Check your truth conditions! Clarifying the relationship between theories of causation and social science methods for causal inference

Ingo Rohlfing* and Christina Isabel Zuber**

*University of Cologne (ORCID: 0000-0001-8715-4771)

**University of Konstanz (ORCID: 0000-0002-6403-7342)

08.08.2018

Abstract

Theories of causation in philosophy ask what makes causal claims true and establish so called truth conditions allowing one to separate causal from non-causal relationships. We argue that social scientists should be aware of truth conditions of causal claims because they imply which method of causal inference can establish whether a specific claim holds true. A survey of social scientists shows that this is worth emphasizing because many respondents have unclear concepts of causation and link methods to philosophical criteria in an incoherent way. We link five major theories of causation to major small and large-n methods of causal inference to provide clear guidelines to researchers and improve dialogue across methods. While most theories can be linked to more than one method, we argue that structural counterfactual theories are most useful for the social sciences since they require neither social and natural laws nor physical processes to assess causal claims.

Acknowledgments

We would like to thank two reviewers for their comments. Previous versions of this article have been presented at the *AnaGramm* workshop on "Social Mechanisms: Methodological Challenges, Empirical Applications & Modeling Techniques", 7.-8. November 2014, Düsseldorf, at the ELP-Colloquium of the Department of Philosophy, University of Konstanz, 11 February 2016 and at the Research Colloquium of the Graduate School of Decision Sciences of the University of Konstanz, 7 November 2017. We thank participants on these occasions for their very helpful comments. Special thanks go to Jochen Briesen, Dina Emundts, Meinard Kuhlmann, Andrea Lailach-Hennrich and Wolfgang Spohn, whose comments have greatly supported our attempt to access a complex philosophical debate. On the philosophical side, we are particularly indebted to Tobias Henschen for numerous constructive discussions that helped us clarify our arguments and detailed comments on the final version of the manuscript. Nathalie Behnke, Simon Franzmann, Ulrich Glassmann, Leonce Röth, Julian Schuessler, Nils Weidmann and Gregor Zons gave very helpful comments from the perspective of political science. Charlotte Brommer-Wierig, Yelda Grönlund and Nancy Deyo provided valuable research assistance. Last but not least, we thank all respondents for taking the time to answer our survey on causation in the social sciences. Ingo Rohlfing received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement nr. 638425, Enhanced Qualitative and Multimethod Research). Both authors contributed equally to the article.

Social scientists use a variety of methods to support causal claims. Whether one method achieves this better than others has been the subject of extensive and heated debate (Moses and Knutsen 2007:chap. 1). This debate has been productive and, in many respects, has provided clarity, for example, by increasing recognition of the difference between effects and mechanisms and the different types of observations that are collected in qualitative and quantitative research (Beach and Pedersen 2016:chap. 1; Collier, Brady and Seawright 2004).

This article adds a further dimension to this debate and argues that social scientists should be aware of philosophical theories of causation, since they have implications for which method of causal inference can establish whether a given causal claim holds true — and which cannot. Theories of causation in philosophy are interested in the meaning of causal claims and the conditions under which they are true and false (Johnson, Russo and Schoonenboom 2017; Paul and Hall 2013:chap. 2). They thereby establish criteria that allow separating causal from non-causal relations and tell us which types of evidence can confirm or, in contrast, disconfirm a causal argument. To take a simple example, if I understand causes as probability-raisers (for example, Gerring 2005:163), "x causes y" means "x precedes y and raises the probability of y, *ceteris paribus*". The statement is true if I observe that the conditional probability of y given x is higher than the conditional probability of y when x is not given. For the choice of methods, the implication is that a causal claim of probability-raising could be established empirically through statistical analysis of a large-N dataset, but not through process tracing in a single case.

The truth conditions of causal claims posited by different philosophies of causation have received limited attention in the debate about methods of causal inference in the social sciences (Brady 2008).[1] As a consequence, quantitative and qualitative methodologists often misconstrue

[1] Aligning methods and theories of causation is *necessary* for assessing empirically whether a causal claim is true, but it is *not sufficient*. Issues of data availability, data quality and adequate measurement affect the level of *epistemic* certainty with which we judge that our claim is likely to be true or not. We leave these issues aside here because they

how those "on the other side" draw causal inferences. For example, characterizing quantitative research as regularity-based (Beach and Pedersen 2013:chap. 3) is wrong because modern quantitative research is methodologically anchored in a counterfactual, interventionist theory of causation (Morgan and Winship 2014:chap. 1). Qualitative research, on the other hand, can draw on a single-case, counterfactual theory of causation for making causal claims (Lewis 1973a), thus invalidating charges that causal inference always requires a larger number of cases (Dowding 2016:151-157) or has to follow the idea of singular processes and mechanisms without difference-making (Beach and Pedersen 2016:chap. 1).

A survey of 109 social scientists conducted for this article additionally showed that researchers hold partially contradictory views on which types of causal claims can be supported by a given large- or small-N method of causal inference. Strikingly, even if two researchers use the same method and both conclude that "x causes y", they often have a different understanding of what "causes" means, implying that they require different patterns of evidence for assessing their claim.

In order to improve the dialogue about methods and provide guidelines for researchers aiming at causal inference, this article introduces five prominent philosophical theories of causation that define causation in terms of *regularities, probability-raisers, counterfactuals, interventions, or physical processes*. The selection is guided by the state of the art in philosophy of causation where variants of all these five theories are still being defended as viable contenders in the quest for the philosophically best theory of causation (see for example Johnson, Russo and Schoonenboom 2017; Paul and Hall 2013:chap. 2).

We then connect these theories to prominent large- and small-N methods of causal inference in the social sciences: *experiments, observational inferential statistics, Qualitative*

have been discussed plenty of times in work on methods.

Comparative Analysis, comparative case studies, process tracing and multi-method research. In presenting the theories and connecting them to social science methods, we discuss which of the five philosophically viable theories are most adequate for the social sciences more generally (see also Reutlinger 2013) and for each method more specifically.

In this article, we clarify the relationship between these theories of causation and methods of causal inference and thereby make two contributions. First, empirical researchers receive guidance on the evidential requirements that each theory of causation raises. This allows them to determine what claim it is they wish to establish and under which conditions this claim is true. On this basis, researchers can make an informed choice of a method that allows them to empirically assess the causal argument.

Second, we argue that most methods are in principle compatible with more than one theory of causation, as long as the connection between so called *type level theories of causation* (relating classes of facts, or events) and large-N methods on one hand, and so called *token level theories* (relating singular facts, or events) and small-N methods on the other hand is adhered to (Paul and Hall 2013:chap. 2). Based on methodological considerations, however, we argue more specifically that counterfactual theories, formalized in *structural causal modeling*, have most value for the social sciences (Halpern 2016; Pearl 2009).

Structural counterfactuals differentiate causal from non-causal relations *without* having to invoke reference to universal social and natural laws, which is notoriously problematic in the realm of the social sciences (George and Bennett 2005:chap. 7; Reutlinger 2013). Equally important, counterfactual structural theories do not require assuming physical processes on the level of particles and waves that are impossible to trace in a social science study. The argument that counterfactuals are valuable for causal inference has been made before for quantitative and qualitative research (for example, Brady 2008; Keele 2015; Lebow 2010). Unlike previous

arguments, however, our claim for counterfactuals embedded in structural causal theorizing is based on a comparative review of the truth conditions of five prominent theories of causation. We therefore show that counterfactual theories of causation are *superior to their alternatives* when it comes to social science research. This is not to say that counterfactual theories are without problems because every theory of causation faces challenges under some circumstances (Paul and Hall 2013:chap. 2). However, if we take theories of causation seriously, which we certainly should, we then need to recognize that some of them posit truth conditions that are much more appropriate for the social sciences than others.

The article is structured as follows. In sections 2 and 3, we introduce five theories of causation central to philosophy of causation and our article. Similar overviews have been provided before (Brady 2008; Johnson et al. 2017), but our discussion is more comprehensive because we consider a broader set of theories and social science methods and we focus specifically on the truth conditions implied by each theory. Each theory is complex, comes in different varieties and could be discussed in book-length. In our context, we have to give a stylized discussion focusing on the essence of each theory and the characteristics that are most relevant from a social science perspective. Section 4 presents the main results of our survey of causation in the social sciences to show how researchers using a variety of methods think about the relationship between candidate truth conditions and methods. In section 5, we show that the views of researchers and methodologists are partly incoherent and use philosophical and methodological arguments to suggest coherent links between the five theories of causation and the six methods of causal inference. We also explain why structural counterfactuals are the best choice given the types of evidence with which we work in the social sciences. Section 6 concludes.

2. Theories of causation and social science research

We delineate the scope of our discussion by distinguishing *ontological (*or *metaphysical), semantic* and *methodological* accounts of causation (Reutlinger 2013:4). *Ontological* theories ask questions such as what causation is; what the relata of the causal relationship are; whether causal facts are constitutive components of the physical world, or whether they can ultimately be reduced to non-causal facts such as natural laws. These are important debates in the philosophy of causation. However, we side with Woodward (2015:3594-3596) in stating that most of these issues lack *practical* implications for making causal claims in applied research outside the realm of fundamental physics. For ease of presentation and in line with common usage in the social sciences, we will speak of *events* or *phenomena* without taking sides in the debate about the relata of the causal relationship.

By contrast, we are interested in philosophers' *semantic* accounts of causation. Semantic theories clarify the *meaning* of causal claims by stating the conditions under which a causal claim is true. They provide a set of criteria that allows us to distinguish causal relationships, for which causal statements are true, from non-causal relationships for which causal statements are false. When we say that "smoking causes lung cancer" or that "the assassination of Archduke Franz Ferdinand triggered World War I", semantic accounts give answers to the question of how we should understand these causal claims and which inferential method and empirical evidence should convince us that they are true.

