

RESEARCH ARTICLE

Searching for a middle ground? A spectrum of views of causality in qualitative research

Chiara Ruffa^{1*}  and Matthew Evangelista²

¹Department of Peace and Conflict Research, Uppsala University and Swedish Defence University, Uppsala, Sweden and

²Department of Government, Cornell University, Ithaca, NY, USA

*Corresponding author. Email: chiara.ruffa@pcr.uu.se

(Received 14 July 2020; revised 5 March 2021; accepted 8 March 2021; first published online 19 April 2021)

Abstract

Qualitative scholars exhibit a wide range of views on and approaches to causality. While some approaches reject causality from the outset, a large strand of qualitative research in political science and international relations does, however, pursue causal explanation. Qualitative scholars nevertheless disagree about what causality means. Our paper reviews what causality means within different strands of qualitative research and how qualitative scholars engage in causal explanations. We focus particular attention on the fertile middle ground between qualitative research that seeks to mimic the statistical model and research that rejects causality entirely. In broad strokes, we understand views of causality as lying on a spectrum and partly overlapping. Along the spectrum, we identify three main clusters: ‘positivist leaning,’ ‘postpositivist leaning,’ and ‘interpretivist leaning.’ Within each cluster, we identify the main traits and provide illustrative examples. We find merit in each of these three clusters of approaches and in the ongoing dialogue among qualitative scholars of different orientations. Understanding similarities and differences in the way various scholars address causality might encourage some to take steps along the spectrum and expand their repertoires to embrace elements of other approaches. By making these distinctions more explicit, we hope to be able to enhance our understanding of different views of causality and the extent to which they overlap and provide the potential for collaboration.

Key words: case study; causality; ethics; interpretivism; positivism

Introduction

Qualitative scholars exhibit a wide range of views on and approaches to causality.¹ Unlike in quantitative approaches, which typically aim to identify a relationship of cause and effect between an independent variable (x) and a dependent variable (y), qualitative researchers treat causality in a more nuanced way and often disagree among themselves about its importance. Qualitative approaches embracing interpretivism and post-structuralism, for example, typically reject causality from the outset, either on ontological grounds—it does not exist—or as irrelevant to their research endeavor (Bray, 2008; Salter and Mutlu, 2013). A large strand of qualitative research in political science and international relations does, however, pursue causal explanation. Qualitative scholars nevertheless disagree about what causality means and, in particular, whether it means the same thing in qualitative as in quantitative research. Our paper reviews what causality means within different strands of qualitative research and how scholars engage in causal explanation. We focus particular attention on the fertile middle ground between qualitative research that seeks to mimic the statistical model and research that rejects causality entirely.

¹By qualitative scholars, we mean researchers engaging with small- N empirical material and intensive strategies of data collection.

In broad strokes, we understand views of causality as lying on a spectrum and partly overlapping. Along the spectrum, we identify three main clusters. We build on the classification scheme proposed by Della Porta and Keating (2008), recognizing that not all scholars feel comfortable in the boxes where we place them and that some qualitative scholars will move around back-and-forth in those boxes during the course of their career. At one end of the spectrum, ‘positivist-leaning’ scholars embrace meanings of causality resembling those used among quantitative scholars, with the explicit use of variable language and with the aim of generalizing to a population of cases. They strive to capture causality by providing plausible accounts of thick processes and patterns, using detailed case studies representing broad phenomena of interest, engage in comparative, counterfactual, and/or process-tracing methods and are particularly attentive to research design. Some of them employ causal analysis in historical research with the goal of producing findings relevant for current public policy, including—in international relations research—foreign and military policies (Table 1).

