementation*, *evaluation*, *maintenance*, *termination*, and *prevention*. There have been several models and detailed characterizations of the policy process and its distinct stages (see Birkland 2016; Dunn 2016; Hill and Varone 2016; Kraft and Furlong

2017), so there is no one absolute or ultimate theory about which stages may or may not obtain in different policy processes. Instead, the model is employed by the researchers to bring systematic coherence and to help clarify the distinct elements, aims, and agents that take part in the policy-making process (Smith and Larimer 2013: 27–33). See Grin (this volume) for a more focused look at some of the stages of the policy process.

Second, there is a plethora of discussions and substantial characterizations of many of the potential enabling and disturbing factors that can affect the outcomes of policy making, such as ideologies, power imbalances, translation problems, political interests, institutional settings, and the like (Cairney 2016). These factors play a crucial role in public policy studies, not only by interfering with, mediating, or counterbalancing the attainment of the final policy goal, but much more specifically by means of systematic interactions at each of the different stages of the policy process (Head 2013; Jasanoff 2013).

The study of policy making as a dynamic process, and the detailed investigation of the different factors that could influence the process at each stage, may explain why public policy scholars have often been rather critical of the simplistic depiction of policy making, and of the restricted understanding of science-policy interactions that are inherent in the EBP movement (Newman 2016). For instance, some authors have criticized how EBP research tends to investigate causal relations between two (or a relatively small well-defined system of) measurable variables, as if they could really be properly understood in isolation from the social context of the policy process in which they operate (e.g., Greenhalgh and Russell 2009; Bédécarrats et al. 2017; Kvangraven 2020b). Other critics have pointed out the role of political and ideological interests behind the (supposedly objective) scientific results obtained using EBP methodologies and are suspicious of undisclosed non-epistemic interests affecting (or biasing) the interaction between scientific researchers and state governments. This has led to the evidence favored by EBP being referred to sarcastically as “policy-based evidence” (e.g., Marmot 2004; Head 2013; Strassheim and Kettunen 2014; Boswell 2018). Remarkably, these lines of criticism have seldom been carefully considered in the philosophical literature.

By taking seriously the relevant insights from public policy research about the dynamic nature of the policy process, philosophers of science might finally be able to open the policy black box and be better equipped to contribute more relevant bottom-up studies about the interaction between causal knowledge and policy making. One way to proceed would be to start by identifying and interpreting the various challenges at each distinct stage of a policy process in relation to a particular policy goal, and only after that perhaps climb up to the rich philosophical library of conceptual frameworks and accounts to gather the appropriate tools that could help us contribute to the understanding of the science-policy interaction under study.

A straightforward and significant consequence of adopting a bottom-up approach like the one advanced here would be that, by considering and interpreting policy making as a process with clear stages (and each stage with distinct aims and requirements), one could analyze and assess separately and more precisely how different types of causal knowledge and evidential techniques play different roles at each distinct stage of the policy process. Thus, instead of general philosophical debates about whether a scientific claim is policy relevant or not, or about which particular type of evidence or inferential method is better or worse for “policy purposes” overall, the discussion could be refocused toward more localized and informative concerns, such as which pieces of causal knowledge, evidential methods, or mediating contextual factors are more or less relevant, and in what ways, to the specific aims and evidential requirements *at each of the different stages* of the policy process.

As an illustration, the longstanding debate on whether difference-making or mechanistic evidence is the best evidence for policy could be expressed in slightly different terms. Instead

of asking what the best evidence is that scientists should be putting forward to support policy prescriptions, or to best attain policy goals, we can now ask: in relation to a particular policy problem, what are the evidential requirements at each of the various stages of the policy process that will be conducive to attaining the policy goal at hand? The “diagnostic stage” may require forms of qualitative evidence, expert opinion, or some type of mechanistic evidence to figure out the causal and contextual characteristics of the policy problem, whereas at other stages, such as “implementation” or “evaluation,” quantitative and probabilistic evidence, perhaps in combination with expert contextual accounts, may have a more substantive evidential role.

