Causal Mechanisms and Process Patterns in International Relations

Thinking Within and Without the Box

Jörg Friedrichs

University of Oxford

Department of International Development

3 Mansfield Road

Oxford OX1 3TB

England

Tel.: ++44/1865/281820

Email: joerg.friedrichs@qeh.ox.ac.uk

Abstract

This article makes the case for process patterns as an alternative to causal mechanisms. Causal mechanisms are explanatory tools to unpack the "black boxes" separating the input and output of models. Unlike causal mechanisms, process patterns do not require such a black box. They refer to recurrent sequences of interaction observed across any number of empirical domains. Scholars can apply them across disciplines when similar processes occur in different domains. The article provides examples from International Relations where scholars have sometimes studied process patterns in all but name.

Keywords

Causal mechanisms; process patterns; methodology; methods; research design

Acknowledgements

Thanks for suggestions and comments are due to Emanuel Adler, Giulia Andrighetto, Peter Hedström, David Houghton, Adam Humphreys, Friedrich Kratochwil, Milja Kurki, and Camelia Voinea.

Causal Mechanisms and Process Patterns in International Relations

Thinking Within and Without the Box

It is widely accepted in International Relations (IR) circles that there are two types of inference: descriptive and causal. Descriptive inference detects general tendencies within an empirical domain, whereas causal inference explains such tendencies. On the condescending note that descriptive inference is the best that can be obtained in inchoate fields of knowledge, King et al. (1994) see causal inference as the ultimate aim of any self-respecting scientist.

From a pragmatic point of view, this pejorative understanding of descriptive inference as the poor sibling of causal inference is unfortunate. Just as a good map may enable a commander to win a military campaign, and just like descriptive statistics is more meaningful to human intuition than inferential statistics, sound descriptive inference is a useful pathfinder through the complexities of the social world (Friedrichs and Kratochwil, 2009).

Besides, descriptive inference has a practical advantage: it is relatively easy to check it against evidence. This is so because descriptive inference always and by definition refers to observable entities. Causal inference is rather different: it often refers to unobservable entities, for example when it relies on logical deductions or the detection of regular associations (or constant conjunctions)

of independent and dependent variables. In such cases, the scholar infers causality without any direct observation of how it operates.

Fortunately, causal inference can be more empirical and descriptive. It is descriptive when we trace causal connections from the empirical record. For example, we observe that Event A is followed by a sequence of events that trigger Event B, thus causing it. This is an exercise in descriptive inference because the operation refers to observable phenomena. At the same time, it is also an exercise in causal inference because we are explaining something. The upshot is that descriptive versus causal inference is a false dichotomy.

Positivists know this in their bones, and their response is refreshingly pragmatic. In addition to relying on conjunctions of independent and dependent variables to formulate causal claims, they rely on causal mechanisms to infer causality descriptively. The most typical procedure is to first identify a regularity and formulate it terms of variables, and then to test it for causality by checking if there are specific linkages, or sequences of linkages, connecting independent and dependent variables (Bunge, 1997). These connectors, called causal mechanisms, inject a descriptive and thus empirical element into causal inference (Hedström and Swedberg, 1998).

Positivists are not alone in pursuing this strategy. Scientific realist have taken similar views on causal or “generative” mechanisms as the descriptive bedrock of causal inference (Bennett, 2013). In its most typical form, the strategy is to

first formulate a causal hypothesis and then test it empirically by identifying the mechanism(s) connecting the input and output factors of the model (Gerring, 2008). This means, as it were, that the researcher must first establish a regularity, with causation relegated to a "black box". She then makes that box translucent so that the causal mechanisms become visible (Figure 1).