Finally, *methodological* theories of causation provide guidance into ways to infer the truth of causal statements on the basis of empirical data in a *particular realm* of research methods, such as, e.g. statistical analysis of observational data or experiments. Since we seek to provide a general overview of theories of causation that also has the goal of diminishing some misunderstandings between qualitative and quantitative researchers, we focus here on semantic

accounts of causation, not the methodological suggestions that discuss causal inference *within,* but not beyond a specific methods camp.[2]

Our discussion is guided by the classification of theories of causation according to the characteristics in table 1. The first column states what makes a causal claim true under each theory alongside the philosopher usually associated with the theory. The second column covers whether the theory belongs to the class of accounts invoking a criterion of *difference-making* or whether it instead focuses on *productive processes* (Paul and Hall 2013:chap. 2). From a difference-making perspective, an event counts as a cause when its presence or absence makes a difference to the presence or absence of the effect (assuming binary causes and effects here for ease of discussion). In a production view on causation, we take an event as a cause when it is connected to the outcome through a process or chain of events.[3]

The last column captures whether the theory locates the truth conditions of causal claims at the *type* or *token* level (Paul and Hall 2013:chap. 2).[4] The token-type distinction is crucial for anchoring methods because large-n methods support causal claims at the level of types and small-n, qualitative methods support claims at the level of tokens. Token-level theories define causation as a relationship between singular, instantiated events such as "Why did the democracies Great Britain and France not go to war with each other during the Fashoda crisis of 1898"? Type-level theories define causation as a relationship between classes of events; they ask, "why do democracies not go to war with other democracies"? Following the convention, moving forward, we use "x" and "y" for token-level, and "X" and "Y" for type-level causes and effects. The type-

[2] For prominent examples of methodological theories of causation, see Spirtes, Glymour and Scheines (2001) and Pearl (2009). We bring in methodological considerations when addressing *individual* methods in section five.

[3] Throughout the discussion, we use binary events as examples for ease of discussion. The same arguments hold for multicategorical events and events measured on a continuous scale.

[4] A third possibility is to define causation as a relationship between *variables* (Woodward 2003). This is often taken to be co-terminous with a type-level account. However, variables can capture type and token level phenomena. We distinguish between type and token because this is analytically more precise.

token distinction shows that it is important to distinguish *singular* cases from *single* cases. A singular case instantiates a causal relation at a specific point in time and space and has characteristics distinguishing it from any other case. A singular case is always a single case, but the reverse does not necessarily hold. We can couch a single case in a singularist perspective by taking time and space into account. Alternatively, we can simply consider a single case as one instantiation of a type-level cause-effect relationship, abstracting away from time and space (Baumgartner 2008:342-348). For example, we might say that "large numbers of refugees entering the European Union cause nationalist policy-responses and the breakdown of intergovernmental policy-making". This causal claim is made on the type-level yet, so far, there is only a single case instantiating it. However, it is not a claim about a singular case because it does not specifically say that "a large number of refugees entering the European Union caused nationalist policy-responses and the breakdown of intergovernmental policy-making *in 2015*".

Whereas the choice of a type or token level causal claim implies whether large- or small-N evidence is needed to infer whether the causal statement does or does not hold true, it does not imply that the respective other level becomes irrelevant because it also performs a valuable function for the overall research process of which causal inference is just one component (see figure 1). If we follow a token theory of causation, we can still arrive at causal claims about type-level relationships, but we do so by *generalizing* over causal claims about token-level phenomena (Glennan 2011:789). Conversely, type-level theories accept that one can make singular causal claims, but only if we conceive of them as instantiations of single cases which again are instantiations of the type-level relationship (Baumgartner 2008:347). In this sense, type-level and token-level theories "advance our knowledge of the general and the specific simultaneously" (Waldner 2015:240). However, the criteria for establishing whether one and the same causal claim holds true are located on either the token or the type level, but not on both.

Table 1. Truth conditions, primary causal focus and level of major theories of causation

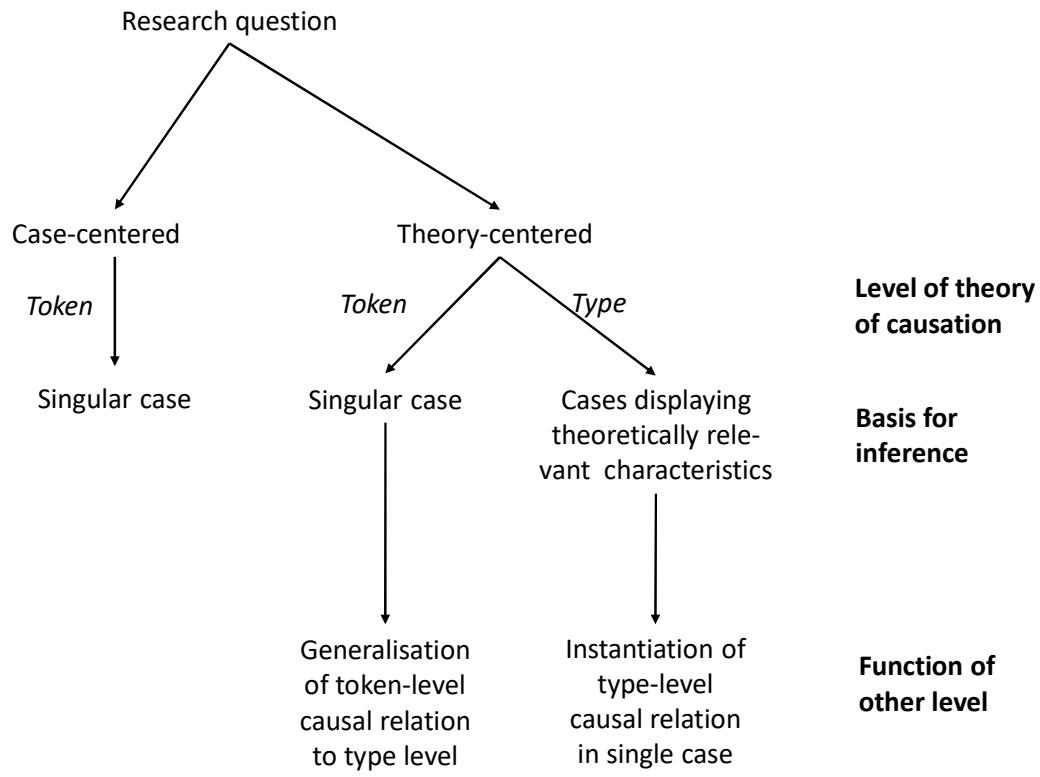
Truth condition (key philosopher)	**Primary causal focus**	**Level**
Regularity (Hume, Mackie, Baumgartner)	Difference-making	Type
Probability raising (early Salmon, Cartwright)	Difference-making	Type
Token-level counterfactual (Lewis)	Difference-making	Token
Intervention/type-level counterfactual (Woodward)	Difference-making	Type
Process (Dowe)	Production	Token

In figure 1, we highlight the implications of these distinctions for empirical research. In *case-centered research* aimed at formulating a comprehensive explanation of a singular case, we are dealing with token-level causal claims and correspondingly with the set of token-level theories of causation discussing the truth condition of claims relating a singular cause to a singular effect. Framing the case as singular denies subsuming it under types of events. In *theory-centered research* aimed at establishing general theory, we can formulate and empirically assess both type and token-level causal claims. Depending on which claim we seek to assess, we then need to collect type- or token-level evidence to draw causal inferences about the validity of our causal claim.

It is important to emphasize that it is the nature of the causal claims, not the research question, which leads to a token- or type-level inferential method. We can formulate a type-level research question such as "Why do democracies not fight each other?" and proceed with a

qualitative study linked to a token theory of causation. For example, Owen (1994) motivates his analysis with a type-level research interest, the democratic peace phenomenon, but draws causal inferences on the token level through multiple process tracing case studies. This does not create a contradiction because he generalizes the qualitative insights to the type-level in a second step. Conversely, we can start with a puzzle posed by a singular case, but consider this case only as an instantiation of a relationship between classes of events and draw type-level causal inferences on the basis of type-level evidence.

Figure 1: Research questions, levels and causal claims



3. Five theories of causation

3.1. Regularity theories

Discussions of causation often take Hume's *regularity account of constant conjunction* (2003) as their point of departure (for example, Brady 2008). Hume specifies three criteria causal relationships must meet: X temporally precedes Y; cause and effect are proximate in time and space (spatio-temporal contiguity); and the effect regularly follows the cause. These criteria apply to the *type level* because we can only require recurring sequences of X and Y when we look at classes of events. To introduce our running example, the claim that "granting territorial autonomy to minorities (X) causes secessionism (Y)" is true for the regularity theorist if and only if each and every time a minority somewhere in the world is granted territorial autonomy, secessionism grows. The researcher investigating this claim would have to look for evidence that granting and not granting autonomy to minorities makes a difference for whether they become secessionist (see Baumgartner 2013:90); all minorities who were granted autonomy are secessionist, those who did not receive autonomy are not secessionist.

