11 Discussion chapter

Shadow boxing in Plato's cave: assessing causal claims

Bertjan Verbeek

Introduction

The purpose of this contribution is to assess how several of the methodological approaches to investigating the role of non-state actors in international security, introduced and applied in this volume, contribute to the possibility of stating plausible causal claims. To that effect (no pun intended) this chapter dissects three contributions that specifically address the issue of causal claims: comparative case study methodology (Chapter 8 by Andreas Kruck on the privatization of security); qualitative case analysis (Chapter 9 by Patrick Mello on armed conflict); and geospatial analysis (Chapter 10 by Alexander De Juan on the emergence of violent conflict). I will organize my own claims around several questions to be put to each individual chapter: how to select a suitable method for answering a particular research question? How to select cases that fit research design and method(s)? What type of knowledge is served, or hindered, by our choice of method(s)? What can be said to have been explained? I will discuss the three chapters in this part of the volume while keeping these questions in mind. As will be shown, there is not always *the* perfect solution: researchers often have to make decisions that involve trade-offs.

One preliminary observation is in order, however: by assessing whether and how specific methodologies allow for causal claims it is presupposed that, at least in principle, causality exists, and that, to different degrees and possibly only indirectly, evidence exists that allows us to agree on the plausibility of causality to have occurred (see Guzzini 2016). Within this spectrum many different positions are still tenable and indeed the International Relations (IR) discipline boasts hugely diverse notions of causality. On the one hand, we find those who, sometimes explicitly modelling themselves after the natural sciences, seek for universal laws of international politics – Waltz's claims about balancing in anarchical systems come closest to that notion (Waltz 1979). On the other hand, we find those who reject the possibility of universal laws, but still agree that even unique historical developments can be unpacked as a sequence of events representing a causal chain in which one phenomenon at t_1 helped bring about another at t_2.

Social Constructivists often find themselves at odds with either position: because universal laws presuppose similar, if not identical, actors and contexts

to be present in the international realm, constructivists argue that actors, their properties, and the contexts in which they find themselves change over time because their social existence (constitution) is subject to social construction. Although this leaves little room for universal laws, for some constructivists this does not render impossible the notion of causality as long as the scholar strives for causal claims about time- and space-bound, comparable phenomena. Hence their emphasis on the importance of so-called scope conditions (e.g. Checkel 2001).

Still, a constructivist would argue, another issue has to be tackled before one can engage in such 'small t theorizing': the issue of mutual constitution. Because, in *social* constructivism, meaning emerges out of interaction between two or more actors, such meaning is mutually constituted in a process in which it is difficult to assume a clear linear process of x leading to y. Rather, for a constructivist, y affects x while, at the same moment, x affects y. This hugely complicates the issue of causal claims, because dissecting the mutual constitutive effect still requires the adoption, if only for pragmatic reasons, of some kind of time sequence. Hence, the position of some constructivists that constructivist research requires engaging in so-called bracketing. Bracketing involves the temporary assumption that one side of the mutual interaction can be considered to be constant in order to describe the other side. In this sense, constructivists borrow from an experimental research design (Price and Reus-Smit 1998; for a different perspective see Fierke 2005: 9–16; Klotz and Lynch 2007: 14–15).

Critical theorists would not necessarily discard the possibility of causal claims. They would point to two major issues: first, because to them empirical phenomena are by definition context bound (in time and place) they would hesitate talking in terms of universal laws and strive for limited causal claims for a circumscribed class of phenomena. Second, they would emphasize that the individual researcher herself is part of the social phenomenon she is studying. They thus add a new dimension to causality: everything the researcher does or does not do may have a causal effect on the phenomenon to be studied. This is diametrically opposed to those scholars in our field (such as, for example, neorealists) who would argue that the investigator is looking in on international reality from the outside and that her presence does not affect the subject under study. A critical theorist would therefore require the scholar to be aware of her position and impact on the matter under investigation.

With these caveats in mind, I will look upon the three chapters on causal claims. My own position is that causality implies a time sequence (even if in the form of constructivist 'bracketing') and that because of this we are implicitly or explicitly thinking in terms of an independent variable preceding a dependent variable. In addition, I start from the notion that establishing a causal claim not only requires an observation of a sequence of events, but above all an explanatory mechanism accounting for *why* the sequence has occurred. Such explanatory mechanisms can only be found in theories, also if we are investigating phenomena that only occur in a specific context of time and space.