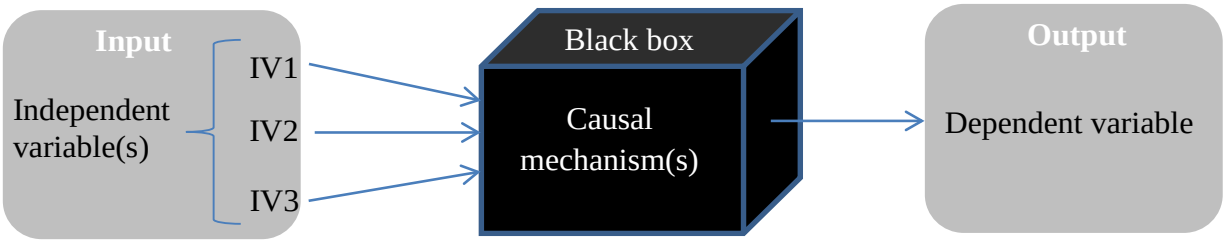

Figure 1: The positivist notion of causal mechanisms

The procedure has merits in some cases, but it is too narrow in many others. With due respect for mechanistic *thinking within the box*, it is sometimes more promising to infer causality by *thinking without the box*. As this article shows, we may systematically study recurring processes by means of causal-descriptive inference even in the absence of a black box. At its core, the strategy is to observe causal processes via the comparative analysis of cases across any number of empirical domains. To differentiate this approach from the black-box notion of *causal mechanisms*, let us call it *process patterns*.

We may define *causal mechanisms* as recurrent sequences of interaction generating a specific outcome, given specified initial conditions. They feature prominently as a methodological tool in many disciplines, including general

and comparative social and political science (e.g. Elster, 1989; Elster, 1999; Stinchcombe, 1991; Bunge, 1997; Hedström and Swedberg, 1998; Steel, 2004; Mayntz, 2004; Gerring, 2008; Gross, 2009; Falleti and Lynch, 2009; Hedström and Ylikoski, 2010). International Relations scholars in particular have offered methodological contributions (e.g. Wendt, 1987; Dessler, 1991; Tilly, 2001; George and Bennett, 2005; Gehring and Oberthür, 2009; Bennett, 2013) and concrete applications (e.g. Young, 1999; Houghton, 2009; Banta, 2013).

My main contention is that *process patterns*, or recurrent sequences of interaction observed across any number of domains, are an interesting methodological alternative that opens promising avenues. Insofar as similar processes occur in different social domains, scholars can profitably engage in transdisciplinary research and collaboration to examine process patterns.

The first section introduces causal pathways as a wider category within which both causal mechanism and other forms of recurrent sequences of interaction, such as process patterns, have their place. In the second section, I present the conventional specification of causal pathways as causal mechanisms. In the third section, I introduce process patterns as a methodological alternative to causal mechanisms. In the fourth and final section, I discuss specific examples related to IR to show how the proposed methodology works in practice.

Causal pathways: the overarching category

There is a problem with the common practice of taking causal mechanisms to encompass *any* recurrent sequence of interaction: it amounts to conceptual stretching. Conceptual stretching occurs when a category is unable to cover all that people subsume under it. When that happens, the remedy is to climb Sartori's (1970) ladder of generality and introduce an overarching category that is sufficiently broad and open to subsume everything intended. In this vein, let me propose causal pathways as an overarching category to causal mechanisms and process patterns alongside other usages (Figure 2).

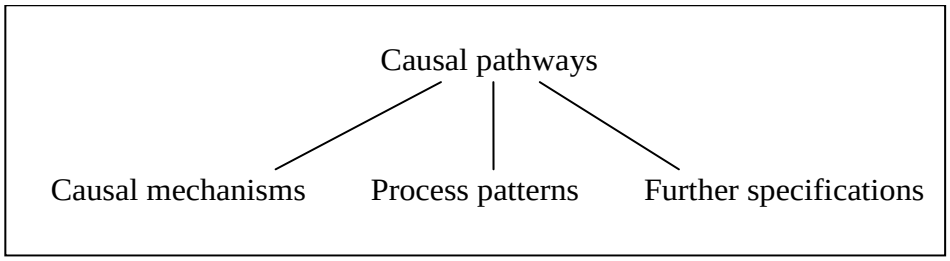

Figure 2: Causal pathways

Let me first demonstrate that conceptual stretching is indeed occurring. In the conventional view, a causal mechanism is "the pathway or process by which an effect is produced." This definition hails from a review article by Gerring (2008), entitled "The mechanismic worldview: thinking inside the box." The purpose of Gerring's definition is to represent the smallest common denominator. Gerring derived it from an impressive array of methodological

contributions, from which he distilled nine different meanings of "mechanism", which he boiled further down into his minimal definition.