Regularities are not sufficient for causation because many non-causal associations meet Hume's criteria. In light of this, the constant conjunction theory has been revised in many ways since its inception (Baumgartner 2008). What is probably the most widely known revision is Mackie's (1974) regularity theory, wherein X is a cause of Y if and only if X is an INUS condition of Y. INUS stands for "an *insufficient* but *non-redundant* part of an *unnecessary* but *sufficient* condition" (Mackie 1974:62).[5] However, whether it does better than its predecessor when it comes to distinguishing causal from non-causal regularities remains questionable (Baumgartner 2013:85-87). Ultimately, both Hume and Mackie must ground regular relations in

[5] Mackie's own example is the claim that a short circuit caused a fire. The short circuit was a necessary part of a conjunction of conditions (that is, it was not sufficient by itself; it had to be at least combined with flammable material). This conjunction was sufficient for causing the fire, but not necessary since fires can also be caused by cigarette butts and flammable material, among a host of other possibilities.

universal causal laws to make this distinction: "thus if we said that a short-circuit here was a necessary condition for a fire in this house, we should be saying that there are true universal propositions from which, together with true statements about the characteristics of this house, and together with the supposition that a short-circuit did not occur here, it would follow that the house did not catch fire." (Mackie 1965: 254, see also Kim 1971: 429). This creates problems in applying this theory to the social sciences, where universal laws are hard to come by (Reutlinger 2013:17).

3.2. Probability-raising theories

At first glance, *probabilistic theories of causation* fare better in this regard. For the probability theorist, causes do not necessitate their effects, as argued by regularity theory, but merely make them more likely (Salmon 1998:chap. 14).[6] X causes Y if and only if the conditional probability of Y given X is higher than the conditional probability of Y when X is not given ($p(Y|X) > p(Y|\text{not-}X)$). The type-level claim that "granting territorial autonomy causes secessionism" can then be true even if there are some minorities who received autonomy and did not increase their secessionism as long as the conditional probability of secessionism *given* autonomy is higher than the probability of secessionism *without* territorial autonomy. In light of this criterion, our researcher would need to look for evidence on a representative sample of minorities and see whether a probabilistic relationship between autonomy and secessionism is in place.

Probability theories can account for systematic associations that fall short of universal, law-like regularities. This is an important advantage when it comes to establishing the truth conditions for social scientists' causal claims, which are often of a probabilistic nature and cannot rely on the existence of deterministic social laws. However, Simpson's paradox illustrates that

[6] For a theory of causation that provides a unified framework for deterministic *and* probabilistic relations, see Spohn's (2006) ranking theory.

probability-raising is not necessary for causation because it shows that an event can lower the probability of Y and still be a cause. The paradox states that any causal relationship between X and Y that holds in a given population can be reversed in a sub-population if a variable correlates with both X and Y. Simple probability-raising theories also fail to offer sufficient criteria for causation. In many instances, including the notorious common-cause problem that is central to observational research, X raises the probability of Y without being its cause. Revised versions of probabilism therefore apply a stricter definition of *ceteris paribus* conditions that hold common causes constant.[7] These developments approximate advanced probabilistic accounts to Woodward's (2003) interventionist theory of causation that draws on type-level counterfactuals to clarify what it means for x to cause y. Counterfactual theories solve most of the problems of regularity and probability theories. They come in both a token- and a type-level variant. Keeping with the order in which they were developed, we discuss the token-level version first.

3.3. Counterfactual dependence and possible worlds

David Lewis (1973b) proposed the counterfactual theory of causation when regularity accounts were still dominant, motivated in particular by regularity theories' failure to distinguish causal from accidental relations. Lewis radically departs from regularities by defining causation as counterfactual dependence at the token level. The basic and intuitive counterfactual definition of a cause is: a singular event x causes a singular event y if and only if had x not occurred, y would not have occurred. For our running example, we now have to replace the type-level claim with a token-level claim, e.g. "granting a new autonomy statute to the Catalans in 2006 (x) caused increased support for secessionist Catalan elites in the elections to the Catalan parliament in 2010 (y)". This claim is true if we can establish that had the Catalans *not* received a new autonomy

[7] For example, this is captured by Cartwright's idea of test situations (1979).

statute in 2006, support for secessionist elites would *not* have risen in 2010. In contrast to the theories previously discussed, large-N evidence about multiple minorities is now irrelevant to establishing causation. Instead, the researcher needs to devise *a possible world* in which Catalans have not received the new autonomy statute and consider whether the Catalans in this possible world would have become secessionist or not (see section 5.2).[8]

Casting causation in terms of counterfactual dependence successfully separates causal relations from mere coincidences and common-cause constellations. If x and y coincide without being causally related, or if both have a common cause, the counterfactual "had x not occurred, y would not have occurred" is simply not true. This crucial advantage of counterfactual dependence comes at a prize. The ability to distinguish causal from non-causal relations requires assessing what would have happened in the *most similar possible world* to the actual world for which we want to make a causal argument (Lewis 1973a:559). A challenge is that we cannot use empirical evidence to reason about what would have happened in a possible world. More fundamentally, cause-effect relationships in the actual world are governed by deterministic natural laws, according to Lewis. But if this is the case, how could things have been different at all? In Lewis' account, a possible world must differ from the actual world either in facts or because the natural laws characterizing the actual world are violated (Lewis 1973a:560). It requires at least a "small miracle" to construct a possible world that differs in some respect from the actual one (Lewis 1979). The degree of difference between the actual world and possible worlds in terms of facts and miracles is important for establishing similarity rankings and choose the most similar possible world needed to assess whether counterfactual dependence obtains.

We need to address one additional issue in anticipation of our discussion of process

[8] For the single case of Catalonia, large-n evidence of relevant observable implications could be instrumental for constructing the counterfactual. For example, one could cite results of a survey of Catalans about how the conferral of autonomy strengthened their identity as Catalans because this can be theorized as an intermediate step toward support of secessionist elites.

tracing in section 5.2. Lewis notes that causal relations are *transitive*, meaning that if x is a cause of z and z a cause of y, then x is also a cause of y. However, counterfactual dependence between events is not necessarily transitive (for example in cases of double prevention) (Paul and Hall 2013:chap. 5). Lewis addresses this problem by requiring that a *transitive causal chain* links x and y, and in this chain counterfactual dependence holds between each individual link. If z counterfactually depends on x and y counterfactually depends on z, then x is a cause of y even without counterfactual dependence between x and y (Lewis 1973a:563). Counterfactual analysis can thereby be extended to chains of events and does not imply "black-boxing" the process connecting causes to their effects.

3.4. Counterfactual dependence and interventions

Woodward's (2003) *interventionism* takes Lewis' token-level counterfactuals to the type level. Taking his starting point in methodological theories of causation that are primarily concerned with identifying causal relations in a quantitative framework (Pearl 2009; Spirtes et al. 2001), Woodward elucidates the sense in which the relations that these theories take to be primitive are causal (Woodward 2003:38). This theory has received considerable appraisal since its inception and is central to current discussions on causation, mechanisms, and explanations (for example, Glennan 2011).

Woodward analyses causation as a relationship between variables. He states that the type-level claim "X causes Y" is true if there is an *intervention* on the variable X that will change the values of the variable Y (Woodward 2003:55): "A necessary and sufficient condition for X to be a direct cause of Y with respect to a variable set V is that there be a possible intervention on X that will change Y (or the probability distribution of Y) when all other variables in V besides X and Y are held fixed at some value by interventions." (Woodward 2003:59) Woodward explains that

interventions need not be man-made and can also occur naturally. The claim that "granting territorial autonomy causes secessionism" is true for the interventionist, if territorial autonomy is or could be distributed through an exogenous process across minorities and we observe that secessionism grows among those minorities that received autonomy, while holding all other causally relevant factors fixed.

Like in Lewis' token-level account, counterfactual dependence between cause and effect is a sufficient criterion for causation. Where Lewis resorted to possible worlds to characterize the relevant token-level counterfactuals, Woodward takes inspiration from Pearl and captures the relevant type-level counterfactuals with structural equations that map possible values of Y onto possible values of X. The relevant counterfactuals now relate variables instead of singular events. If the requirements of the interventionist theory hold, causal relations between variables describe "what the response of *Y would be if* a certain sort of change in the value of *X* were to occur" (Woodward 2003:39, original emphasis).

3.5. Process theories

Unlike the theories presented so far, process or physical theories of causation do not aim to clarify the meaning of the concept of causation as we use it in ordinary and scientific causal claims. Instead, they provide an ontological account of what causal processes *are* in light of what science tells us about the world (Dowe 2000:1-11). Nonetheless, we integrate process theories into our discussion, because some qualitative methodologists have argued that they provide an adequate conceptual basis for process tracing as a method of inference (for example George and Bennett 2005:chap. 7). Process theories dispense with the idea that causation requires difference-making and see causation as *production* instead. Under the process account, causation requires a spatiotemporally continuous process transmitting some physical structure from a token cause to

its token effect. Process theories are located at the token level because this is where the processes unfold. Whether a *type* of process is causal then depends on *generalization* over the singular causal processes (Dowe 2000:96). The token claim in our example is then true, if there is a continuous *physical connection* initiated by the signing of the autonomy statute in 2006 that ended with Catalans voting in favour of secessionist parties in 2010. This means that our researcher now must turn to evidence based on process observations.

According to Dowe's process theory (2000:210-215), a process is causal if it manifests a conserved quantity, while an interaction between two processes is causal if one process exchanges a conserved quantity with another process. When we throw a stone at a window, our arm first transfers a conserved quantity, namely energy, to the stone, which transfers its energy to the window, which shatters as a consequence of the exchange. While this notion of causation seems intuitively plausible, the theory has several disadvantages. First, it struggles to reliably distinguish causal processes from non-causal processes (Hausman 2002:720-721). Dowe resorts to the law of conservation to distinguish causal from non-causal processes, yet the central notion of conserved quantities makes it difficult to imagine the application of the process theory beyond physics and, even in that domain, its applicability has been questioned (Hausman 2002:718-719). Using this theory for causal claims and causal explanation in the social sciences seems particularly farfetched (Woodward 2003:354). Dowe leaves open what the particular conserved quantities are, but states that current scientific theory is our best guide towards them and mentions mass-energy, linear momentum, and charge as plausible candidates (Dowe 2000:94). It seems highly unlikely, if not impossible, to describe the fine-grained exchange of quantities involved in causing any but the most basic physical phenomenon (Dowe 2000:353-354).