How to select a suitable method for answering a particular research question?

One often hears the suggestion that one's research question determines one's research method. Yet, that claim is still no shortcut to a pragmatic decision on method. First of all, we need to distinguish between research design and methods in the more narrow sense. Second, we need to distinguish methods to collect data from methods to analyse data. One's research design is related to one's conceptual model which in turn reflects the research question as well the essential elements of the theory one considers plausible for tackling the research question. Theory could be conceived in a narrow sense (assumptions leading to the deduction of testable hypotheses) or in a wider sense (a set of related concepts that delineate the empirical phenomena we wish to study), but it always implies some kind of reduction of complexity as well as applying an interpretive filter that allows us to make sense of the world. E.g., realists do this when they present world politics as an anarchic system of states (despite the presence of other actors and forms of governance) on the grounds that only states have the capacity to organize in such a way that they can survive in such a surrounding.

Andreas Kruck's research (see Chapter 8) seeks to explain the variation in degree and nature of privatization of security in four different countries. Andreas Kruck opts for a comparative case study of France, Germany, the United Kingdom and the United States. Now the research question itself does not preclude other methods: it allows for a large-*n* study of privatization. His operational definition of privatization would certainly facilitate that. Yet, Andreas Kruck seeks to accomplish more than correlation between variables: he wants to see *whether* the causal mechanisms underlying such a correlation can be unveiled. In order to trace causality, one could, instead, consider opting for a large-*n* study supplemented with one or more case studies, a strategy sometimes preferred if a large *n* is available (Landman 2005). Andreas Kruck, however, opted for four case studies in order to establish both the 'comprehensive explanation' and the causal claims. He might have done so for pragmatic reasons because large-scale standardized data for numerous countries that would be necessary for a large-*n* study (supplemented by case studies) were not available as ready-made data sets and might be hard to gather by the individual researcher. He calls this 'zooming in' and 'linking up'. This is a plausible strategy for situations in which a large *n* is absent and a comparative case study seems warranted. At any rate, the broader lesson is that the number of (potentially) available *n* is an important determinant for the choice of an adequate method. The number of available cases may in turn depend on a whole range of factors including the real world frequency of a phenomenon, availability of (reliable) data and the researcher's time and resources for gathering new data. So researchers might want to ask themselves: would my research question allow for large *n* analyses, or does it suggest a comparative case study? Which additional considerations regarding data collection and data analysis would this demand from my research?

Patrick Mello (see Chapter 9) is looking for an explanation of the choices by 30 democracies of whether or not to participate in the 2003 Iraq War. Patrick Mello suggests Qualitative Comparative Analysis (QCA) as an appropriate method to tackle this specific question. Although his research involves few – if any – non-state actors, his empirics indicate that QCA can be a viable avenue for specific forms of research into non-state actors. QCA is a method of analysis that enables us to establish whether individual variables or specific combinations of variables constitute necessary or sufficient conditions for a specific outcome to occur by formally attaching weights to qualitative variables and counting the actual occurrence of all possible logical combinations of independent variables. Now QCA offers particular advantages to comparative case studies: when *n* is too small to engage in probability tests, yet the number of cases is substantive, QCA offers a method to make a statement about causal links between independent and dependent variables (or conditions and outcomes, as researchers using QCA prefer to call them) – importantly, it is a method that facilitates replication. The additional advantage is that QCA allows for making plausible statements about there being various causal pathways to the same outcome (equifinality). Furthermore, QCA permits identifying clear outlier cases, which could then be investigated in depth through single case studies. Hence, it might be useful for a lot of studies of non-state actors in international security, if researchers asked themselves: does my research allow for the possibility of different causal pathways leading to the observable outcome? Would in my research design QCA be useful to depict various such pathways?

Alexander De Juan (see Chapter 10) seeks to investigate the relationship between reduced availability of natural resources and the advent and intensity of violent conflict. His analysis involves data pertaining to the situation in the Sudanese region of Darfur. Interestingly, Alexander De Juan's research question springs from an empirical observation that geographical patterns of violent conflict do not coincide with geographical patterns of deteriorating supply of vital natural resources. This contradicts the common-sense idea that people would be prepared to use violence over extreme lack of access to (dwindling) resources. The observation induces him to formulate a theory that basically states that dwindling resources force people to migrate to areas of plenty and that this migration causes tension between old and new users of these resources. This observation allows for a conceptual model accounting for these two movements (first, migration patterns; second, use of violence) and for readily testable hypotheses, which could be approached both with qualitative and with quantitative methods. Alexander De Juan's approach thus highlights the importance of formulating testable hypotheses. This points to another set of questions relevant for causal research: does my research design allow me to formulate straightforward expectations? Which type of methods would allow me to put these to the test?