The problem is that Gerring's minimal definition is not broad enough to cover the gamut of legitimate research practices that sail under the flag of causal mechanisms. This becomes apparent when we take Mahoney's (2001: 579-580) glossary of 24 definitions. Some of the definitions listed in Mahoney's glossary fall under Gerring's minimal definition, but others don't.

An example of the first group is the definition of a mechanism as "the process by which one variable influences the other, in other words, how it is that X produces Y" (Kiser and Hechter, 1991: 5). Another example is their definition as "analytical constructs that provide hypothetical links between observable events" (Hedström and Swedberg, 1998: 13). Unlike these two, other definitions are incompatible with Gerring's minimal definition. For example, Elster (1999: 1) defines mechanisms as "frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences." It is impossible to subsume this definition under Gerring's minimal definition because, where Gerring talks about a specific effect, Elster talks about indeterminate consequences.

The result is stunted perception. According to Gerring (2010: 1500), causal mechanisms "are not at variance with traditional practices in the social sciences and thus hardly qualify as a distinct approach to causal assessment."

Based on his positivist understanding of mechanisms as "thinking inside the box", Gerring cannot find any "sharp point of contrast between 'mechanismic' and 'covariational' social science" (Gerring, 2010: 1505).

Since the term "causal mechanisms" is too narrow as a common denominator, let me suggest the umbrella term of "causal pathways," or *recurrent sequences of interaction that a researcher may invoke in the descriptive inference of causal explanations*. This broad umbrella term remains open to diverse specifications, as exemplified by the contrast between causal mechanisms (*thinking within the box*) and process patterns (*thinking without the box*).

Causal mechanisms: thinking within the box

Causal mechanisms are *recurrent sequences of interaction generating a specific outcome, given specified initial conditions.* In his 1989 book, Elster saw them as the "nuts and bolts", or "cogs and wheels" of social science, connecting observable inputs with observable outputs in what would otherwise be the erratic black boxes of empirical generalization (Elster, 1989: 3).

The fundamental idea behind this so-called I-M-O, or Input-Mechanism-Output model (Hedström and Swedberg, 1998: 9) is that one should not prematurely assume causality when specified initial conditions (inputs, e.g. independent variables) are empirically associated with specified outcomes

(outputs, e.g. dependent variables). One should assume causality only when a causal mechanism can be shown to intervene in what would otherwise be the black box separating a model's inputs from its outputs.

The model works for both the hypothetic-deductive model of theory testing and the variable-oriented model of statistical inference. In the former, mechanisms link inputs and outputs by demonstrable causal accounts and thus establish causality for covering laws (Hempel, 1965). In the latter, they similarly connect the independent and dependent variables of statistical models (Lieberman, 2005; Imai et al., 2011). Like positivists of different stripes and colors, scientific realists are also interested in causal mechanisms as the descriptive bedrock of causal inference (Bennett, 2013).

Thus understood, causal mechanisms are an indispensable supplement to conventional modes of empirical generalization. The point is that, after identifying a covering law or a co-variation between variables, one must open the black box in order to see if input and output are actually connected in a causal way. Without an identifiable mechanism, covering laws and statistical correlations are likely to be spurious and cannot be taken as causal.

> Discovering correlations between assumed causes and effects is not good enough. Explaining means showing *how* a cause leads to an effect. (…) Democracies do not seem to fight each other. But saying that democratic domestic structures explain peace among democracies is unsatisfactory.