Second, any theory of physical causation is unable to subscribe to the idea of *causation by omission* (Barros 2013:458-459). It is given when the *absence* of a phenomenon is linked to the

occurrence of an outcome. A classic example in philosophy is that not watering a plant causes the plant to die, yet there is no physical connection between not-watering and the plant's dying. The inability to account for causation by omission is an exclusive problem of process theories because all other theories we have addressed can accommodate such situations.[9]

We summarize our reasoning in this section in table 2. Using our running example, we relate each theory of causation to a corresponding token- or type-level causal statement and the evidence needed to confirm the statement. The example once again highlights that causal statements about *types* of events have truth conditions that require us to collect *type*-level evidence for drawing causal inferences, whereas *token*-level evidence is required to judge whether causal statements about *singular* events hold true.

Table 2. Causal statements and required positive evidence by theory of causation

Theory	**Causal statement**	**Evidence**
Regularity (type)	Granting territorial autonomy to minorities causes secessionism.	Territorial autonomy for minorities is always followed by secessionism while no autonomy is followed by non-secessionism.
Probability raising (type)	Granting territorial autonomy to minorities causes secessionism.	The probability of secessionism is higher for minorities with territorial autonomy than for minorities without territorial autonomy.
Token-level counterfactual (token)	Granting a new autonomy statute to the Catalans in 2006 caused an increase in support for secessionist elites in 2010.	In the closest possible world where the Catalans did not receive autonomy, their support for secessionist elites did not increase.

[9] Dowe (2000:chap. 6) tries to circumvent this problem by arguing that cases of omission do not in fact involve genuine causation at all.

Type-level counterfactual (type)	Granting territorial autonomy to minorities causes secessionism.	After having distributed autonomy among minorities through an exogenous causal process, secessionism grows among minorities that received autonomy.
Process (token)	Granting a new autonomy statute to the Catalans in 2006 caused an increase in support for secessionist elites in 2010.	Physical processes connect the new autonomy statute of 2006 to voters' choices in favor of secessionist elites in 2010.

The next section assesses the extent to which these insights are reflected in how social science researchers and methodologists think about causation and how they relate the candidate truth conditions of causal claims we discussed to popular social science methods of causal inference: *experiments*, *statistical analysis of observational data*, *Qualitative Comparative Analysis* (QCA), *comparative case studies*, *process tracing,* as well as *multi-method research* (MMR) that combines a large and a small-N method. We then contrast how researchers link truth conditions to methods with our own arguments about coherent and incoherent linkages based on philosophical and methodological considerations.

4. Linking truth conditions to social science methods: a survey of social scientists

Between July 2017 and October 2017, we conducted a "Practitioners' Survey on Causation in the Social Sciences".[10] We identified empirical researchers using all methods discussed in this article based on their publications with methods listed as 'Topics' in Web of Science.[11] In addition, we selected methodologists working on each of the six methods to also capture the views of experts.

[10] A detailed description of the survey can be found in the appendix to this article.

[11] Researchers who mention the used method in the title, abstract or as a keyword might be systematically more interested in methods than researchers who don't. Interest in methods might in turn be correlated with theories of causation, which would bias the analysis against finding the kinds of uncertainties we describe below.

The survey then asked respondents for the method or methods they primarily use in their research to validate our initial assignment of methods.[12] 109 respondents answered our first, open question about their understanding of causation and 94 moved on to answer also the second, closed question on how one can link methods to theories of causation. The small number of respondents, their distribution across the six methods and the sample selection method limit the degree to which we can generalize insights from the survey. Within these limits, however, the responses illustrate the variety of perspectives on causation within the discipline, as well as the uncertainties surrounding researchers' views on the relationship between the truth conditions of causal claims and methods adequate for assessing these claims empirically.

To reveal respondents' own definition of causation, we asked the following open question:

> *Causal inference in the social sciences means determining whether an empirical association between X and Y is causal. How would you define causation? What is your criterion for determining that X is a cause Y?*

We used qualitative content analysis to code 85 users' and 24 experts' answers to the open question.[13] The coding frame consisted of two broads sets of categories. The first set contained the philosophical theories of causation presented above plus a residual category called "unclear" for answers that did not fit either theory. The second set of categories captured the distinction between type- and token-level, again supplemented with a residual category "unclear" if the answers did not tell us whether causation was defined as a type- or a token-level relation.

A first finding is that the majority of answers make no explicit reference to the type- or the

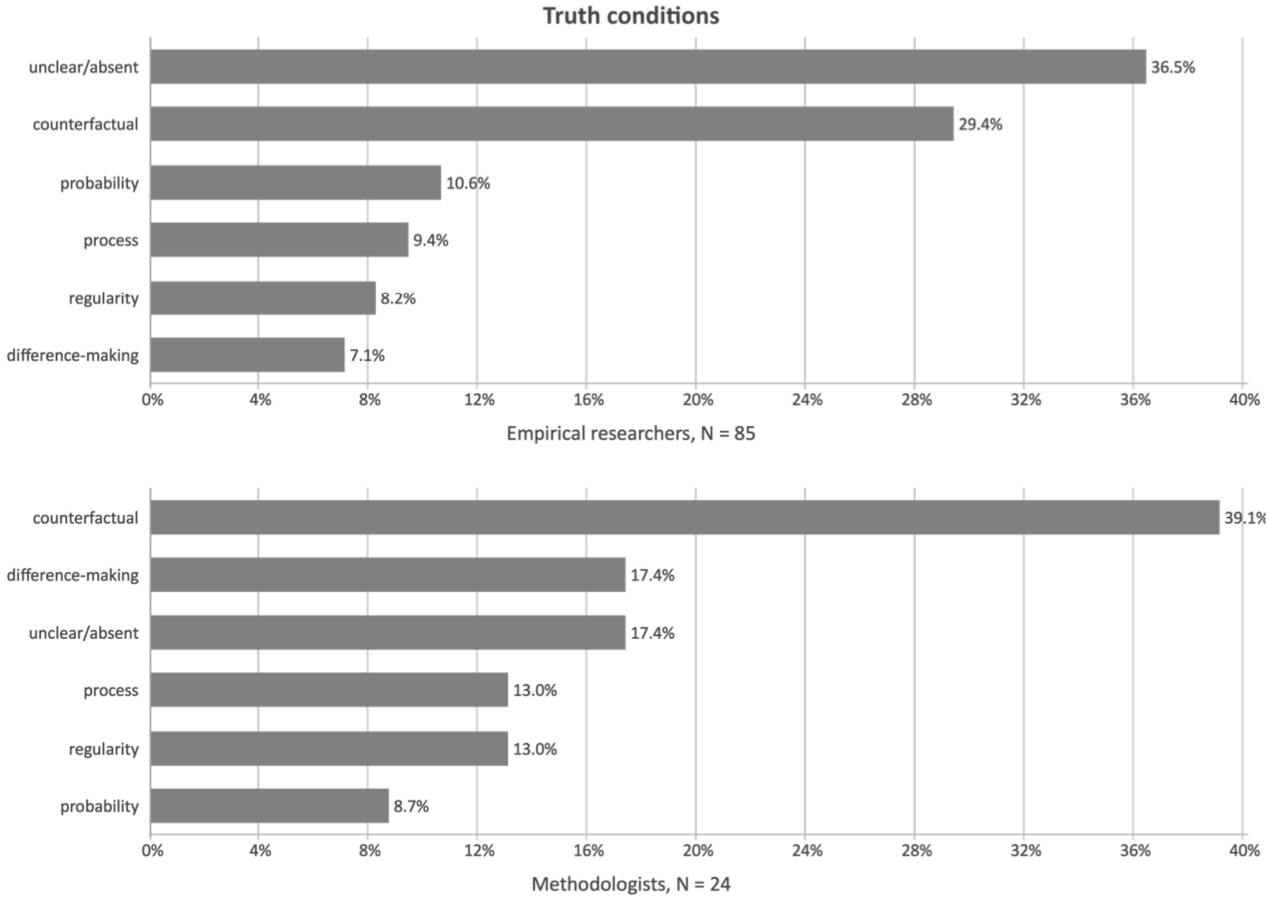
[12] Publication-based methods correspond to self-assigned methods for 71% of respondents.

[13] The answers were coded by two coders following the coding manual in the appendix. Inter-coder reliability testing showed convergence in applying the philosophical categories for 78 of 109 answers. Coders then sat together and discussed the remaining 31 answers resolving all controversies. A coding protocol is available upon request.

token-level. This suggests that respondents do not see this crucial philosophical distinction as an important characteristic of their concept of causation. For the theories of causation, a majority of empirical researchers presented answers that we coded as "unclear" (see figure 2 below), either because the answer mentioned criteria other than those suggested by any theory of causation or because criteria from different theories were being mixed up. Counterfactual accounts of causation were dominant among those responses for which we could identify clear criteria and among the group of methodologists.

The dominance of a counterfactual concept of causation might reflect the growing prominence of experimental and quasi-experimental methods of causal inference in the social sciences, which finds its counterpart in the popularity of counterfactual theories of causation in the philosophy of science. Probabilism — the most active field in the philosophy of causation before type-level counterfactual theories became available — continues to be a bit more popular among users than among methodologists, where it is the least frequently invoked concept of causation.