Data collection and units of analysis

In terms of data collection, it requires a decision on the unit of analysis: the case to be selected is supposedly part of a class of phenomena about which one's theory is expected to provide a meaningful interpretation. For Andreas Kruck (see Chapter 8) this seems to be all 'Western "strong" and democratic' countries that have to some extent privatized security. From that group he selects four. Although quite a considerable number for a comparative case study, one could debate whether more may have been better. However, researchers often have to limit their analysis for pragmatic reasons (how much time do I have for my research? What data are available?). Andreas Kruck offers three explanatory theories (functional, political and ideational). In particular, from the perspective of 'covariation analysis' – one of the three case study techniques Andreas Kruck discusses – one might argue that establishing a hint of causation would require holding constant two of three 'theories' each time one 'theory' is 'put to the test'. Four case studies would thus not yet yield sufficient insight into the relative strength of the three theoretical approaches. The four case studies do however yield a sizeable number of empirical observations (beyond IVs and DV) which – through 'congruence analysis', the second case study technique he mentions – can be compared with the predictions of the three theories. Moreover, Andreas Kruck's case selection would allow for tracing the causal mechanism within each approach, but then one case per approach would suffice. The question then is which case comes closest to representing each theoretical approach: the expectations formulated by Andreas Kruck make France the best candidate for the political approach (weakest parliament); the UK for the ideational approach (as the most neoliberal system) and possibly the USA for the functional angle (being the system most geared toward cost saving). Indeed, one would thus engage, as did Andreas Kruck, with structured-focused comparison as a research method (George 1979a): 'focused' implies a plausible link with the theory that informs the research (here: functional, political and ideational accounts of privatization). 'Structured' implies that the selected cases still have to be argued in terms of comparability: certainly, in Andreas Kruck's chapter, France, the UK and USA have, or seek, a leading role in world politics and thus would have a comparable 'interest' in an effective security apparatus. This would be less plausible for Germany. Ultimately, however, achieving both goals of comprehensive explanation and zooming in may profit from additional cases. This specific study thus raises general questions for all kinds of causal research projects: is it clear from my research design how many different theoretical claims are under investigation? Is it clear at what level of analysis my theories seek to account for phenomena? How large is the universe of cases related to these theories? When juxtaposing various theories, how to ensure that my selected cases are, on the one hand, comparable, yet differing in terms of the independent variables my theories identify?

Alexander De Juan (Chapter 10) would like to be able to make claims about all non-state actors engaged in violent conflicts. His theory is about situations in

which natural resources are dwindling. Of course, what constitutes salient resources differs across countries and even within countries and what counts as a salient resource may partly be the product of social construction. Nevertheless, Alexander De Juan seeks to engage with the literature that focuses on natural resources in a traditional sense (related to land, water and air) often in the context of climate change (Buhaug *et al.* 2010). Importantly, Alexander De Juan persuasively argues that resource deterioration, migration and conflict often occur at the subnational level. He rightfully pleas for the collection and analysis of data at the subnational, i.e. regional, level. He proposes the use of geospatial analysis as a method to collect and interpret such regional data, in his case from Darfur in Sudan. This disaggregation is certainly welcome, but if one assumes that actors or structures at the national level affect the situation and choices at the regional level, this should be included in the conceptual model and one should seek to hold constant the impact of the national level. Alexander De Juan elaborately discusses the pros and cons of different approaches to measure phenomena at the regional level and persuasively argues to use specifically sized grids as the optimal unit of observation (rather than administrative units, etc.). Rather than opting for single size grids, he proposes two different-sized circles around villages. Generally, choosing the 'right' level of analysis is a crucial but often tricky task, not only but especially for researchers of non-state actors in international security who have to ask themselves: to what extent does my research design give a lead to the level of analysis at which my theory offers an explanation? What does this imply for the actual unit of observation I select in my empirical research? In selecting cases can I control for the impact of variables at different levels of analysis?