> It needs to be shown, for example, how parliaments and civil society groups scrutinize executives and thereby reduce the likelihood of war. Saying that smoking kills is a superficial statement. To explain why many smokers die prematurely, the effects of nicotine on lungs and other body parts need to be revealed (Daase and Friesendorf, 2010: 12).

Searching for causal mechanisms means to check if there are linkages connecting the input and output of a given model and thus to evaluate causality empirically. From this perspective, causal reconstruction starts with an empirical generalization about a specific population or sample of cases, and subsequently identifies the mechanism(s) that deterministically or probabilistically produce the outcome (Mayntz, 2004). Scholars following this template are not entitled to make causal statements unless and until they find appropriate mechanisms operating within their black boxes.[1]

To do this kind of analysis involves a series of analytical steps. First, you need a population or sample composed of equivalent entities (democracies; tobacco smokers). Next, you need to derive a tentative empirical generalization (democracies don't fight; smokers get cancer). This may take the form of a covering law, correlation, probabilistic statement, or combination of necessary

[1] Critics object that causal mechanisms are nothing more than a shorthand for temporal sequences of intervening variables. This is not convincing, as causal mechanisms are often non-linear: interaction effects, feedback loops, etc.

and sufficient causation (Rihoux and Ragin, 2009; Goertz and Mahoney, 2012). Finally, you open the black box to see if there is a causal mechanism linking input and output, independent and dependent variables, stochastic events, or conditions and outcome. Process tracing is the method of choice for searching such linkages (Beach and Pedersen, 2013).

There is nothing wrong with this. On the contrary, it is a valuable antidote against the spurious attribution of causality. However, it is not particularly friendly to transdisciplinary research and collaboration. It always presupposes an empirical generalization (e.g. democratic peace) over a specified domain of entities (e.g. states). Since the domain is likely to rest squarely within an established discipline (e.g. IR) or sub-discipline, cross-fertilization among areas of research that are substantively remote from each other is unlikely.

Process patterns: thinking without the box

Let us revert to the overarching category of causal pathways as recurrent sequences of interaction. As I will show, causal pathways need not be mechanistic, nor do they have to unpack black boxes. Instead, it is possible to think entirely without the box—what I call process patterns, or *recurrent sequences of interaction observed across any number of domains*.

To clarify, consider two sets of examples.

1. What do the following situations have in common: a panic in a theatre; a bank run; and the famous race to the bottom after the world economic crisis of 1929?
2. Again, what do the following situations have in common: mobbing an ambitious colleague; bullying a nerd; and a smear campaign against an incorruptible politician?

As the examples indicate, we are capable of noticing patterned regularities across wildly different domains. We do not need an empirical generalization over a specified domain before we can open, as it were, the black box and identify a mechanism to confirm that causality is at work. Instead, we can altogether suspend thinking in terms of black boxes and take process patterns as the heuristic starting point of our analysis.

This entails a methodological shift from population-based to pattern-based case selection; and from variable-driven to process-driven research (cf. Rescher, 1996; Jackson and Nexon, 1999). We do not need to first draw a sample from a given population such as “nation-states” or “democracies”, then stipulate a domain-specific empirical generalization such as “democratic peace”, and finally test it for causal mechanisms. Instead, we can think of all process patterns of a particular type as constituting a universe of cases to be examined through a systematic and rigorous procedure of comparative analysis. This turns the usual procedure on its head.

In population-based case selection and variable-driven research, the researcher starts with a population of ontological givens, draws a sample, derives empirical generalizations, and finally seeks to identify causal mechanisms. In pattern-based case selection and process-driven research, the researcher starts with an ontologically thinner categorization, namely an observed process pattern that constitutes a universe of potential cases across any number of domains. Sampling then follows pragmatic criteria, with the cases selected revealing as much interesting variation as possible. The objective is to arrive at a precise and fully explicit articulation of the observed regularity.