As soon as we realize that policy making cannot be reduced to changing the value of a unidimensional output variable by intervening in its causes, it becomes clear that there is no reason to presuppose that there is one universal evidential recipe that would be *the best* for policy making. Hence, general philosophical claims of the form “the such-and-such type of evidence is necessary for policy” are, from a policy-making perspective, rather ambiguous and uninformative. Is the evidence meant to contribute to the diagnosis of the policy problem? To the deliberation of potential alternative solutions? To the decision process about which is the best solution in this case? To the implementation plan? For what precise evidential need in the policy process is the evidence supposed to be crucially relevant? Moreover, policy problems are extremely diverse in character and scope, and the details and required stages of the policy process to deal with such problems will be just as diverse.

Another potential consequence of adopting a bottom-up approach could be a reconsideration of the difficulties concerning evidence amalgamation or integration, and the role of mixed methods of causal inference in relation to policy making. The idea behind current research on these topics is that employing a variety of types of evidence to establish a causal claim makes the claim epistemically more robust. So philosophers have been investigating how it is possible to combine different types of evidence to support a causal conclusion. However, different types of evidence are often incompatible, discordant, or incomparable, which has been considered problematic for evidence amalgamation (Kuorikoski and Marchionni 2016).

In policy making, different pieces and kinds of evidence have concurrent roles throughout the policy process, not necessarily to support a single causal hypothesis in isolation but rather to justify or support the several interrelated but distinct causal claims that are relevant at each stage of the policy process. Different types of evidence and evidential methods are all “combined” so to speak through the whole policy process, but they need not be amalgamated, or integrated, or mixed, into a single research study or in the form of an ultimate causal result. Thus, formal issues like incompatibility or incomparability of evidence are not automatically problematic in relation to the use of different types of evidence in policy making, at least not in the same way in which these issues may be problematic for assessing and establishing a scientific causal hypothesis by using a variety of types or sources of evidence. Nevertheless, a careful and fresh analysis of what kinds of evidence integration or amalgamation may or may not be required *within each of the different stages* of the policy process appears to be a wide unexplored area for philosophical investigation.

Similarly, in relation to the philosophy of mixed-methods research, which purports to combine qualitative and quantitative evidence to establish a causal hypothesis in the same study design (see Teddlie and Tashakkori 2009; Parkkinen et al. 2018; Runhardt 2021; Shan 2022; Kuorikoski and Marchionni 2023), an entirely new question to explore would be: which kinds of evidential requirements within a process of policy making may call for employing mixed-methods research? The key point to note is that the role of mixed methods in policy making need not be identical to their role in scientific practice, since the main concern of

policy making is not necessarily to mix different types of evidence to establish *one single* causal claim but rather to use all available relevant (varieties of) evidence to assess and support *different claims throughout the distinct stages* of the policy process.

42.8 Conclusions

The ways in which scientific causal knowledge relates to policy making in real-life situations are much more intricate than existing philosophical accounts manage to convey. In this chapter, I surveyed the main lines of philosophical research on the policy relevance of causal knowledge and argued that regardless of the available valuable contributions to the topic, the “policy” side of the science-policy interaction has mostly remained unexplored. I claimed that the existing research tends to follow a top-down approach, which is mainly motivated by conceptual philosophical concerns about causal inference and scientific evidence, and contented with making general claims about the (potential) policy relevance of the posited philosophical conclusions.

After pointing out the limitations of the top-down approach, I put forward some initial programmatic ideas toward an alternative bottom-up philosophical approach to the policy relevance of causal knowledge. My proposal draws on insights from public policy research, namely, that policy making is a dynamic process, including distinct stages, with distinct aims and evidential requirements emerging at each stage of the policy process. Thus, different forms of evidence and causal methods are called upon to play diverse roles throughout the entire process. This view contrasts with the existing philosophical research that tends to take the notion of a “policy goal” as a unidimensional outcome variable.