Further along the spectrum, we find a cluster of approaches, which we label ‘post-positivist leaning’ that challenge some of the positivist assumptions: they acknowledge causality as complex and not easily captured and they place progressively less emphasis on associations between variables and evince more awareness of context and complexity. Designing research to maximize external validity may still be relevant but to a lesser degree. They may use causal and variable language, but often as ‘grammar’ to clarify the objective and purpose of research, without claiming to have demonstrated causality. Recognition of subjectivity and ‘positionality’ of the researcher vis-à-vis the research object becomes increasingly important. Moving further to a third cluster, which we label as ‘interpretivist-leaning’, we find researchers who are skeptical of quantitative standards of causality but who explore what they call situated, local, or singular causality (Hopf, 1998, 2013; Fujii, 2010; Lebow, 2015; Pouliot, 2007). These studies call into question several of the positivist assumptions and emphasize, along with the post-positivists, the importance of positionality and inter-subjectivity in the research endeavor.

We find merit in each of these three clusters of approaches and in the ongoing dialogue among qualitative scholars of different orientations. Qualitative researchers pursuing causal explanations, and situated along the spectrum we posit, often find themselves stuck between a rock and a hard place: their efforts to embrace causality are rejected by quantitative researchers for failing the standards of statistical inference (but see McKeown, 1986; Goodin and Tilly, 2006) and by critical scholars as misguided. Nevertheless, we find value in a middle ground that strives to capture causality, however imperfectly, while remaining aware of context and complexity. Understanding similarities and differences in the way various scholars address causality might encourage some to take steps along the spectrum and expand their repertoires to embrace elements of other approaches. Positivist-leaning qualitative scholars may benefit, for example, from reflecting more on context and positionality. Conversely, interpretivist-leaning ones may be tempted to think about the potential generality of their cases and consider plausible causal elements identified in their research. Without claiming to offer an exhaustive review, we proceed by providing examples of research that falls into the clusters along our posited spectrum.

Our paper has six parts. First, we provide a broad-stroke overview of three main clusters of usages of causality in qualitative research. Second, we elaborate on positivist-leaning assumptions of causality. Third, we review those approaches that relax some of the positivist assumptions and embrace post-positivism. Fourth, we reflect on interpretivist-leaning approaches and provide some examples. Fifth, we bring those three approaches together, highlighting what they still have in common. Sixth and last, we draw some conclusions in which we elaborate on the importance of our study for social science research in general.

Views of causality in qualitative research

Qualitative researchers disagree about the importance of causality in their research and they disagree about what causality means. Some of their differences partly mirror deeper ontological

Table 1. 'Approaches epistemologies and views of causality in the social sciences?' (Della Porta and Keating, 2008: 23)

	Positivist	Post-positivist	Interpretivist	Humanistic
Ontological issues				
Does social reality exist?	Objective, realism	Objective, critical realism	Objective and subjective as intrinsically linked	Subjective science of the spirit
Is reality knowable?	Yes, and easy to capture	Yes, but not easy to capture	Somewhat, but not as separate from human subjectivity	No, focus on human subjectivity
Epistemological issues				
Relationship between the scholar and his/her object	Dualism: scholar and objective are two separate things, inductive procedures	Knowledge is influenced by the scholar, deductive procedures	Aims at understanding subjective knowledge	No objective knowledge is possible
Form of knowledge	Natural laws (causal)	Probabilistic law	Contextual knowledge	Empathetic knowledge

disagreements, on whether social reality exists and is knowable, and divergent epistemological foundations regarding the relationship and possible distinctions between the researcher and the object of research. Based on key epistemological and ontological distinctions, Della Porta and Keating (2008: 23) have identified four approaches in social and political science: positivist, post-positivist, interpretivist, and humanistic. While they touch upon it, they do not explicitly elaborate on the notion of causality in qualitative research. We build on their categories for our discussion of causality.

Positivism, for Della Porta and Keating (2008), remains a relatively distinct approach because of its three fundamental assumptions: first, that reality is knowable and easy to capture; second, that the scholar and the object of knowledge are two separate things; and third, that the main form of knowledge is law-governed (and, hence, causal). By contrast, post-positivism acknowledges the existence of social reality, but one that remains 'only imperfectly knowable' and influenced by the scholar (Della Porta and Keating, 2008: 24). Interpretivists question or relax the assumptions of the other approaches. 'Here, objective and subjective meanings are deeply intertwined' and reality cannot be separate from human subjectivity (Della Porta and Keating, 2008: 24). Epistemologically, understanding the context is what matters most. Lastly, Della Porta and Keating (2008) introduced what they call humanistic approaches that place subjectivity at their center and are not interested in causality. From this perspective, social science is therefore 'not an experimental science in search of laws but an interpretative science in search of meaning' (Geertz, 1973: 5 cit. in Della Porta and Keating, 2008: 25). That the cluster we call interpretivism might also endorse this claim reinforces our point about the overlapping nature of these categories. In this paper, we do not explicitly address humanistic approaches, except to the extent that they overlap with interpretivism.