Harking back to the first example above (bank runs etc.), let us take the process pattern whereby self-regarding behavior is collectively self-defeating in that its private pursuit reinforces the propensity of others to engage in the same kind of self-regarding behavior, thus leading to the worst possible outcome for everybody (akin to a prisoner's dilemma). It is easy to see that, apart from the process pattern they share in common, the universe of cases based on this pattern is not composed of ontologically homogenous entities. You would never put theatre panics, bank runs, and the world economic crisis of 1929 into the same basket unless you had a prior understanding of the process pattern at work in all of these cases. Yet, once you have identified the process and constituted a pattern-based case universe, you can sample from it and set in motion a systematic procedure of comparative research.

You are then able to ask interesting research questions. What do all manifestations of the process pattern have in common, what are the specificities of particular sub-sets, and what are the idiosyncrasies of individual cases? For instance, what are the commonalities and differences between theatre panics and bank runs? What role do context, actor constellations, and trigger events play (Falleti and Lynch, 2009; Scharpf, 1997; Pierson, 2004)? To what extent can institutions like fire exits and financial reinsurance transform the dynamics of theatre panics and bank runs? Are there any discontinuities or path dependencies? What determines when a discontinuity kicks in, and whether the pattern goes down one path rather than another? How does all of this affect the outcome? For example, why are some theatre panics and bank runs more devastating than others?

At its core, the analysis of process patterns is an exercise in descriptive inference. Descriptive inference, including the analysis of recurrent patterns of interaction, is valuable in and of itself as it increases our ability to understand and contextualize social phenomena. When a social process falls under a process pattern, this puts our mind at rest and thus may serve as an explanation in a broader sense than is conventionally associated with causal inference.

Indeed, process patterns can be of many different kinds. Sometimes they can be developed through classical categorization (Sartori, 1970). Sometimes it is better to formulate them in ideal-typical terms (Weber, 2004 [1904]). Sometimes they are mechanical, and sometimes they require “understanding”

the internal viewpoint of the participants (Weber, 2004 [1921]; Hollis and Smith, 1990). Sometimes they constitutive explanations providing an answer to a "how possible" question (Wendt, 1998). At other times, they will follow other explanatory protocols. In short, process patterns are causal in a broad, catholic sense that encompasses not only efficient causes but also many other forms of explanation: empathetic, constitutive, etc. (Kurki, 2008).

Insofar as process patterns are about causality in motion, it is possible to repackage them as causal mechanisms whenever they follow specific trigger events, occur in specific contexts, and lead to specific outcomes. Thus, it is possible to move seamlessly from one form of causal inference to the other. It is important to note, however, that process patterns are more than a mere prelude to the study of causal mechanisms. The more interesting and productive part of a research cycle will often happen at the stage of exploring process patterns. Moreover, process patterns may just as well be a prelude to other pursuits such as discourse analysis and ethnographic fieldwork.

A research strategy that studies process patterns comparatively in different settings is well suited to support transdisciplinary research. It enables the researcher to detect regularities across divergent realms and thus to generate knowledge that can "travel" (Sartori, 1970). How far a specific process pattern can and should travel will depend on empirical and substantive considerations. At least in part, however, it is also up to evaluative choices of the researcher to decide at what level(s) of abstraction a process pattern is most useful. A

process pattern formulated in a more abstract way will travel further than one specified in more specific terms. Depending on how far a process pattern is required to travel, the researcher will formulate and reformulate its definition and thus the criteria for inclusion in the universe of possible cases.

Examples from research practice

In methodology, causal pathways are often reduced to causal mechanisms. In practice, however, there have been many interesting contributions falling under the rubric of process patterns for all but name. As the following examples suggest, the study of process patterns is indeed a rewarding task.

To begin with, game theorists have been studying process patterns for a long time. Take for example the emergence of “tit for tat” behavior as a recurrent phenomenon in situations of iterated prisoner’s dilemmas. Axelrod mentions phenomena as divergent as stickleback fish and divorcing women as cases where one can observe the emergence of “tit for tat” behavior (Axelrod, 1997: 5-6; Axelrod, 1984). Although he does not use the term “process pattern”, Axelrod is talking about a processual pattern with an extraordinary ability to travel across the most divergent areas of the social world, from the animal kingdom to IR (for even bolder attempts, see Scheffer, 2009; Steel, 2007).