Patrick Mello's (Chapter 9) research question makes the state the unit of analysis. Different theories point to variables at different levels (e.g. realism to military power; liberalism to parliamentary veto power). The advantage of QCA is that combinations of variables at different levels can point to causal pathways, thus allowing for the identification of the complementarity of theories (which may not be the same as the desired integration of theories). Yet the specific research question comes at a price: Patrick Mello's research question lumps together the decision to participate in the 2003 intervention in Iraq and the decision to contribute (militarily) to the post-war reconstruction. This obscures the position of countries such as Italy and the Netherlands, which opted for political support for the war and, later, for military support for reconstructing Iraq. In this case this would affect the so-called truth table (Table 9.5) and the conclusions regarding necessary and sufficient conditions. It does not jeopardize the valuable contribution of QCA per se, but it once more underlines the vital interest of linking precise research questions to theory before the use of specific methods can be assessed. Researchers should therefore always ask themselves: is my research question sufficiently specific about the phenomenon to be accounted for? Do my theories speak to precisely this phenomenon?

The objective of assessing the causal mechanism implies a data analysis method that would allow one to do just that. Andreas Kruck has chosen *process*

tracing to accomplish that aim (see George 1979b; Checkel 2008). His strategy of assessing whether covariation over time or across countries takes place allows him to draw plausible conclusions regarding causal links. Andreas Kruck complies with the prescription that process tracing requires us to pose the same questions to comparable cases and to standardize the data collection of all cases. In doing so, he avoids the problems of thick description, which may give valuable qualified and more complex information, but which entails the risk of losing sight of arguing that causal patterns can or cannot be observed.

The unravelling of causal mechanisms is, however, a challenging endeavour for researchers. The conceptual model that was chosen by Alexander De Juan, for example, cannot yet fully account for the causal chain: we know from the mobilization literature that the step from frustration to aggression requires the incorporation of processes of organization and mobilization (see Tilly 1978); we need to establish whether migrants moved because of the threat rather than for other reasons; we need to know whether variation in conflict in migration areas can be linked to differences in organization and mobilization. Still, in principle, such issues could be measured by taking the geospatial angle, provided data at the grid level are available.

The broader lesson from these examples is that (the growing number of) researchers who are interested in thoroughly studying causal mechanisms will have to address a number of crucial questions: does my research design specify the causal mechanism(s) that may underlie the theoretical approaches to answering my research question? Does my chosen method of analysing data, and the data itself, allow me to draw conclusions about the presence of the identified causal mechanisms? Are my selected cases sufficiently comparable that inferences about causal mechanism can be made?

In Patrick Mello's QCA approach, causality is effectively defined as identifying necessary and sufficient conditions. This renders important insights into his scientific problem: he identifies two pathways to participation in the 2003 Iraq War: the presence of a right-wing executive combined with absence of constitutional restrictions, and the absence of military strength combined with the absence of constitutional restrictions combined with the presence of parliamentary veto rights. Despite these results it remains unclear how these pathways represent the causal mechanisms belonging to the four theoretical approaches Mello sets out to test and, at some point, integrate: institutionalism, realism, liberalism and constructivism. Undoubtedly, this is addressed in the research from which he draws (Mello 2014), but the exact link with theory could have been developed more. Such a shortcoming is a frequent lacuna; maybe because linking empirics to theory is a very demanding task. If this link is underdeveloped, however, it may render it difficult to fully assess the explanatory power of the theories used (or of alternative ones). In the present QCA-chapter this makes it more difficult to assess under what circumstances QCA needs to be supplemented with process tracing in order to detect the causal mechanisms that can be traced back to theory (see George and Bennett 2005: 163). Again, this underlines the usefulness of a research design that

specifies the causal mechanism(s) that may underlie the theoretical approaches to answering one's research question.

What kind of knowledge is generated?

Andreas Kruck has succeeded in plausibly arguing that causal patterns have been observed – the goals of zooming in. Yet, it in order to meet the wider objective of a comprehensive explanation it might be useful to specify more precisely the wider class of phenomena and to place the theories that guide the case selection in the context of larger debates in International Relations about what behaviour to expect from states towards privatization of security, e.g. whether sovereign states can be expected to privatize security in the first place. It is interesting that new puzzles and blind spots are identified but it remains unclear whether they speak to the original research question. However, as said before, researchers often have to limit their analysis due to practical reasons, which can lead to trade-offs. From Alexander De Juan's analysis we learn that geospatial analysis helps us falsify the direct link between resources depletion and violent conflict and establishes the plausibility of an indirect link via migration behaviour to resource-plenty areas. Geospatial analysis enables us to expand the conceptual model (to account for causal mechanisms) to the extent that data at the grid level are available. Generalization of knowledge would nevertheless require controlling for influences at the national (and possibly transnational) level.