Because of these overlapping features, most qualitative research, we argue, lies on a spectrum of positivist-, post-positivist- and interpretivist-leaning approaches, even if for heuristic purposes they are portrayed in tabular form as a typology (Table 2). Moreover, despite our notion of a spectrum, we do not claim that epistemological and ontological views correspond in a linear way to views on causality. Within each ontology or epistemology, several different views of causality may co-exist, ranging from more formal views looking for an association between an independent and a dependent variable to more contextual views of causality within the same epistemological approach. There is, however, a broad correspondence between ontological and epistemological views, understandings of causality and emphasis on different aspects of the research endeavor.

Those that we label as positivist-inclined constitute a large cluster of approaches, ranging from ones putting qualitative work at the service of causal inference to ones that may use causal language as a grammar to explain parsimoniously the objectives and scope of their research, to those

Table 2. A spectrum of views on causality in qualitative research

	Positivism	Post-positivism	Interpretivism
Views of causality	Causality as relation between independent and dependent variable	Complex causality; Language of variables sometimes used as grammar	Contextual causality OR No causality
Emphasis on research design for generalizability; often cases used as complement for quantitative research	→		
Emphasis on positionality, ethics and data collection	→		
Examples	<ul style="list-style-type: none"> • Evangelista (1986) • Knopf (1998) • Krause (2017) • Bove <i>et al.</i> (2020) • Talmadge (2015) 	<ul style="list-style-type: none"> • Historical institutionalist approaches • Mahoney <i>et al.</i> (2016) • Skocpol (1979) • Lebow (2015) • Ruffa (2018) 	<ul style="list-style-type: none"> Fujii (2010), Pouliot (2007), Hopf (1998, 2013), Bode and Karlsrud (2018), Von Billerbeck (2017), Autesserre (2010)

that adopt a limited view of causality, avoid using the language of variables, yet still make nuanced causally informed claims. What most of these positivist-leaning approaches share in common is the goal of maximizing external validity in the service of generalization. In this respect, they pay particular attention to research design and the question of what something is a ‘case of.’ Post-positivist scholars are more critical of the possibilities of generalizing beyond the case(s) under study and, typically, are more cognizant that it is never fully possible to distinguish the object that is being studied from the subject of such study. In their view, causality is complex and difficult to capture. Further along the spectrum, we find another broad category—interpretivism—where some scholars recognize causality only as contextual and others reject it entirely. Overall, interpretivists are interested in a better understanding of how the researcher influences the field and vice versa and how the inter-subjective understanding of certain phenomena is constituted. Scholars of this approach are exceptionally reflexive about strategies of qualitative data collection—particularly interviews and ethnography—and their ethical and practical implications.

By making these distinctions more explicit, we hope to be able to enhance our understanding of different views of causality and the extent to which they overlap and provide potential for collaboration. In the remainder of the paper, we further elaborate on these three different views of causality by providing illustrative examples from existing research.

Positivist understandings of causality

Many scholars among the diverse group working in a positivist orientation acknowledge the importance of developing and testing theories with a causal logic that establishes a connection, for example, between independent and dependent variables of interest. Notwithstanding important differences, these scholars all identify a universe of cases, where their proposed respective theories should be relevant. They design their qualitative research in a way that aims to generalize beyond the case(s) at hand and they aspire to fulfill certain positivist criteria, such as external validity and falsifiability. Attentive to issues such as selection bias, they understand causality in a way similar to how quantitative scholars do.