Likewise, complex systems theory and agent-based modelling examine recurrent features of social life such as adaptive behavior and cybernetic feedback (Axelrod, 1997; Cederman, 1997). From the vantage point of complex system theory, the social world is of mindboggling complexity but there are emergent structures, or islands of order, even in the most chaotic systems. Agent-based modelling provides a useful tool to get a handle over the emergence of order in such situations (Hedström and Ylikoski, 2010: 62-64). From the viewpoint of complex systems theory and agent-based modelling, the best way for the researcher to proceed is arguably to take process patterns as starting points and explore as much as possible around them.

Another hotbed of pattern-based research, in the tradition of Robert Merton's middle-range theories, is the academic movement of analytical sociology (Hedström and Bearman, 2009; Demeulenaere, 2011). To cite only the most classical example, self-fulfilling and self-negating prophesies (Merton, 1948; Biggs, 2009) are observed over any number of social domains, and their relevance for IR has been demonstrated (Houghton, 2009).

Denial is an interesting case in point. Having presented a comprehensive treatment of denial elsewhere (Friedrichs, 2014), I can be brief here. Denial is the habit of treating a real problem as if it were a nonissue. A real problem is one that makes us suffer regardless of whether or not we acknowledge it.

Thus understood, denial is a ubiquitous social and psychological phenomenon. Any kind of problem, from personal trauma to planetary challenges such as climate change, can be an object of denial. Denial occurs at the individual level, e.g. when a cancer patient disavows her terminal illness. It also occurs in social constellations, from the denial of marital infidelity to the denial of race discrimination. Some cases of denial are deeply political, e.g. when a nation state such as Turkey denies genocide. A particularly worrisome manifestation happens at the global level, where humankind is in denial of the fact that infinite growth on a finite planet is impossible. This includes widespread denial of climate change and resource depletion (Friedrichs, 2013).

As the examples suggest, and as is typical for process patterns, denial occurs across different social domains and levels of analysis. Although, or precisely because, denial can occur in multiple spheres of experience, it is amenable to systematic empirical scrutiny when understood as a process pattern. What is more, denial comes in different varieties depending on the problem at stake. A problem may be *tractable or intractable*, and it may be *permanent or escalating*. Without trying to be exhaustive, let us consider two such varieties to reveal the heuristic fruitfulness of the process-pattern perspective.

The most straightforward case of denial is the disavowal of a *permanent and intractable* problem. In such cases, folk wisdom has it that “If we can do something to solve the problem, we will perhaps do it. If there is nothing we

can do, we should try to forget the problem." The reason is that, while denial does not solve the problem itself, it reduces the discomfort of admitting it.

An example from IR is the way post-war Japan has handled the international stigma from the atrocities committed during the Asia-Pacific War. With due care, this can be compared to the disavowal of permanent problems in other domains, e.g. the way the Nixon Administration initially handled the Watergate Affair, or the way liberal white Australia has ignored until recently the suffering inflicted on the Aborigines. In all of these cases, denial may be morally objectionable but it is easy to see how the deniers half-consciously or (more likely) unconsciously serve their perceived self-interest by denying a reputational problem that they understand as intractable.

A more pathological case is the denial of a problem that escalates yet, until a certain point, *would* be tractable if appropriate action *were* taken. For example, during the years leading up to the 2008 financial meltdown, national and international public servants ignored whistleblowers criticizing the lack of market regulation until it was too late. Or, to cite another example, until the Euro crisis of 2010/11 there was a regime of silence in public circles around the fairly obvious way Greece had concealed its lack of fiscal austerity.

The same pattern occurs at the domestic level, with some of Europe's ageing societies blissfully ignoring the unsustainability of their public pension system, with predictably devastating consequences. Another case in point is

the way Hitler and his stalwarts, against all odds, refused to admit the deterioration of their military position. The denial of escalating yet tractable problems leads to tragedy because, with hindsight, it turns out that some action could and should have been taken at an earlier point in time.