Patrick Mello's research underlines the possibility of making causal claims even when the number of cases is too low for quantitative analysis. Because QCA forces the researcher to be extremely explicit about operational definitions and the attaching of weights to variables it facilitates generalization. Although he does not discuss whether his empirical case (states' decisions regarding the 2003 Iraq War) is representative for all armed conflict, it is safe to say that the two causal pathways are plausibly valid for democracies' decisions to go to war in the post-Cold War era. The additional advantage of QCA in identifying outlier cases will help us to conduct supplementary in-depth case studies that trace the conditions that trigger deviant choices. These reflections on the three specific studies invite a final set of questions that will be relevant for many, if not all research projects seeking to make causal claims: is it clear in my research which class of phenomena I seek to theorize? Is it clear whether I consider them time- and space-bound and how that may limit the size of the phenomenon I seek to account for? Is it clear how this relates to the larger debates in IR theory?

Conclusion

First and foremost, the choice of method(s) follows one's conceptual model, which itself is based upon a theory-driven research question. In that sense, each method applied in this section of the volume may be a wise strategy depending on the problem at hand. The chapters in this volume make clear, however, that research question and conceptual model merit further explicit deliberation and

justification. This is necessary to be able to determine the nature of the knowledge generated (e.g. in terms of generalizability) and to assess whether the chosen methods of data collection and data analysis actually make forms of causal claims possible. Although all three methods (Structured Focused Comparison; Geospatial Analysis; QCA) seriously offer a tool to make plausible claims about causality, the chapters show that the former two, but possibly also the latter, may need additional tools (such as process tracing) to reveal underlying causal mechanisms. In order to be able to do so, we need to be more explicit about the causal mechanism that our theories suggest. All in all then, although we may have come closer to making inferences about causality, we may still be in Plato's cave observing shadows.

References

Buhaug, Halvard, Nils Petter Gleditsch, and Ole Magnus Theisen. 2010. 'Implications of climate change for armed conflict', pp. 75–101 in *Social Dimensions of Climate Change: Equity and Vulnerability in a Warming World*, R. Mearns and A. Norton, eds. Washington: The World Bank.

Checkel, Jeffrey T. 2001. 'Why Comply? Social Learning and European Identity Change'. *International Organization* (55)3: 553–588.

Checkel, Jeffrey T. 2008. 'Process Tracing', pp. 114–127 in *Qualitative Methods in International Relations. A Pluralist Guide*, A. Klotz and D. Prakash, eds. Houndmills: Palgrave.

Fierke, Karin M. 2005. *Diplomatic Interventions. Conflict and Change in a Globalizing World*. Houndmills: Palgrave.

George, Alexander L. 1979a. 'Case Studies and Theory Development: the Method of Structured Focused Comparison', pp. 43–69 in *Diplomacy. New Approaches to History, Theory and Policy*, P.G. Lauren, ed. New York: The Free Press.

George, Alexander L. 1979b. 'The Causal Nexus between Cognitive Beliefs and Decision Making Behavior: The "Operational Code" Construct', pp. 95–124 in *Psychological Models in International Politics*, L.S. Falkowski, ed. Westview: Boulder.

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

Guzzini, Stefano. 2016. 'Power and Cause'. *Journal of International Relations and Development*. doi:10.1057/s41268-016-0002-z.

Klotz, Audie, and Cecilia Lynch. 2007. *Strategies for Research in Constructivist International Relations*. Armonk: M.E. Sharpe.

Landman, Todd. 2003. *Issues and Methods in Comparative Politics: An Introduction*. London: Routledge.

Mello, Patrick A. 2014. *Democratic Participation in Armed Conflict: Military Involvement in Kosovo, Afghanistan, and Iraq*. Houndmills: Palgrave Macmillan.

Price, Richard, and Christian Reus-Smit. 1998. 'Dangerous Liaisons? Critical International Theory and Constructivism'. *European Journal of International Relations* (4)3: 259–294.

Tilly, Charles. 1978. *From Mobilization to Revolution*. New York: McGraw-Hill College.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Reading: Addison Wesley.