For instance, in her study about the battlefield performance of authoritarian regimes, Talmadge (2015: 37) writes ‘because I want to maximize external validity, I begin by seeking

out cases that display what Slater and Ziblatt call “typological representativeness”—cases that represent the full range of possible variation of the dependent variable’. Krause (2017: 13), in proposing ‘movement structure theory’ (MST) to account for the success of armed nationalist groups in his book, *Rebel Power*, explicitly embraces the desiderata of large-*n* studies. He writes ‘as much as I would have liked to test MST across the universe of cases in the future... I had to start with a small number of cases (...) I selected movements with diverse values on the dependent variable.’ As these studies suggest, a defining feature of much qualitative positivist research has been its dependence on criteria developed to evaluate quantitative research.

Such dependence likely stems from the persistent influence of the 1994 study by King and co-authors, *Designing Social Inquiry*—a book intended to establish the ground rules for what constitutes good qualitative research. Mahoney and Goertz (2006: 228) reflect on how King and co-authors offered quantitative work as the standard to which qualitative researchers should aspire when it comes to causation. Their book ‘was explicitly about qualitative research, but it assumed that quantitative researchers have the best tools for making scientific inferences, and hence qualitative researchers should attempt to emulate these tools to the degree possible.’ Beck had a similar reaction: ‘when I first read King and co-authors (1994) (KKV), I was excited by the general theme of what seemed to me the obvious, but usually unspoken, idea that all political scientists, both quantitative and qualitative, are scientists and governed by the same scientific standards.’ The authors, he suggested, ‘clearly were preaching to the qualitative researcher, and the subject of the sermon was that qualitative researchers should adopt many of the ideas standard in quantitative research. Thus, to abuse both the metaphor and KKV, one might compare KKV to the view of the Inquisition that we are all God’s children’ (Beck, 2006: 347).

Other scholars, such as Brady *et al.* (2006), while embracing a pluralistic view of methodology, nonetheless remain pretty much wedded to a particular, quantitatively inspired, view of causality. Regarding methods, they contrast data set observations (DSOs) using a quantitative logic of comparison with causal process observations (CPOs) which are well suited to study context, process, and mechanisms. While DSOs are variables and cases that constitute ‘the basis for correlation and regression analysis,’ CPOs are ‘insights or pieces of data that provide information about context, process, or mechanism and that contribute distinctive leverage in causal inference’ (Brady *et al.*, 2006: 353). They find merits in both approaches but highlight what they consider the drawbacks of CPOs when not used according to their own preferred approach: ‘information that could be treated as the basis for CPOs may simply be presented in a narrative form, with a different analytic framing than we have in mind here’ (*ibidem*, 328).

Unlike the authors of *Designing Social Inquiry*, Mahoney and Goertz (2006: 227) do not seek to subsume qualitative research into the standards of quantitative work but write instead of ‘two traditions’ or ‘alternative cultures’ which adopt distinct positions on a range of issues, including causality. Seeking to explain particular cases, they argue, ‘qualitative researchers often think about causation in terms of necessary and/or sufficient causes’ and adopt appropriate methodological choices, such as Mill’s methods of agreement and difference and explanatory typologies (*ibidem*: 232). Constructing typologies and arraying cases in a cell matrix is especially well suited for conducting a controlled comparison of historical cases (George and McKeown, 1985; George and Bennett, 2005: ch. 11). In this type of qualitative work it is *not* recommended to select cases randomly (Seawright and Gerring, 2008; Nielsen, 2016; Ruffa, 2020). On the contrary, strategic case selection helps to evaluate causal claims.

Comparative case studies using the most similar system and most different system designs, for example, may help to discard potential confounders and ‘isolate’ the independent variable of interest. These designs can be fruitfully combined with a process-tracing method. As Bennett and Checkel (2015: 29) point out, ‘most-similar cases rarely control for all but one potentially causal factor, and process tracing can establish that other differences between the cases do not account for the difference in their outcomes.’ This is the approach Evangelista (1986) took in early work on the determinants of deployment of US nuclear weapons in Europe, using a

comparative case design to identify factors that were present in decisions both for and against the deployment of certain weapons (e.g. perceptions of a Soviet threat, need to resolve alliance political disputes with a 'military fix') in order to isolate a necessary cause that was present only in the decisions to deploy (what he called technological entrepreneurship). In a similar fashion, Knopf (1998) conducted comparative case studies (along with a large-*n* analysis) to identify peace movements as a necessary condition of US agreement to limit nuclear weapons through negotiations with the Union of Soviet Socialist Republics (USSR).