Even if the specifics of my proposal as presented here require further elaboration and debate, the main purpose of this chapter is to make evident why the policy process should not remain unanalyzed in philosophical accounts that aim at offering meaningful insights about the ways science interacts with policy making. Opening the black box of the policy process and engaging in bottom-up research are crucial steps toward bringing out philosophical studies of the policy relevance of causal knowledge that (in line with Worrall’s hope) could, moreover, be of “real practical value.”

42.9 Related chapters

Andersen, F., Anjum, R. L., and Rocca, E. “When decisions must be based on partial causal knowledge”.

Grin, J. “Causation in policy science: knowledge, power, meaning, agency and context”.

Jukola, S. “From evidence to policy”.

Ghiara, V. “Causal evidence and the social determinants of health”.

Ilardo, M. L., and Reiss, J. “Causation, regulation, and the assessment of Adverse Events Following Immunization”.

Kelly, M. P. “Evidence, causation, guidelines and regulation”.

Carusi, A. “Science to policy through Adverse Outcome Pathways”.

Acknowledgments

I am extremely grateful to Melinda Fagan, Alexander Gebharter, Virginia Ghiara, Magdalena Małecka, and the editors of the volume, Phyllis Illari, and Federica Russo, for their valuable feedback on previous versions of this chapter. I have discussed some of the ideas in this chapter at the 13th conference of the International Network for Economic Method, San Sebastián,

2017; and at the symposium “Reasoning about Evidence,” Ghent, 2019. This research was supported by the Academy of Finland, project 330524, “Opening the policy black-box in evidence-based policy.”