Figure 2. Classifying respondents' own concepts of causation



After respondents had defined their concept of causation, we introduced a table listing the truth conditions of five theories of causation in the lines, and the six methods of causal inference in the columns.[14] We asked participants whether they believed that a given method meets the criterion for causal inference implied by each theory of causation. For each method, a participant could check none, one, or multiple theories of causation, respectively tick "other" (a criterion not listed) or "don't know". Figure 3 summarizes answers to the closed-question by method and standardizing the responses so that the shares add up to 1. [15] For example, about 20 percent of all

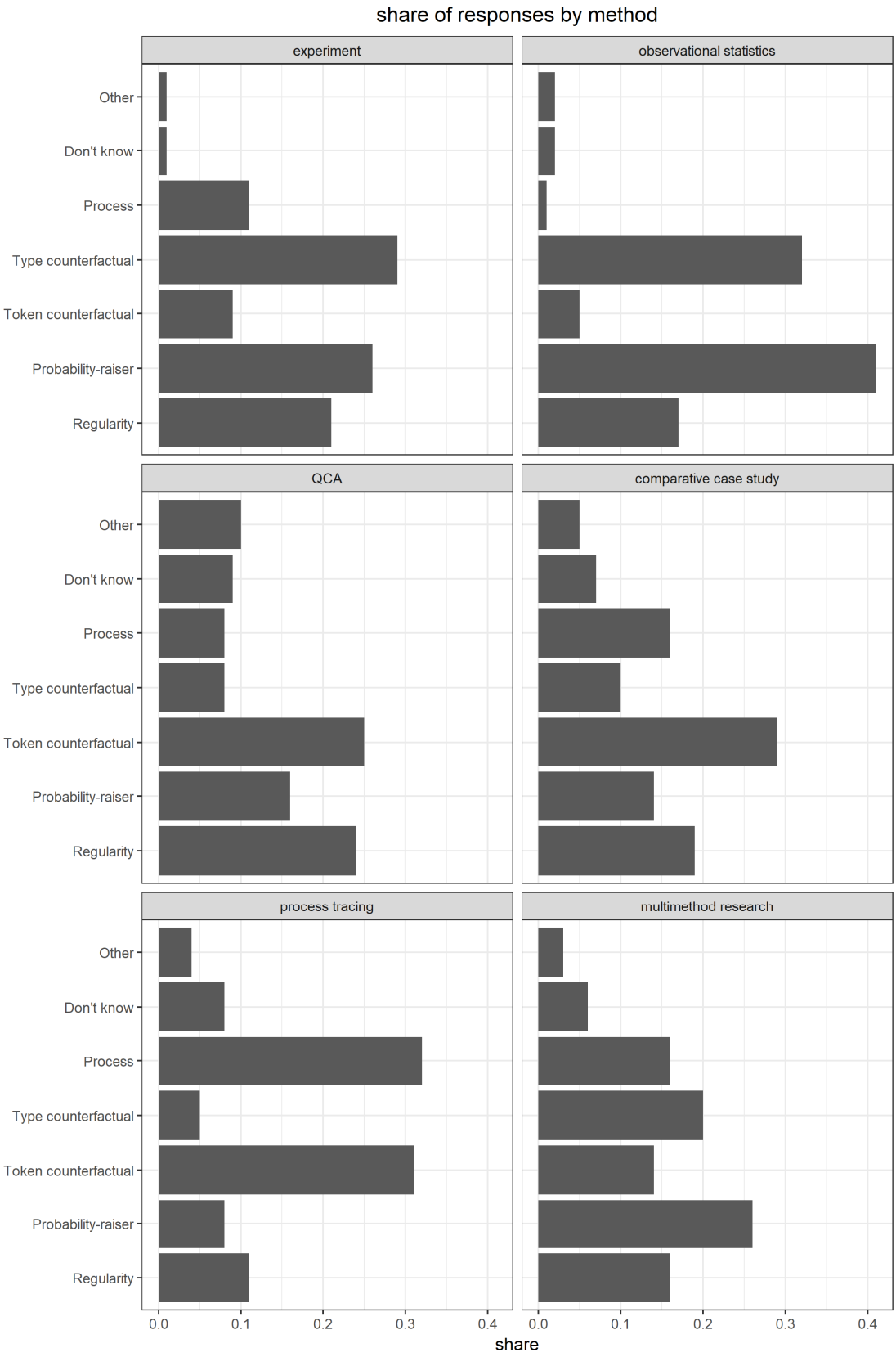
[14] See the appendix for the exact table and wording. We did not mention the name of the theory of causation to avoid priming effects. Participants were not able to return to their answers to open question 1 after they had seen the table in question 2.

[15] In additional analyses we disaggregated the findings by users *versus* methodologists and by those who work with a method *versus* those who do not. The patterns remain similar (see Appendix).

linkages were made between experiments and the truth condition of regularity theories, reflecting the belief that experiments can be used to assess whether regularities obtain.

The main insight from this analysis is the diversity of perspectives on which causal claims can be evaluated with which method. Each method is linked to at least two criteria with a share of 20 percent or more. This shows that researchers are ambiguous about the exact type of causal claim that each method can substantiate. To pick one example, about 30 percent of all responses state that process tracing substantiates causal claims by assessing whether physical processes obtain. The same share of responses, however, establishes a link between process tracing and token counterfactuals, thus believing that process tracing substantiates causal claims through difference-making. Just knowing that someone applies process tracing therefore does not allow us to infer what type of causal claim she made and whether she used (and should have used) *process* evidence or (counterfactual or comparative) *difference-making* evidence in her study. The practical consequence is that empirical scholars should become aware of and be transparent about the truth conditions of their causal claims. Only mentioning the method can breed confusion about *how* inferences were made and about what would count as valid confirming and disconfirming evidence in a follow-up study on the same causal relationship.

Figure 3: Results of the survey for closed question (*n*=94)


share of responses by method
experiment
observational statistics
QCA
comparative case study
process tracing
multimethod research
Other
Don't know
Process
Type counterfactual
Token counterfactual
Probability-raiser
Regularity
0.0
0.1
0.2
0.3
0.4
share

5. Linking truth conditions to social science methods: state of the art and recommendations

The descriptive insights reflect the heterogeneity of views on how to infer causation among social scientists and methodologists. They also show that researchers are unaware of the crucial distinction between type and token level truth conditions of causal claims. In this section, we now move on to differentiating coherent from incoherent linkages between causal claims, truth conditions and social science methods. We structure the discussion of methods along the common distinction between quantitative and qualitative methods (counting QCA as qualitative) and discuss MMR on its own terms. Based on the philosophical discussion, we argue that the minimum requirement is to match the level on which the truth condition is located with the level on which a method processes evidence: type-level theories have to be combined with large-N methods and token-level theories with small-N methods. This leaves more than one possibility for each method because there are several type- and token-level theories of causation. For each method, we therefore review the methodological state of the art on causal inference in order to single out the theory of causation that is considered most adequate for empirical research in the social sciences, recommending structural counterfactual theories of the token and type variant over their alternatives.

5.1 Theories of causation in quantitative research

Experiments

Experiments are probably the method for which it is easiest to determine the adequate theory of causation. Woodward's interventionist theory mirrors experiments very closely and forms the basis of current experimental research (Baetu 2016). Both of the other type-level theories, i.e. regularity and probability theory, are also plausible candidates. Advanced versions of probabilism require conditions that resemble Woodward's point that x be brought about by an intervention.

Since interventionism resembles the experimental logic — that is nowadays also becoming the gold standard for statistical inference of observational data (Keele 2015) — using this theory instead of any version of probabilism is straightforward (Woodward 2003:33). The random assignment of units to the treatment and control groups represents an intervention, which is precisely the necessary and sufficient condition for causation in Woodward's theory. In fact, Woodward repeatedly refers to experiments in his exposition of the interventionist theory and relates his hypothetical interventions to actual experiments (Woodward 2003:chap. 1).

Furthermore, it is worth noting that the potential outcomes framework, the dominant methodological framework for experimental research, starts with the *individual treatment effect* (ITE), that is, it takes the token level as the starting point and asks for an individual's value on the outcome when it does and does not receive the treatment. The ITE cannot be estimated because we can never observe one and the same unit receiving and not receiving treatment. Because of this fundamental problem of causal inference, the potential outcomes framework moves on to the type level. It formulates the average treatment effect (ATE) or another treatment effect of interest and derives the identification assumptions that must be met to get a valid estimate of a variable's effect (Morgan and Winship 2014:43-46). This is in line with the interventionist theory because it centers on type-level relationships and, at the same time, argues that they are *realized* on the token level (Woodward 2003:40). By contrast, if we linked token-level counterfactuals to experiments (as do about 10 percent of survey responses), we would have to look for causal relationships at the level of each individual participant of an experiment. This is not what the experiment gives us because we can only estimate average treatment effects on the type level.

Inferential statistics with observational data

The statistical analysis of observational data used to be anchored in probabilistic theories of

causation (Salmon 1998:chap. 13). This seems to be natural, given that much of statistics is about the calculation of probabilities. However, just like experiments, inferential statistics can currently also be anchored in type-level counterfactual accounts of causation (Keele 2015:314), like Woodward's interventionism. Pearl's (2009) account of directed acyclic graphs (DAG), which is a methodological theory of causation, provides the link between the philosophical, semantic theory of interventionist causation and observational inferential statistics. DAGs can be loosely defined as causal maps of our theoretical expectations formalized in structural equations. In the ideal setting, the DAG shows us that a causal effect is non-parametrically identified and that it can, in principle, be correctly estimated (Morgan and Winship 2014:chap. 3).[16] Alternatively, we might exploit a natural intervention to identify and estimate marginal effects without having to rely on DAGs (Keele 2015). Type-level counterfactuals can therefore be seen as the unifying element of interventionist theories of causation, DAGs and the potential outcomes framework underpinning experiments and inferential statistics of observational data.