The causal claims of such studies were vulnerable to various criticisms, and these should be borne in mind by qualitative researchers undertaking case studies. Historians, for example, have pointed to the 'indeterminacy of historical interpretation' in itself: if historians disagree about a particular case, social scientists cannot uncontroversially deploy those cases in comparative perspective to yield valid theoretical insights (Njølstad, 1990), even less so when other social scientists disagree with their interpretation of the cases (Risse-Kappen, 1986). The notion of 'configurative causality,' developed by Senghaas (1972, 1990) suggests that if one variable is removed from a complicated process that generates armaments or conflict more generally, other variables could play a redundant role and produce the same conflictual result. The notion is similar to what Bennett and Checkel (2015), Mahoney and Goertz (2006), and others have described as the problem of 'equifinality.' As George and McKeown explained, 'as is often the case in social science, the phenomenon under study has complex multiple determinants or alternative determinants rather than single independent variables of presumed causal significance' (George and McKeown, 1985: 27).

Positivist-oriented scholars have developed ways to bolster the validity of process-tracing studies by connecting them to a well-recognized body of existing theory using the methods of congruence and counterfactual reasoning (George and Bennett, 2005). The underlying idea here is that outcomes are congruent with the expectations of a particular theory (Blatter and Haverland, 2012 cit. in Ruffa, 2020: 1144).

George and McKeown (1985: 30–31) have cautioned, however, that 'ways must be found to safeguard against unjustified imputation of a causal relationship on the basis of mere consistency, just as safeguards have been developed in statistical analysis to deal with the possibility of spurious correlation.' They particularly call attention to the risk that the consistency is spurious and based on 'idiosyncratic aspects of the individual case'. Ruffa (2020: 1144) has advocated for the congruence method as a pragmatic approach to the complexities of fieldwork. She was too cautious that 'the method of congruence is insufficient when it comes to establishing causal mechanisms,' and should be complemented by verifying 'the consistency of observable implications of theory at different stages of the causal mechanism.'

A related way to assess the generalizability of a causal relationship inferred from the congruence method is simple counterfactual reasoning. 'In considering some plausible counterfactual situation chosen for its general similarity to the situation actually analyzed,' write George and McKeown (1985: 33–34), 'the researcher may attempt the following thought experiment: If my argument about causal processes at work in the observed case is correct, what would occur in the hypothetical case if it were governed by the same causal processes?' Scholars have proposed a number of criteria that good counterfactual exercises should meet in order to strengthen their plausibility. Among them, 'consistency with well-established theoretical generalizations' and 'consistency with the empirical evidence' resemble the congruence method (Levy, 2015: 395–396; Campbell, 2018).

A study that combines process-tracing, counterfactuals, congruence, and two types of comparison is Evangelista's book on the development of new weapons in the Soviet Union and the United States during the Cold War (Evangelista, 1988a). Based on the notion of 'domestic structure' in the literature on comparative foreign policy (Evangelista, 1997), the author argued that the highly centralized and secretive Soviet state would tend to stifle technological innovation, whereas in the United States, strong societal pressures from scientists, weapons laboratories, and military corporations would promote new technologies from the 'bottom up,' and find