Bibliography

- Achinstein, P. (2003) *The book of evidence*, Oxford: Oxford University Press.
- Angrist, J. D., and Pischke, J.-S. (2010) “The credibility revolution in empirical economics: How better research design is taking the con out of econometrics,” *Journal of Economic Perspectives* 24(2): 3–30.
- Angrist, J. D., and Pischke, J.-S. (2015) *Mastering metrics: The path from cause to effect*, Princeton: Princeton University Press.
- Anjum, R. L. (2016) “Evidence based or person centered? An ontological debate,” *European Journal for Person Centred Healthcare* 4(2): 421–429.
- Anjum, R. L., and Mumford, S. D. (2017) “A philosophical argument against evidence-based policy,” *Journal of Evaluation in Clinical Practice* 23(5): 1045–1050.
- Aronson, J. K., La Caze, A., Kelly, M. P., Parkkinen, V., and Williamson, J. (2018) “The use of mechanistic evidence in drug approval,” *Journal of Evaluation in Clinical Practice* 24(5): 1166–1176.
- Ashcroft, R. E. (2004) “Current epistemological problems in evidence based medicine,” *Journal of Medical Ethics* 30(2): 131–135.
- Banerjee, A. V., and Dufló, E. (2011) *Poor economics: A radical rethinking of the way to fight global poverty*, New York: Public Affairs.
- Bareinboim, E., and Pearl, J. (2013) “A general algorithm for deciding transportability of experimental results,” *Journal of Causal Inference* 1(1): 107–134.
- Beach, D. (2021) “Evidential pluralism and evidence of mechanisms in the social sciences,” *Synthese* 199(3/4): 8899–8919.
- Beach, D., and Pedersen, R. B. (2013) *Process-tracing methods: Foundations and guidelines*, Ann Arbor: University of Michigan Press.
- Bédécarrats, F., Guérin, I., and Roubaud, F. (2017) “All that glitters is not gold: The political economy of randomized evaluations in development,” *Development and Change* 50(3): 735–762.
- Bennett, A., and Checkel, J. T. (2014) *Process tracing: From metaphor to analytic tool*, Cambridge: Cambridge University Press.
- Benson, K., and Hartz, A. J. (2000) “A comparison of observational studies and randomized, controlled trials,” *New England Journal of Medicine* 342(25): 1878–1886.
- Birkland, T. (2016) *An introduction to the policy process*, New York: Routledge.
- Black, N. (1996) “Why we need observational studies to evaluate the effectiveness of health care,” *British Medical Journal* 312(7040): 1215–1218.
- Blanco-Sequeiros, S., and Mireles-Flores, L. (2024) “Reorienting external validity analysis,” *TINT Working Paper*, University of Helsinki, Helsinki.
- Borgerson, K. (2009) “Valuing evidence: Bias and the evidence hierarchy of evidence-based medicine,” *Perspectives in Biology and Medicine* 52(2): 218–233.
- Boswell, J. (2018) “What makes evidence-based policy making such a useful myth?” *Governance* 31(2): 199–214.
- Bunge, M. (1997) “Mechanism and explanation,” *Philosophy of the Social Sciences* 27(4): 410–465.
- Cairney, P. (2016) *The politics of evidence-based policy*, Basingstoke: Palgrave Macmillan.
- Cartwright, N. D. (1979) “Causal laws and effective strategies,” *Noûs* 13(4): 419–437.
- Cartwright, N. D. (2006) “Well-ordered science: Evidence for use,” *Philosophy of Science* 73(5): 981–990.
- Cartwright, N. D. (2009) “Evidence-based policy: what’s to be done about relevance?” *Philosophical Studies* 143(1): 127–136.
- Cartwright, N. D. (2010) “What are randomised controlled trials good for?” *Philosophical Studies* 147(1): 59–70.
- Cartwright, N. D. (2012) “Will this policy work for you? Predicting effectiveness better: How philosophy helps,” *Philosophy of Science* 79(5): 973–989.
- Cartwright, N. D. (2013) “Knowing what we are talking about: Why evidence doesn’t always travel,” *Evidence and Policy* 9(1): 97–112.
- Cartwright, N. D., and Efstathiou, S. (2011) “Hunting causes and using them: Is there no bridge from here to there?” *International Studies in the Philosophy of Science* 25(3): 223–241.