The counterfactual foundations of modern inferential statistics are worth emphasizing because quantitative research is often misrepresented by qualitative methodologists in this regard. This can create unnecessary divides between scholars working with different methods. It is sometimes argued that inferential statistics invoke a regularity (Beach and Pedersen 2013:29) or a probabilistic understanding of causation (Goertz and Mahoney 2012:45). In principle, both of these type-level theories can be plausibly linked to statistical methods. However, the current literature converges on type-level counterfactuals. Analyzing evidence on the type level therefore need not imply a regularity understanding and estimating marginal effects in terms of probabilities need not imply that probability-raising is the criterion used to distinguish a causal from a non-causal relation. This shows how an awareness of the truth conditions of causal claims

[16] One does not necessarily estimate the true effect because issues related to estimation such as sampling variability might interfere.

can avoid misunderstandings between scholars working with different methods.[17]

5.2 Theories of causation in qualitative research

Qualitative Comparative Analysis (QCA)

QCA is a technique that processes cases with the aim of determining relations of *sufficiency* and *necessity* between one or multiple terms and an outcome.[18] A term is sufficient if its presence is associated with the presence of the outcome, and a term is necessary if the outcome is present only if the term is present. Set relations are not causally interpretable *per se*, but can be interpreted in causal terms if couched in a regularity theory of causation. Semantically, the definitions of sufficient and necessary relations entail a regularity understanding because they mandate that the outcome is *always*, not just *mostly* given when the term is present (sufficiency), and that the term is always present when the outcome is present (necessity).

Underpinning QCA with regularities is most visible in Baumgartner's (2009) work. He formulates a revised version of Mackie's INUS theory to develop the coincidence analysis algorithm (CNA) that is used to derive a solution from a truth table. A central idea that guides the generation of a QCA solution is '*non-redundancy'*. A condition is non-redundant or a non-redundant element of a configuration if its presence and absence makes a difference to the presence and absence of the outcome. Only conditions that are non-redundant according to the CNA algorithm qualify as causal, meaning that regularity theories incorporate an element of difference-making (Paul and Hall 2013:chap 2).

In contrast to CNA, the causal foundation of the more widely used *Quine-McCluskey*

[17] Our discussion points to a problem of Mahoney's '*Unified theory of causality*' (2008). He neglects that inferential statistics (what he calls "population-oriented research") has a sound basis in Woodward's interventionist theory and does not require a regularity theory of causation. Contrary to Mahoney, type-level theories of causation also do acknowledge that causation *is* instantiated in individual cases (Runhardt 2015:1297).

[18] A term can be a single condition or a conjunction of conditions (also called a configuration), or a disjunction of terms.

(QMc) algorithm has, to the best of our knowledge, never been discussed. QMc follows a different routine for determining non-redundant conditions (Baumgartner 2009). However, this is not important here because they share the goal of inferring that necessary and sufficient relations are causal and define non-redundancy as making a difference to the outcome. We argue here that because of the overall interest in set relations, QCA using QMc is also best anchored in a regularity theory of causation.[19]

Comparative case studies

Comparative case studies, understood as the classic pairwise comparison following the idea of the most-similar design (for example, Slater and Ziblatt 2013), have received considerable attention over the last decades. However, the debate has centered on technical aspects, neglecting the fundamental question about the underlying theory of causation (for example, Rohlfing 2012:chap. 4, 5). A token-level counterfactual theory is most reasonable because a counterfactual aims to determine whether the outcome counterfactually depends on the cause, which is identical to the inferential goal of a most-similar comparison. Similarly to Woodward's type-level counterfactuals and large-n research, counterfactual truth conditions at the token level do not require constructing a counterfactual to assess a claim empirically. Instead, we can use the criteria of counterfactual theory to determine the best possible *empirical* case for comparison.[20]

However, we have seen in section 3.3 that Lewis' original account of similarity rankings

[19] A complication concerns the use of remainders (that is, truth table rows without cases) because they require counterfactuals that do not sit easily with a regularity theory. We leave this issue aside here. We also leave aside alternative proposals that see QCA as generalizing over token-level causal claims, because they do not use QCA for causal inference (Rohlfing and Schneider 2018). Inquiries into sufficiency and conjunctions have also been made in the context of a type-level counterfactual framework (VanderWeele and James M. 2008), but no coherent QCA framework has developed out of this yet.

[20] A reviewer wondered whether comparative case studies could not also be rooted in a type-level regularity theory. In principle, this is possible because a type-level analysis can be done with any number of cases (Hofmann and Baumgartner 2011). However, the analysis would then confront the problem of having to presuppose that universal type level regularities hold true.

between possible worlds is problematic because his theory assumes that deterministic natural laws govern relationships in the actual world and requires at least minor deviations from facts and natural laws to construct possible worlds. The reliance on natural laws and "miracles" as deviations from them makes Lewis' theory of counterfactual dependence unattractive for the social sciences.

Instead, we can again use a structural causal modeling framework inspired by Woodward and Pearl, but this time following Halpern's (2016) token level variant.[21] The causal model now formalizes our theoretical knowledge or expectations about how causes are related to each other and to the outcome *in a singular case*. An exogenous intervention on a cause, rather than possible worlds, is what creates a difference between the actual case and a counterfactual case (Halpern 2016:71).[22]

When we use a structural causal model, the intervention on a cause becomes the basis for determining the best possible case for an empirical comparison.[23] In addition to solving the problem with Lewis' theory, it improves the comparison by forcing the researcher to ask herself: why is there the difference in the (purported) cause in the first place? Is it because of confounding, which would be a problem, or because of exogenous variation? An example of choosing two cases that differ because of exogenous variation is Hale's (2011) analysis of the impact of formal versus informal power-sharing on democratization. Hale identifies a case of formal power-sharing in the Ukraine and of informal power-sharing in Kyrgyzstan. According to Hale, the two cases are most similar to each other. They only differ in their type of power sharing and Hale explains that this difference is attributable to an exogenous factor that is unrelated to democratization. In Ukraine, a constitution mandating formal power-sharing was available at the

[21] Woodward (2003:section 2.7) also discusses how type-level structural models can be used to analyse token causation.
[22] As for the type-level equivalent, a token-level structural model is not reductionist.
[23] Whether this case is available is a separate matter.

time of the revolution in 2005 and Hale explains in detail that the availability of the constitution was independent of the process and outcome of the revolution. No such draft constitution had been available in Kyrgyzstan. Although Hale does not explicitly use a theory of counterfactual dependence, his study is instructive for the value of the idea of exogenous interventions for comparative case studies.

Process tracing

At first glance, the "natural" theory of causation for process tracing seems to be a process theory because both are concerned with processes operative in singular cases (we turn to "mechanistic theories" as a second, seemingly plausible alternative below). Relating process tracing to a physical theory would pick up on the forcefully made claim that we have to follow an asymmetric understanding of causation in process tracing (Beach and Pedersen 2013, 2016; Blatter and Haverland 2012:chap. 1), because process theory sees causation as production, not difference-making.

There are two reasons why process tracing should not be anchored in a physical theory of causation. First, physical theories capture causal processes at the level of physics where conserved quantities such as energy or charge are being exchanged. It seems epistemically impossible to go from such physical processes to social processes, for example, accounting for the democratic peace phenomenon (Woodward 2003:350-356; Kutach 2014:49-50). Critical realist definitions of causal mechanisms refer to the transfer of energy or information between agents or entities (George and Bennett 2005:chap. 7), or the transmission of causal forces from x to y (Beach and Pedersen 2013:29). These accounts of processes appear very similar to process theories of causation. However, how exactly such an "information transmission" could be seen as

causal if the truth conditions are conserved quantities remains questionable.[24] Furthermore, it seems that the problem of distinguishing causal from pseudo processes would re-emerge since Dowe solved it only via establishing exchange of conserved quantities in a closed system where the law of conservation holds, something that seems unlikely if the quantity exchanged is information.

Second, even if we granted that a "socialized" physical theory were available, the problem of accounting for causation by 'omission' would remain (see above). Marrying process tracing to a process theory of causation would make it impossible to assess claims such as "the US intelligence agencies *failed* in intervening in the preparations of the 9/11 attacks because of insufficient inter-agency cooperation". We cannot trace the exchange of energy, information or anything else between the intelligence agencies because they did not interact, and we cannot trace an interaction between the agencies and the terrorists because they did not interact. This is an important shortcoming because there is no fix available beyond declaring omissions as non-causal (see above). In contrast, a counterfactual theory can handle omissions and does not require tracing a physical exchange of conserved quantities. We therefore consider it most appropriate to anchor process tracing in a token-level theory of counterfactuals (Waldner 2014, 2015).

Lewis' take on the transitivity problem highlights how this is achieved on an abstract level. For every component of the mechanism, be it called 'entity' and 'activity' as Machamer, Darden and Craver do (2000), or something else, we evaluate whether the process would be different from the observed process at time t+1 if one element had been different at time t. If we can counterfactually confirm this for each segment of the process, we can conclude that we have

[24] The level at which an explanation bottoms out depends on the research question and can be at the level of individuals, groups or macro entities (Machamer et al. 2000). As we explain below, this matter should be kept separate from the question of how we infer that a mechanism is causal. (We thank a reviewer for asking about the relationship between inferences and bottoming out.) It might be possible to adapt Dowe's process theory to process tracing (or the other way round), but this requires work on this topic that still needs to be done.

traced a *causal* mechanism, rather than a mere sequence of events. Recent calls for increasing the quality of causal inference from process tracing through *comparative* process tracing in a most similar case fall into this line of thinking (Bennett and Checkel 2014).