Conclusion

While it is appropriate in many situations to rely on causal mechanisms as tools for *thinking within the box*, it is often more fruitful to study process patterns as devices for *thinking without the box*. As we have seen, this has important methodological implications. The passage from causal mechanisms to process patterns implies a shift from population-based to pattern-based case selection; and from variable-driven to process-driven research.

There are two main reasons why process patterns hold promise. First, they offer an opportunity for transdisciplinary research and collaboration. The world is punctuated by recurrent sequences of interaction, and process patterns can span any number of social domains. Because this is so, research investigating process patterns is likely to cross disciplinary boundaries.

Second, process patterns commit us to open-ended research. We are more likely to encounter surprises and genuinely learn something new if we map out recurrent social processes first; and only then ponder over which social

domains they apply, and whether and to what extent we can or should repackage them in terms of conventional causal models and frameworks.

Besides, they mitigate the “small n” problem. IR scholars can often mobilize only a small number of cases that is insufficient for robust generalization. For example, nuclear deterrence during the Cold War initially appeared unique and unprecedented. When scholars started to analyze it as a process pattern (mutual assured destruction, MAD), the strategic logic became apparent and it was possible to gain insights from analogues in other fields.

References

Axelrod RM. (1984) *The Evolution of Cooperation,* New York: Basic Books.

Axelrod RM. (1997) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration,* Princeton: Princeton University Press.

Banta B. (2013) Analysing discourse as a causal mechanism. *European Journal of International Relations* 19: 379-402.

Beach D and Pedersen RB. (2013) *Process-Tracing Methods: Foundations and Guidelines,* Ann Arbor: University of Michigan Press.

Bennett A. (2013) The mother of all isms: causal mechanisms and structured pluralism in International Relations theory. *European Journal of International Relations* 19: 459-481.

Biggs M. (2009) Self-fulfilling prophecies. In: Hedström P and Bearman P (eds) *The Oxford Handbook of Analytical Sociology.* Oxford: Oxford University Press, 294-314.

Bunge M. (1997) Mechanism and explanation. *Philosophy of the Social Sciences* 27: 410-465.

Cederman LE. (1997) *Emergent Actors in World Politics: How States and Nations Develop and Dissolve,* Princeton: Princeton University Press.

Daase C and Friesendorf C. (2010) Security governance and unintended consequences. In: Daase C and Friesendorf C (eds) *Rethinking Security Governance: The Problem of Unintended Consequences.* London and New York: Routledge, 1-20.

Demeulenaere P. (2011) Analytical sociology and social mechanisms. Cambridge: Cambridge University Press.

Dessler D. (1991) Beyond correlations: toward a causal theory of war. *International Studies Quarterly* 35: 337-355.

Elster J. (1989) *Nuts and Bolts for the Social Sciences,* Cambridge: Cambridge University Press.

Elster J. (1999) *Alchemies of the Mind: Rationality and the Emotions,* Cambridge: Cambridge University Press.

Falleti TG and Lynch JF. (2009) Context and causal mechanisms in political analysis. *Comparative Political Studies* 42: 1143-1166.

Friedrichs J. (2013) *The Future Is Not What It Used to Be: Climate Change and Energy Scarcity,* Cambridge, MA: MIT Press.

Friedrichs J. (2014) Useful lies: the twisted rationality of denial. *Philosophical Psychology* 27: 212-234.

Friedrichs J and Kratochwil F. (2009) On acting and knowing: how pragmatism can advance international relations research and methodology. *International Organization* 63: 701-731.

Gehring T and Oberthür S. (2009) The causal mechanisms of interactions between international institutions. *European Journal of International Relations* 15: 125-156.

George AL and Bennett A. (2005) *Case Studies and Theory Development in the Social Sciences,* Cambridge, MA: MIT Press.

Gerring J. (2008) The mechanismic worldview: thinking inside the box. *British Journal of Political Science* 38: 161-179.