- Cartwright, N. D., and Hardie, J. (2012) *Evidence-based policy: A practical guide to doing it better*, Oxford: Oxford University Press.
- Cartwright, N. D., and Stegenga, J. (2011) "A theory of evidence for evidence-based policy," in P. Dawid, W. Twining, and M. Vasilaki (eds.) *Evidence, inference and enquiry, proceedings of the British Academy* 171. Oxford: Oxford University Press: 289–319.
- Clarke, B., Gillies, D., Illari, P., Russo, F., and Williamson, J. (2014) "Mechanisms and the evidence hierarchy," *Topoi* 33(2): 339–360.
- Claveau, F. (2011) "Evidential variety as a source of credibility for causal inference: Beyond sharp designs and structural models," *Journal of Economic Methodology* 18(3): 233–253.
- Claveau, F. (2012) "The Russo-Williamson thesis in the social sciences: Causal inference drawing on two types of evidence," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43(4): 806–813.
- Claveau, F., and Mireles-Flores, L. (2014) "On the meaning of causal generalisations in policy-oriented economic research," *International Studies in the Philosophy of Science* 28(4): 397–416.
- Claveau, F., and Mireles-Flores, L. (2016) "Causal generalisations in policy-oriented economic research: An inferentialist analysis," *International Studies in the Philosophy of Science* 30(4): 383–398.
- Cochrane, A. L. (1972) *Effectiveness and efficiency: Random reflection on health services*, London: Nuffield Provincial Hospitals Trust.
- Cohen, J., and Easterly, W. (2009) *What works in development? Thinking big and thinking small*, Washington: Brookings Institution.
- Collins, H., and Evans, R. (2007) *Rethinking expertise*, Chicago: University of Chicago Press.
- Cook, T. D., and Campbell, D. T. (1979) *Quasi-experimentation: Design and analysis issues for field settings*, Boston: Houghton Mifflin.
- Crasnow, S. (2012) "The role of case study research in political science: Evidence for causal claims," *Philosophy of Science* 79(5): 655–666.
- Crasnow, S. (2017) "Process tracing in political science: what's the story?" *Studies in History and Philosophy of Science Part A*, 62: 6–13.
- Deaton, A. S. (2009) "Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development," *NBER Working Paper No. 14690*, The National Bureau of Economic Research, Cambridge, MA.
- Deaton, A. S., and Cartwright, N. D. (2018) "Understanding and misunderstanding randomized controlled trials," *Social Science & Medicine* 210: 2–21.
- Dunn, W. (2016) *Public policy analysis: An integrated approach*, New York: Routledge.
- Eells, E. (1991) *Probabilistic causality*, Cambridge: Cambridge University Press.
- Elster, J. (1989) *Nuts and bolts for the social sciences*, Cambridge: Cambridge University Press.
- Favereau, J., and Nagatsu, M. (2020) "Holding back from theory: Limits and methodological alternatives of randomized field experiments in development economics," *Journal of Economic Methodology* 40(1): 1–21.
- Gerring, J. (2007) *Case study research: Principles and practices*, Cambridge: Cambridge University Press.
- Ghiara, V. (2022) "Taking the Russo-Williamson thesis seriously in the social sciences," *Synthese* 200: 481, <https://doi.org/10.1007/s11229-022-03924-3>
- Gillies, D. (2011) "The Russo-Williamson thesis and the question of whether smoking causes heart disease," in P. Illari, F. Russo, and J. Williamson (eds.) *Causality in the sciences*. Oxford: Oxford University Press: 110–126.
- Glennan, S. (2017) *The new mechanical philosophy*, Oxford: Oxford University Press.
- Glennan, S., and Illari, P. M. (2017) "Mechanisms in economics," in S. Glennan, and P. Illari (eds.) *The Routledge handbook of mechanisms and mechanical philosophy*. London: Routledge.
- Glymour, C. (1997) "A review of recent work on the foundations of causal inference," in V. McKim, and S. Turner (eds.) *Causality in crisis?* Notre Dame: University of Notre Dame Press: 201–248.
- Gondolf, E. W. (2015) "The evidence-based practice movement," in R. A. Scott, S. M. Kosslyn, and N. Pinkerton (eds.) *Emerging trends in the social and behavioral sciences*. Hoboken: John Wiley & Sons.
- Greenhalgh, T., and Russell, J. (2009) "Evidence-based policymaking: A critique," *Perspectives in Biology and Medicine* 52(2): 304–318.
- Grüne-Yanoff, T. (2016) "Why behavioural policy needs mechanistic evidence," *Economics and Philosophy* 32(3): 463–483.