Making their own causal foundation explicit will force researchers to specify more clearly when they are convinced that a link in the chain of events they are studying is in fact *causal* and, thus, of explanatory relevance: if they subscribe to processes as truth conditions, this is the case if they have observed the transfer of energy/information; if they subscribe to counterfactuals, this will be the case if they can think of an adequate counterfactual or study a most similar case wherein the absence of the previous event leads to the absence of the subsequent event in the causal chain.

In social science and history, counterfactuals in the study of singular cases are not uncommon. David Lewis is rarely referenced in this body of literature and possible worlds are more loosely understood (for example, Mahoney and Barrenechea 2017), but the social scientists' and historians' toolbox for constructing counterfactuals shows striking similarities to the rules for ranking possible-worlds (in particular Lebow 2010:chap. 2). Rules such as "minimum rewriting" and "logical consistency" aim to change as little as possible and as much as necessary about the actual world to obtain the closest possible world.[25]

Structural causal modeling can clarify which counterfactuals are adequate in process tracing studies and thus serve a similar, yet not fully identical role as in comparative case studies. The structural model can codify the informal guidelines for constructing counterfactuals developed by qualitative methodologists and make the counterfactual analysis more transparent.

[25] Mahoney and Barrenechea introduce different types of token-level counterfactuals based on the set-theoretic concepts of necessity, sufficiency, INUS conditions and SUIN conditions (2017). The problem with using set theory for the study of *singular* cases is that the origin and meaning of the sets is unclear. Sets are groups of cases which is equivalent to a type-level perspective. As we explained before, this creates a mismatch between levels and is inadequate for making singular causal claims about counterfactual dependence.

For example, if we are interested in x taking on a different value than in the actual world, the causal model specifies the antecedents of x that would be kept fixed in the counterfactual analysis and thus allows us to assess the "realism" of the intervention (Lebow 2010:50-51). Another informal criterion is *interconnectedness*. It refers to the counterfactual processes that unfold from the counterfactual change of x (Lebow 2010:50). Again, these processes become more transparent in a structural model because it requires us to specify the path connecting the hypothetical intervention on x to the outcome y.

A structural approach also solves a hitherto overlooked problem of process tracing encountered by researchers committed to determinism (for example, Beach and Pedersen 2016:chap. 2). Like Lewis himself, they need to answer the question how the actual world could ever be different if it is governed by determinism. A structural account married to the idea of exogenous interventions offers a way out of this problem for determinists.

The argument that counterfactual truth conditions can guide causal inference in process tracing seems to be at odds with recent arguments in the process-tracing literature that deny the possibility of counterfactual reasoning, in their account in favor of a *mechanistic* view of causation (Beach and Pedersen 2016:40). We endorse the idea of *explaining* by specifying mechanisms and conceive of them as systems comprising entities and activities (or actors and their behavior) (Machamer et al. 2000:3). There are two reasons why this does not speak against linking process tracing to token-level counterfactuals in causal analysis. First, the philosophers who invented the idea of mechanisms as systems and who are often referenced in this context, most notably Glennan (2011:section 37.4), Craver (2007:chap. 3) and Craver and Darden (2012:chap. 8) themselves argue that a counterfactual theory is useful for determining whether a mechanism and its entities and activities are causal. This has so far been ignored in the social science debate about mechanisms. Second, it is wrong to argue that counterfactuals, also called

"logical parlor games" and "natural dirty experiments" (Beach and Pedersen 2016:40), are uninformative for drawing inferences about a single case.[26] As we explained above, a major strand of research on actual causation and events in single cases in philosophy builds on a counterfactual understanding of causation (Lewis 1973; Halpern 2016).[27] This is a meaningful endeavor because the counterfactual theory establishes a precise truth condition for *making causal claims about actual events in a given case*. Counterfactuals are not an end in themselves, but a *means* for establishing the truth of causal claims about actual singular phenomena (Kutach 2014:75).

5.3. Multi-method research

We understand multi-method research (MMR) as the combination of a statistical large-n analysis with process tracing (Lieberman 2005), or the integration of QCA with process tracing (Schneider and Rohlfing 2013).[28] MMR has been criticized as incoherent because the large-n and the small-n analysis are accused of applying incommensurable logics of inference (Chatterjee 2013). Our discussion of statistical observational research, QCA and process tracing shows that this critique is attributable to a misconception of the underlying theories of causation. Our discussion of philosophies of causation has shown that the truth-conditions for causal claims are *either* located on the type level and need to be supported by a type-level method of causal inference, *or* on the token level and demand to be supported by a method generating inferences on the token level. In this light, there cannot be incommensurable logics of inference because we follow one logic of inference on one level. Instead of trying to align causal inferences made on

[26] "Actualism" is pitted against counterfactuals, but this is a false dichotomy because possibilism is the opposite position to actualism.
[27] Baumgartner (2013) develops an actual regularity theory of causation, showing that there are a multitude of ways to draw causal inferences on actual cases.
[28] There are more variants of MMR than these two, but these are most often applied in the empirical literature.

two levels based on two theories, we can either describe type-level patterns to assess the *generalizability* of a token-level causal inference or describe the succession of events in a singular case to illustrate the *instantiation* of a type-level causal inference (see section 2). Thus, both methods always play a valuable role in multi-method research, but only one method contributes to causal inference about one and the same causal claim, since a causal claim is either formulated as a type, or a token-level claim.

For example, we could learn from the case of Catalonia and a corresponding counterfactual that granting more autonomy to the Catalans in 2006 did *cause* more people to vote secessionist in 2010. We could then determine whether we find the same pattern of autonomy and secessionism for minorities worldwide without using the insights for inferring that autonomy is *causal* for secessionism at the type level. In an alternative form of MMR, we could commit ourselves to Woodward's type-level theory and use process tracing to substantiate the premise that the type-level effects are underpinned by processes in singular cases.[29] In this case, we would choose a causal identification strategy for a dataset covering a large number of minorities worldwide to infer that autonomy *causes* secessionism. We could then pick a typical case to illustrate *how* autonomy leads to secessionism. The process analysis could be confined to describing a chain of events, without engaging in systematic counterfactual reasoning at each link in the chain because the causal relationship has already been inferred in the large N analysis.[30]

5.4. The broader picture

Our discussion of theory-method linkages highlighted an important difference between experiments and quantitative analysis of observational data on one hand, and qualitative medium

[29] Runhardt (2015) can be interpreted along these lines with regard to MMR designs.

[30] The type-level analysis would not shed light on the causal mechanism linking autonomy to secessionism. However, neither would process tracing, since it could only contribute token-level evidence, and, in this example, we are now testing whether a type-level relation obtains.

and small-n research on the other. Experiments and the quantitative analysis of observational data are currently firmly anchored in an interventionist theory of causation, defining causation in terms of structural relations between variables. The tight linkage exists because philosophers of science, and Woodward in particular, actively engaged with problems of causal inference in quantitative research and current practices in empirical social science studies when developing their theory of causation.

The picture looks different for qualitative methods, where exchanges between theorists of causation, methodologists and empirical researchers are not as close. In the field of QCA, the debate has recently shifted towards regularities thanks to Baumgartner's work and there is no serious, fully developed contender available at the moment. In the literature on comparative case studies and process tracing, explicit references to theories of causation are rare. Counterfactuals are mentioned for comparative case studies, but not systematically linked to the corresponding theory of causation, while the discussion on process tracing centers on the role of mechanisms in explanation rather than on theories of causation. Multi-method research supports this impression because it expects complementary methods to have synergistic causal value. We have argued that this is misguided because theories of causation locate the truth conditions of causal claims either on the type or the token level, which implies that for one and the same causal claim, only one of the two methods can provide the adequate evidence for valid causal inference.

Building on the philosophical and methodological literature, we have shown that apart from QCA, all methods can be aligned with at least two theories of causation. However, we have also argued that a counterfactual structural foundation is the superior solution in those cases where a method matches the truth conditions of more than one theory of causation. The plea for counterfactuals is scarcely new to quantitative researchers because it is the foundation of their methods. It is also not new to qualitative researchers because the counterfactual, potential

outcomes model has been proposed as ideal for qualitative methods in *Designing Social Inquiry* (DSI, King, Keohane and Verba 1994:chap. 3). In our reading of the post-DSI debate, this has led qualitative methodologists to see difference-making views of causation and counterfactuals as something inherently linked to quantitative research (for example Beach and Pedersen 2016:chap. 9). Our arguments have shown that this take on the debate ignores token-level counterfactual theories positing truth conditions that are adequate and fully compatible with small-N and single-case qualitative research.

6. Conclusion

In this article, we have clarified the relationship between social science methods of causal inference and philosophical theories of causation and aimed to provide methodologists and empirical researchers with an overview of the state of the art in the philosophy of causation. A survey of empirical researchers and methodologists showed that researchers hold a variety of views on concepts of causation and that many of these views are unclear or mix elements of different philosophical theories of causation. The survey also showed that many uncertainties surround the question of how theories of causation and methods of inference are related to each other, uncertainties that we aimed to resolve in this article.

Following the philosophical distinction between type- and token-level theories of causation, we showed that both large-N and small-N methods are capable of valid causal inferences. Type theories investigate whether relations between types of events or general properties are causal, whereas token theories tell us whether a relationship between singular events is causal. Adhering to the connection between type-level truth conditions and large N methods on one hand, and token-level truth conditions and small-N methods on the other still leaves room for different theory-method linkages: Experiments and inferential statistics of

observational data can assess causal claims interpreted in terms of regularities, probabilities or interventions. Process tracing can be linked to token-level counterfactuals or the process theory of causation. Comparative case studies are tied to the notion of causation as counterfactual dependence at the token level and, at the moment, the only option for QCA is the interpretation of causal relations in terms of type-level regularities.