opportunities to do so in a decentralized and fragmented state apparatus. On the Soviet side, once innovation was adopted—usually in response to its development in the United States, resources could be effectively mobilized to copy and put into mass production, in a ‘top-down’ process. These expectations were congruent with theories of economic and organizational innovation (Rogers, 1983) and with the comparative historical literature on late-industrializing countries (Gerschenkron, 1963; Moore, 1966). Because he literally hypothesized a *process* of five stages—different ones for each country—Evangelista’s study lent itself to a method of process-tracing, relying mainly on historical materials. It was comparative between the two countries and within each country (comparing the development of several weapons technologies). It aspired to positivist criteria such as *falsifiability*, with a search for disconfirming evidence at each stage, based on counterfactual hypotheses. The theory would be undermined, for example, if a Soviet innovation were developed and promoted by scientists early in a bottom-up fashion, before the United States had developed something comparable, or if the US government responded in a top-down way to a Soviet innovation before any US scientists had developed the weapon. Although Evangelista made no claims for another common positivist criterion—*replicability*—a recent study suggests that his basic comparative framework has endured. Employing Evangelista’s bottom-up and top-down generalizations and five-stage comparative framework, Meyer *et al.* (2020) examined the (ultimately unsuccessful) development of radiological weapons in the United States and the Soviet Union. Relying on extensive archival data, interviews, and declassified documents from each country, their meticulous process-tracing exercise largely confirmed Evangelista’s comparative generalizations about US and Soviet behavior.

In sum, the qualitative scholars we call positivist-leaning have sought to meet many of the criteria of positivist research, such as external validity and falsifiability. Their small- or medium-*n* studies rely on a variety of methods, including process-tracing, structured focused comparisons, CPOs, and counterfactual reasoning.

Challenging (some) positivist assumptions on causality: post-positivism

Many qualitative scholars are unwilling to embrace certain positivist assumptions, for example, that single variables exert causal influence in similar ways across a wide range of purportedly similar cases. Reflected in the critiques by George and McKeown and others is a basic challenge to those assumptions. Instead, the critics argue, context matters. Goodin and Tilly made this point some years ago, and—along with many of the authors we have cited here—they chose the work of King and co-authors as the focal point for their argument. They start by claiming that ‘although any thinking political analyst makes some allowances for context, two extreme positions on context have received surprisingly respectful attention from political scientists during recent decades: the search for general laws, and postmodern skepticism’ (Goodin and Tilly, 2006: 7). At one extreme, the ‘search for general laws’ approach considers context as noise—a misunderstanding that stems from its notion of causality. Goodin and Tilly label King and coauthors’ contribution as a ‘spirited, influential and deftly conciliatory synthesis of quantitative and qualitative approaches to social science’ (*ibidem*). Even though apparently conciliatory, King and co-authors still end up arguing that ‘the final test for good social science is its identification of casual effects,’ defined as ‘the difference between the systematic component of observations made when the explanatory variable takes one value and the systematic component of comparable observations when the explanatory variable takes on another value’ (King and co-authors, 1994: 82 *cit.* in Goodin and Tilly, 2006: 7). ‘This seemingly bland claim,’ they maintain, ‘turns out to be the thin edge of the wedge’ (Goodin and Tilly, 2006: 8), in effect a claim that neither context nor mechanisms matter—only the relationship between variables. Postmodern skepticism, at the other extreme, suggests that context is the very object of political analysis: ‘the complex, elusive phenomenon we must interpret as best as we can’ (Goodin and Tilly, 2006: 8). In this

view, a qualitative researcher explores a research question by working with nuances and considering context but eschews any claims about causal effects.

We insist that there is a vast middle ground between the model of qualitative research that takes statistical regression as its inspiration and one that completely rejects the notion of causality. Scholars have staked out sections of that ground and tilled it using various tools, including the process-tracing and comparative methods we discussed in the previous section. In a further critique of *Designing Social Inquiry*, Lebow (2015: 408) argues that the statistical model for qualitative research ‘would make sense only in a world in which systematic factors explained a significant percentage of the variance. There is no reason to believe that we live in such a world, and King and co-authors offer no evidence in support of this most critical assumption.’ He is skeptical of a ‘naïve understanding of cause’ that ‘builds on the concepts of succession and continuity, and the assumption that some necessary connection exists between them’ (ibidem, 409). Yet he is not a ‘postmodern’ skeptic if that means rejecting causality altogether. Instead, he identifies ‘a large category of events that are causal but non-repetitive,’ what he calls *singular causation* (ibidem, 409). Consistent with George and McKeown’s (1985: 27) emphasis on ‘complex multiple determinants or alternative determinants rather than single independent variables of presumed causal significance’ is Lebow’s advocacy of ‘causal maps.’ Along with Goodin and Tilly, and *contra* King and co-authors accord high priority to context:

Neither behavior nor aggregation can usually be attributed to single causes. To tease out multiple causes and relationships among them we need to employ the so-called factual and counterfactual arguments. They help us identify pathways that might qualify as causal and construct multiple causal narratives, or what I describe as causal maps. Narratives of this kind allow richer depictions of the world and more informed judgments about possible underlying ‘causes,’ the level at which they are found, and their key enabling conditions. The mapping of many outcomes and their possible causes will ultimately provide us with a more useful understanding of the world than the search for regularities, covering laws, or the properties philosophical realism associates with things. But we must be clear that context is almost always determinate, so causal maps and any generalizations they allow are at best starting points for forecasts, never for predictions (Lebow, 2015: 410).

Challenging the assumptions of approaches that understand the context of noise and causation as amenable to general laws opens up space for a broader view on causality, one that we associate with post-positivist approaches. Several constructivists and historical institutionalists, in particular, have paved the way for this to happen. They are less interested in maximizing external validity and more sensitive to understanding positionality. While some scholars may still talk with the language of variables, this becomes merely grammar. For instance, when Ruffa in her own research argues that military cultures influence the ways in which soldiers interpret the context in which they operate and hence how peacekeepers behave, she phrases her argument causally even though she cannot establish for sure that military cultures *cause* the behavior of peacekeepers (Ruffa, 2018). Her research makes use of a positivist-leaning research design: she identifies two pairs of cases—French and Italian units deployed in both the UN mission in Lebanon and the North Atlantic Treaty Organization (NATO) mission in Afghanistan. Within each mission, those pairs of French and Italian units are argued to display features of a most-similar system design, with several factors that are similar among those cases. By focusing on two very different kinds of missions, Ruffa aims to maximize external validity. Notwithstanding all these steps, Ruffa’s work is based on ethnographic data—namely observation, interviews, focus groups, and questionnaires. The kind of material she works with—which is highly subjective—and the very questioning of her own positionality in relation to such complex sites of investigation, makes her approach distinct from the positivist one and fitting within the middle-ground which we label post-positivist box.

Historical institutionalism has played a crucial role in developing theories that entail a more sophisticated understanding of causality. As Bennett and Elman (2006: 250) put it ‘this requires that we adapt and develop our methods, whether formal, statistical or qualitative to address the kinds of complexity that our theories increasingly entail.’ They identify ‘several different phenomena that exhibit causal complexity, including tipping points, high-order interaction effects, strategic interaction, two-directional causality or feedback loops, equifinality, and multi-finality’ (Bennet and Elman, 2006: 251). Several of those phenomena are not well suited to be studied using traditional statistical and qualitative methods (Bennett and Elman, 2006: 264).

Complex causality entails precisely a greater understanding of context and, as historical institutionalists remind us, of the crucial importance of time and sequences. A key element for understanding the existence of causal relation is time. Critical junctures, gradual change, and path dependence all entail the use of causal temporal concepts. To understand patterns of change one has to unpack both the causal processes and the temporal processes that constitute the pattern of change. To illustrate, when talking about critical junctures: ‘a relatively short period in time during which an event or set of events occur that has a large and enduring subsequent impact’ (Mahoney *et al.*, 2016: 7). While antecedent conditions may be important, it is only the critical juncture that is sufficient for leading to the outcome. Skocpol’s (1979) celebrated study, *States and Social Revolutions*, provides a classic example of the inadequacy on single-variable approaches to historical explanation. Her account of the French, Russian, and Chinese revolutions credits the outcomes to a conjuncture of peasant rebellion, loss in war, and breakdown of the state apparatus that could not be explained through any statistical methods. She provides comparative evidence from other states—England, Prussia, Japan—that manifested some, but not all, of the elements of a successful social revolution in the proper configuration.