Gerring J. (2010) Causal mechanisms: yes, but…. *Comparative Political Studies* 43: 1499-1526.
Goertz G and Mahoney J. (2012) *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences,* Princeton: Princeton University Press.
Gross N. (2009) A pragmatist theory of social mechanisms. *American Sociological Review* 74: 358-379.
Hedström P and Bearman P. (2009) The Oxford Handbook of Analytical Sociology. Oxford: Oxford University Press.
Hedström P and Swedberg R. (1998) Social Mechanisms: An Analytical Approach to Social Theory. Cambridge: Cambridge University Press.
Hedström P and Ylikoski P. (2010) Causal mechanisms in the social sciences. *Annual Review of Sociology* 36: 49-67.
Hempel CG. (1965) *Aspects of Scientific Explanation and other Essays in the Philosophy of Science,* New York: Free Press.
Hollis M and Smith S. (1990) *Explaining and Understanding International Relations,* Oxford: Clarendon Press.
Houghton DP. (2009) The role of self-fulfilling and self-negating prophecies in international relations. *International Studies Review* 11: 552-584.
Imai K, Keele L, Tingley D, et al. (2011) Unpacking the black box of causality: learning about causal mechanisms from experimental and observational studies. *American Political Science Review* 105: 765-789.
Jackson PT and Nexon DH. (1999) Relations before states: substance, process and the study of world politics. *European Journal of International Relations* 5: 291-332.
King G, Keohane RO and Verba S. (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research,* Princeton: Princeton University Press.
Kiser E and Hechter M. (1991) The role of general theory in comparative-historical sociology. *American Journal of Sociology* 97: 1-30.
Kurki M. (2008) *Causation in International Relations: Reclaiming Causal Analysis,* Cambridge: Cambridge University Press.
Lieberman ES. (2005) Nested analysis as a mixed-method strategy for comparative research. *American Political Science Review* 99: 435-452.
Mahoney J. (2001) Beyond correlational analysis: recent innovations in theory and method. *Sociological Forum* 16: 575-593.
Mayntz R. (2004) Mechanisms in the analysis of social macro-phenomena. *Philosophy of the Social Sciences* 34: 237-259.
Merton RK. (1948) The self-fulfilling prophecy. *Antioch Review* 8: 193-210.
Pierson P. (2004) *Politics in Time: History, Institutions, and Social Analysis,* Princeton: Princeton University Press.
Rescher N. (1996) *Process Metaphysics: An Introduction to Process Philosophy,* New York: State University of New York Press.
Rihoux B and Ragin C. (2009) Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Los Angeles: SAGE.

Sartori G. (1970) Concept misformation in comparative politics. *American Political Science Review* 64: 1033-1053.
Scharpf FW. (1997) *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research,* Boulder: Westview.
Scheffer M. (2009) *Critical Transitions in Nature and Society,* Princeton: Princeton University Press.
Steel D. (2004) Social mechanisms and causal inference. *Philosophy of the Social Sciences* 34: 55-78.
Steel D. (2007) *Across the Boundaries: Extrapolation in Biology and Social Science,* Oxford: Oxford University Press.
Stinchcombe AL. (1991) The conditions of fruitfulness of theorizing about mechanisms in social science. *Philosophy of the Social Sciences* 21: 367-388.
Tilly C. (2001) Mechanisms in political processes. *Annual Review of Political Science* 4 21-41.
Weber M. (2004 [1904]) The 'Objectivity' of Knowledge in Social Science and Social Policy. In: Whimster S (ed) *The Essential Weber: A Reader.* London: Routledge, 387-404.
Weber M. (2004 [1921]) Basic sociological concepts. In: Whimster S (ed). London: Routledge, 311-333.
Wendt A. (1998) On constitution and causation in international relations. *Review of International Studies* 24: 101-118.
Wendt AE. (1987) The agent-structure problem in international relations theory. *International Organization* 41: 335-370.
Young OR. (1999) The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms. Cambridge, MA: MIT Press.