Causal Knowledge and the Process of Policy Making

- Guyatt, G., Cairns, J., Churchill, D., Cook, D., Haynes, B., et al. (1992) "Evidence-based medicine: A new approach to teaching the practice of medicine," *JAMA: The Journal of the American Medical Association* 268(17): 2420–2425.
- Haack, S. (1993) *Evidence and inquiry: Towards reconstruction in epistemology*, Oxford: Blackwell.
- Head, B. (2013) "Evidence-based policymaking: Speaking truth to power?" *Australian Journal of Public Administration* 72(4): 397–403.
- Hedström, P. (2005) *Dissecting the social: On the principles of analytical sociology*, New York: Cambridge University Press.
- Hill, M., and Varone, F. (2016) *The public policy process*, New York: Routledge.
- Hitchcock, C. R. (2001) "Causal generalizations and good advice," *The Monist* 84(2): 219–242.
- Holland, P. W. (1986) "Statistics and causal inference," *Journal of the American Statistical Association* 81(396): 945–960.
- Hoover, K. D. (2001) *Causality in macroeconomics*, Cambridge: Cambridge University Press.
- Howick, J. (2011) *The philosophy of evidence-based medicine*, Chichester: Wiley-Blackwell.
- Howson, C., and Urbach, P. (2005) *Scientific reasoning: The Bayesian approach*, Chicago: Open Court Publishing.
- Huntington-Klein, N. (2022) *The effect: An introduction to research design and causality*, New York: Chapman and Hall/CRC Press.
- Illari, P. M. (2011) "Mechanistic evidence: Disambiguating the Russo-Williamson thesis," *International Studies in the Philosophy of Science* 25(2): 139–157.
- Jasanoff, S. (2013) "The science of science advice," in R. Doubleday, and J. Wilsdon (eds.) *Future directions for scientific advice in Whitehall*. Cambridge: Cambridge Centre for Science and Policy: 62–68.
- Jiménez-Buedo, M. (2011) "Conceptual tools for assessing experiments: Some well-entrenched confusions regarding the internal/external validity distinction," *Journal of Economic Methodology* 18(3): 271–282.
- Jiménez-Buedo, M., and Russo, F. (2021) "Experimental practices and objectivity in the social sciences: Re-embedding construct validity in the internal–external validity distinction," *Synthese* 199(3/4): 9549–9579.
- 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.
- Khosrowi, D. (2019a) "Extrapolation of causal effects: Hopes, assumptions, and the extrapolator's circle," *Journal of Economic Methodology* 26(1): 45–58.
- Khosrowi, D. (2019b) "Trade-offs between epistemic and moral values in evidence-based policy," *Economics and Philosophy* 35(1): 49–71.
- Khosrowi, D. (2021) "Evidence-based policy," in C. Heilmann, and J. Reiss (eds.) *The Routledge handbook of the philosophy of economics*. London: Routledge.
- Khosrowi, D. (2022) "What's (successful) extrapolation?" *Journal of Economic Methodology* 29(2): 140–152.
- Kraft, M. E., and Furlong, S. R. (2017) *Public policy: Politics, analysis, and alternatives*, California: Sage.
- Kuorikoski, J., and Marchionni, C. (2016) "Evidential diversity and the triangulation of phenomena," *Philosophy of Science* 83(2): 227–247.
- Kuorikoski, J., and Marchionni, C. (2023) "Evidential variety and mixed-methods research in social science," *Philosophy of Science* [pre-pub] <https://doi.org/10.1017/psa.2023.34>
- Kvangraven, I. H. (2020a) "Nobel rebels in disguise: Assessing the rise and rule of the randomistas," *Review of Political Economy* 32(3): 305–341.
- Kvangraven, I. H. (2020b) "Impoverished economics? A critical assessment of the new gold standard," *World Development* 127: 104813. <https://doi.org/10.1016/j.worlddev.2019.104813>
- Leuridan, B., Weber, E., and Van Dyck, M. (2008) "The practical value of spurious correlations: Selective versus manipulative policy," *Analysis* 68(4): 298–303.
- Little, D. (1991) *Varieties of social explanation: An introduction to the philosophy of social science*, Boulder: Westview.
- Machamer, P., Darden, L., and Craver, C. F. (2000) "Thinking about mechanisms," *Philosophy of Science* 67(1): 1–25.
- Marchionni, C. (2017) "Mechanisms in economics," in S. Glennan, and P. Illari (eds.) *The Routledge handbook of mechanisms and mechanical philosophy*. London: Routledge: 423–434.