We have argued that in those cases where several theories of causation are compatible with a method, structural counterfactual theories that come in both a type and a token variant are to be preferred. Among the type theories, counterfactual theories allow for causal inference and explanation without the need to make reference to universal laws to tell causal from accidental relations. Among the token theories, token-level structural counterfactuals are superior to a process view because they neither require us to establish whether a causal hypothesis holds at the level of (likely unobservable and untraceable) physical processes, nor do they ask us to make the unrealistic assumption that social systems are closed systems governed by the law of conservation. In addition, a counterfactual perspective accounts for causation by omission. A structural perspective is also superior to a possible-worlds perspective that requires miracles to explain how a possible world could come about in the first place.

It might have come as a surprise to some that there is not "one true theory of causation", but that we have discussed choices from a repertoire instead. From a philosophical perspective, all theories of causation face their challenges and have unresolved problems. However, it is important to know philosophers of causation are always aiming to construct scenarios in which a theory *ceases* to offer necessary or sufficient criteria for making correct causal claims. In other words, a theory of causation is supposed to work in every possible situation one can think of. This differs from empirical social science research where we can craft research designs and search for cases that diminish or even eliminate these problems. For example, the method of

difference is defined by the goal to avoid overdetermination, which is a classic problem in philosophy of causation, by comparing two cases that only differ on the outcome and on one cause. We might not always achieve the ideal and have to deal with multiple challenges, but we should maintain the goal of reaching it guided by a proper theory of causation.

References

Baetu, Tudor M. 2016. "From Interventions to Mechanistic Explanations". *Synthese*. 193(10):3311-3327.

Barros, D. Benjamin. 2013. "Negative Causation in Causal and Mechanistic Explanation". *Synthese*. 190(3):449-469.

Baumgartner, Michael. 2008. "Regularity Theories Reassessed". *Philosophia*. 36(3):327-354.

Baumgartner, Michael. 2009. "Inferring Causal Complexity". *Sociological Methods & Research*. 38(1):71-101.

Baumgartner, Michael. 2013. "A Regularity Theoretic Approach to Actual Causation". *Erkenntnis*. 78(1):85-109.

Beach, Derek and Rasmus Brun Pedersen. 2013. *Process-Tracing Methods*. Ann Arbor: University of Michigan Press.

Beach, Derek and Rasmus Brun Pedersen. 2016. *Causal Case Studies*. Michigan: University of Michigan Press.

Bennett, Andrew and Jeffrey Checkel. 2014. "Process Tracing: From Methodological Roots to Best Practices". Bennett, Andrew and Jeffrey Checkel (ed.): *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press: 1-37.

Blatter, Joachim and Markus Haverland. 2012. *Designing Case Studies: Explanatory Approaches in Small-N Research*. Basingstoke: Palgrave Macmillan.

Brady, Henry A. 2008. "Causation and Explanation in Social Science". Box-Steffensmeier, Janet M., Henry Brady and David Collier (ed.): *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press: 217-270.

Cartwright, Nancy. 1979. "Causal Laws and Effective Strategies". *Noûs*. 13(4):419-437.

Chatterjee, Abhishek. 2013. "Ontology, Epistemology, and Multimethod Research in Political Science". *Philosophy of the Social Sciences*. 43(1):73-99.

Collier, David, Henry E. Brady and Jason Seawright. 2004. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology". Brady, Henry E. and David Collier (ed.): *Rethinking Social Inquiry. Diverse Tools, Shared Standards*. Lanham: Rowman & Littlefield: 229-266.

Craver, Carl F. 2007. *Explaining the Brain*. Oxford: Clarendon Press.

Craver, Carl F. and Lindley Darden. 2012. *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago: University of Chicago Press.

Dowding, Keith. 2016. *The Philosophy and Methods of Political Science*. London: Palgrave.

Dowe, Phil. 2000. *Physical Causation*. Cambridge: Cambridge University Press.

George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Gerring, John. 2005. "Causation: A Unified Framework for the Social Sciences". *Journal of Theoretical Politics*. 17(2):163-198.

Glennan, Stuart. 2011. "Singular and General Causal Relations: A Mechanist Perspective". McKay Illari, Phyllis , Renee Russo and Jon Williamson (ed.): *Causality in the Sciences*. Oxford: Oxford University Press: 789-817.

Goertz, Gary and James Mahoney. 2012. *A Tale of Two Cultures: Contrasting Qualitative and Quantitative Paradigms*. Princeton: Princeton University Press.

Hale, Henry E. 2011. "Formal Constitutions in Informal Politics: Institutions and Democratization in Post-Soviet Eurasia". *World Politics*. 63(4):581-617.

Halpern, Joseph Y. 2016. *Actual Causality*. Cambridge, MA: MIT Press.

Hausman, Daniel M. 2002. "Physical Causation: Phil Dowe, Physical Causation". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. 33(4):717-724.

Hofmann, Urs and Michael Baumgartner. 2011. “Determinism and the method of difference”. 26(2): 155-176.

Hume, David. 2003. *A Treatise of Human Nature*. Mineola, N.Y.: Dover Publications.

Johnson, R. Burke, Federica Russo and Judith Schoonenboom. 2017. "Causation in Mixed Methods Research: The Meeting of Philosophy, Science, and Practice". *Journal of Mixed Methods Research*. advance access.

Kim, Jaegwon. 1971. “Causes and Events: Mackie on Causation”. *The Journal of Philosophy*. 68(14):426-441.

Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology". *Political Analysis*. 23(3):313-335.

King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Kutach, Douglas. 2014. *Causation*. Cambridge: Polity.

Lebow, Richard Ned. 2010. *Forbidden Fruit: Counterfactuals and International Relations*. Princeton: Princeton University Press.

Lewis, David. 1973a. "Causation". *The Journal of Philosophy*. 70(17):556-567.

Lewis, David. 1973b. *Counterfactuals*. Cambridge: Harvard University Press.

Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow". *Noûs*. 13:455-476.

Lieberman, Evan S. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research". *American Political Science Review*. 99(3):435-452.

Machamer, Peter, Lindley Darden and Carl F. Craver. 2000. "Thinking About Mechanisms". *Philosophy of Science*. 67(1):1-25.

Mackie, John L. 1965. “Causes and Conditions”. *American Philosophical Quarterly*. 2(4):245-264.

Mackie, John L. 1974. *The Cement of the Universe: A Study of Causation*. Oxford: Clarendon Press.

Mahoney, James. 2008. "Toward a Unified Theory of Causality". *Comparative Political Studies*. 41(4-5):412-436.

Mahoney, James and Rodrigo Barrenechea. 2017. "The Logic of Counterfactual Analysis in Case-Study Explanation". *The British Journal of Sociology*. advance access

Morgan, Stephen L. and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research (2nd Ed.)*. New York: Cambridge University Press.

Moses, Jonathon Wayne and Torbjørn L. Knutsen. 2007. *Ways of Knowing: Competing Methodologies in Social and Political Research*. Houndmills: Palgrave Macmillan.

Owen, John M. 1994. "How Liberalism Produces Democratic Peace". *International Security*. 19(2):87-125.

Paul, L. A. and Ned Hall. 2013. *Causation: A User's Guide*. Oxford: Oxford University Press.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference (2nd Ed.)*. Cambridge: Cambridge University Press.

Reutlinger, Alexander. 2013. *A Theory of Causation in the Social and Biological Sciences*. Basingstoke: Palgrave Macmillan.

Rohlfing, Ingo. 2012. *Case Studies and Causal Inference: An Integrative Framework*. Basingstoke: Palgrave Macmillan.

Rohlfing, Ingo and Carsten Q. Schneider. 2018. "A Unifying Framework for Causal Analysis in Set-Theoretic Multimethod Research". *Sociological Methods & Research*. 47(1):37-63.

Runhardt, Rosa W. 2015. "Evidence for Causal Mechanisms in Social Science:

Recommendations from Woodward's Manipulability Theory of Causation". *Philosophy of Science*. 82(5):1296-1307.

Salmon, Wesley C. 1998. *Causality and Explanation*. New York: Oxford University Press.

Schneider, Carsten Q. and Ingo Rohlfing. 2013. "Combining Qca and Process Tracing in Set-Theoretic Multi-Method Research". *Sociological Methods & Research*. 42(4):559-597.

Slater, Dan and Daniel Ziblatt. 2013. "The Enduring Indispensability of the Controlled Comparison". *Comparative Political Studies*. 46(10):1301-1327.

Spirtes, Peter, Clark Glymour and Richard Scheines. 2001. *Causation, Prediction, and Search (2nd Ed.)*. Cambridge: MIT Press.

Spohn, Wolfgang. 2006. "Causation. An Alternative". *British Journal for the Philosophy of Science*. 57(1):93-119.

VanderWeele, Tyler, J. and Robins James M. 2008. "Empirical and Counterfactual Conditions for Sufficient Cause Interactions". *Biometrika*. 95(1):49-49.

Waldner, David. 2014. "What Makes Process Tracing Good? Causal Mechanisms, Causal Inference, and the Completeness Standard in Comparative Politics". Bennett, Andrew and Jeffrey Checkel (ed.): *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press: 126-152.

Waldner, David. 2015. "Process Tracing and Qualitative Causal Inference". *Security Studies*. 24(2):239-250.

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, James. 2015. "Methodology, Ontology, and Interventionism". *Synthese*. 192(11):3577-3599.