- 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.
- Marmot, M. G. (2004) "Evidence based policy or policy based evidence?" *British Medical Journal* 328(7445): 906–907.
- Martini, C., and Boumans, M. (eds.) (2014) *Experts and consensus in social science*, Dordrecht: Springer.
- Mayo, D. (1996) *Error and the growth of experimental knowledge*, Chicago: University of Chicago Press.
- Mireles-Flores, L. (2022) "The evidence for free trade and its background assumptions: How well-established causal generalisations can be useless for policy," *Review of Political Economy* 34(3): 534–563.
- Moneta, A., and Russo, F. (2014) "Causal models and evidential pluralism in econometrics," *Journal of Economic Methodology* 21(1): 54–76.
- Morgan, L., and Winship, C. (2007) *Counterfactuals and causal inference: Methods and principles for social research*, Cambridge: Cambridge University Press.
- Muller, S. M. (2014) "Randomised trials for policy: a review of the external validity of treatment effects," *A Southern Africa Labour and Development Research Unit Working Paper Number 127*, SALDRU, University of Cape Town, Cape Town.
- Muller, S. M. (2015) "Causal interaction and external validity: Obstacles to the policy relevance of randomized evaluations," *The World Bank Economic Review* 29(S1): S217–S225.
- Muller, S. M. (2020) "The implications of a fundamental contradiction in advocating randomized trials for policy," *World Development* 127: 104831, <https://doi.org/10.1016/j.worlddev.2019.104831>
- Muller, S. M. (2021) "Randomised trials in economics," in H. Kincaid, and D. Ross (eds.) *A modern guide to philosophy of economics*. Cheltenham: Edward Elgar: 90–126.
- Newman, J. (2016) "Deconstructing the debate over evidence-based policy," *Critical Policy Studies* 11(2): 211–226.
- Olsen, R. B., Orr, L. L., Bell, S. H., and Stuart, E. A. (2013) "External validity in policy evaluations that choose sites purposively," *Journal of Policy Analysis and Management* 32(1): 107–121.
- 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*, Cham: Springer.
- Pearl, J. (2000) *Causality: Models, reasoning, and inference*, Cambridge: Cambridge University Press.
- Pearl, J., and Bareinboim, E. (2014) "External validity: From do-calculus to transportability across populations," *Statistical Science* 29(4): 579–595.
- Peters, J., Langbein, J., and Roberts, G. (2017) "Generalization in the tropics: Development policy, randomized controlled trials, and external validity," *Ruhr Economic Papers* 716. RWI, Berlin.
- Petty, G. (2006) *Evidence-based teaching: A practical approach*, Cheltenham: Nelson Thornes Ltd.
- Reiss, J. (2007) "Do we need mechanisms in the social sciences?" *Philosophy of the Social Sciences* 37(2): 163–184.
- Reiss, J. (2011) "Empirical evidence: Its nature and sources," in I. Jarvie, and J. Zamora-Bonilla (eds.) *Handbook of philosophy of social science*. London: SAGE: 551–576.
- Reiss, J. (2019) "Against external validity," *Synthese* 196(8): 3103–3121.
- Rol, M. (2008) "Idealization, abstraction, and the policy relevance of economic theories," *Journal of Economic Methodology* 15(1): 69–97.
- Runhardt, R. W. (2021) "Evidential pluralism and epistemic reliability in political science: Deciphering contradictions between process tracing methodologies," *Philosophy of the Social Sciences* 51(4): 425–442.
- Runhardt, R. W. (2022) "Limits to evidential pluralism: Multi-method large-N qualitative analysis and the primacy of mechanistic studies," *Synthese* 200(2): 1–23.
- Russo, F., and Williamson, J. (2007) "Interpreting causality in the health sciences," *International Studies in the Philosophy of Science* 21(2): 157–170.
- Ruzzene, A. (2012) "Drawing lessons from case studies by enhancing comparability," *Philosophy of the Social Sciences* 42(1): 99–120.
- Ruzzene, A. (2014) "Process tracing as an effective epistemic complement," *Topoi* 33(2): 1–12.
- Ruzzene, A. (2015) "Policy-making in developing countries: From prediction to planning," *Journal of Economic Methodology* 22(3): 264–279.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., and Richardson, W. S. (1996) "Evidence based medicine: What it is and what it isn't," *British Medical Journal* 312(7023): 71–72.

- Sampson, R. J., Winship, C., and Knight, C. (2013) "Translating causal chains: Principles and strategies for policy-relevant criminology," *Criminology and Public Policy* 12(4): 587–616.
- Scheines, R. (1997) "An introduction to causal inference," in V. McKim, and S. Turner (eds.) *Causality in crisis?* Notre Dame: University of Notre Dame Press: 185–200.
- Schork, N. J. (2015) "Personalized medicine: Time for one-person trials," *Nature* 520(7549): 609–611.
- Selinger, E., and Crease, R. P. (eds.) (2006) *The philosophy of expertise*, New York: Columbia University Press.
- Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002) *Experimental and quasi-experimental designs for generalized causal inference*, Boston: Houghton-Mifflin.
- Shan, Y. (2022) "Philosophical foundations of mixed methods research," *Philosophy Compass* 17(1): e12804 [pre-pub] <https://doi.org/10.1111/phc3.12804>
- Shan, Y., and Williamson, J. (2021) "Applying evidential pluralism to the social sciences," *European Journal for Philosophy of Science* 11(4): 1–27.
- Silverman, D. (2001 [1993]) *Interpreting qualitative data: Methods for analysing talk, text, and interaction*, London: SAGE.
- Simon, H. A. (1954) "Spurious correlation: A causal interpretation," *Journal of American Statistical Association* 49: 467–492.
- Smith, K. B., and Larimer, C. W. (2013) *The public policy theory primer*, Boulder: Westview Press.
- Spirtes, P., Glymour, C., and Scheines, R. (1993) *Causation, prediction, and search*, Berlin: Springer.
- Stafford, M. C., and Mears, D. P. (2015) "Causation, theory, and policy in the social sciences," in R. A. Scott, S. M. Kosslyn, and N. Pinkerton (eds.) *Emerging trends in the social and behavioral sciences*. Hoboken: John Wiley & Sons.
- Steel, D. (2004) "Social mechanisms and causal inference," *Philosophy of the Social Sciences* 34(1): 55–78.
- Steel, D. (2010) "A new approach to argument by analogy: Extrapolation and chain graphs," *Philosophy of Science* 77(5): 1058–1069.
- Stegenga, J. (2011) "Is meta-analysis the platinum standard of evidence?" *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 42(4): 497–507.
- Stegenga, J. (2022) "Evidence of effectiveness," *Studies in History and Philosophy of Science* 91: 288–295.
- Strassheim, H., and Kettunen, P. (2014) "When does evidence-based policy turn into policy-based evidence?" *Evidence and Policy* 10(2): 259–277.
- Suppes, P. (1970) *A probabilistic theory of causality*, Amsterdam: North-Holland.
- Swann, G. M. P. (2006) *Putting econometrics in its place: A new direction in applied economics*, Cheltenham: Edward Elgar.
- Teddle, C., and Tashakkori, A. (2009) *Foundations of mixed methods research*, London: Sage.
- Weber, E. (2007) "Social mechanisms, causal inference, and the policy relevance of social science," *Philosophy of the Social Sciences* 37(3): 348–359.
- Woodward, J., and Hitchcock, C. (2003) "Explanatory generalizations, part I: A counterfactual account," *Noûs* 37(1): 1–24.
- Worrall, J. (2002) "What evidence in evidence-based medicine?" *Philosophy of Science* 69(S3): 316–330.
- Worrall, J. (2007) "Evidence in medicine and evidence-based medicine," *Philosophy Compass* 2(6): 981–1022.
- Ylikoski, P. (2017) "Social mechanisms," in S. Glennan, and P. Illari (eds.) *The Routledge handbook of mechanisms and mechanical philosophy*. London: Routledge: 401–412.