An interpretivist view of causality

If we move further along the causality spectrum, we find the scholarship that is less interested in the generalizability of theories and findings and more focused on how we can capture knowledge, given the challenges of separating the subject from the object of research. Following this logic, research design becomes less relevant, as do considerations relating to generalization. To illustrate, when von Billerbeck (2017: 7) introduced the core argument of her book about local-ownership in UN peacekeeping, she first contextualizes ‘the issue of local ownership through an analysis of the discourse of ownership’ before moving on to examine ‘the understandings of local ownership on the part of UN and national actors as well as the different ways in which the UN operationalizes the concept.’ In order to do so, she analyzes and interprets a wealth of interviews, ranging from in-depth structured, semi-structured, and unstructured, which she reflects upon. Importantly, though, her research design does not seek to maximize external validity. She writes that ‘while this study focuses on the UN’s peacekeeping practice as a whole, in order to illustrate how the UN operationalizes local ownership and to capture the national perspective on ownership, I use the UN peacekeeping mission in DRC (MONUC) as a primary case study’ (von Billerbeck, 2017: 7). She complements this case with other less detailed shadow cases but her main focus remains Congo and its profound contextual richness.

Another relevant example is an article by Bode and Karlsrud (2018: 458) grappling with the obvious gap between the apparently solid normative foundations regarding the protection of civilians and the wide variation in the implementation of the norm. They argue that normative ambiguity is a fundamental feature of international norms. Combining a critical-constructivist approach to norms with practice theories, they explicitly adopt an interpretivist approach. In their article, one does not find a substantive discussion on the research design they chose but rather a discussion about their strategies of data collection. Their empirics draw ‘on an interpretivist analysis of data generated through an online survey, a half-day workshop and interviews with selected delegations.’

Contextual considerations and reflections on data collection in qualitative work are of key importance to capture causality and, perhaps even more importantly, to understand the specific context in which causal processes unfold. Non-standardized measures are best captured by in-depth immersions in specific contexts (Kapiszewski *et al.*, 2015). When researchers delve into data collection, they need to make use of such empirics in the most ethical and most beneficial way for their own research. Strategies of qualitative data collection are essentially of three kinds: document analysis, interviewing, and observation. The aim of this section is not to provide a detailed overview of those but rather to discuss the dangers and opportunities that arise when striving to capture causality in qualitative research, particularly in relation to ethical considerations and causality itself.

The first—and to some extent less ‘problematic’ from an ethics points of view—kind of data collection entails document analysis, ranging from archival documents to the policy reports, diaries and military doctrines, to name a few. While they may be hard to obtain, these materials are crucial to trace a posited causal process but also somewhat less problematic from a positionality point of view, since the researcher cannot affect it. Scholars may disagree on the interpretation of documents, so making them available to reviewers and skeptics bolsters the credibility of the method.

The second kind of data collection concerns interviewing, including oral histories, focus groups and individual interviews. In all three kinds of interviewing, a researcher’s self-presentation and personality, the respondent’s perception of the interviewer’s identity and personal traits interact with the research context and shape the interpersonal dynamic of an interview and thus the data collected. In some cases, the interviewer’s physical traits—skin color, most notably—or gender can make a difference in facilitating or hindering interviews and can also entail consequences for the subjects of the interviews. In one case, an African–American researcher was able to overcome a general anti-American sentiment to interview subjects in post-conflict Rwanda and gain information about relations between Hutu and Tutsi that would not have been confided to a White colleague (Davenport, 2013). The same scholar found that ‘when doing research in Northern Ireland my blackness helped me initially interact with the Catholic/Republican community but did nothing to assist me with the Protestant/Loyalist one. Interestingly, the former seemed to view African Americans as a similarly oppressed minority who fought and won against a similarly oppressive majority. They even had a mural of Frederick Douglas on one of the main streets, and also referred to themselves as the “Blacks” or N-words of Europe’ (Davenport, 2013). Half of the population in some traditional societies is out of bounds for male researchers, but open to meeting with female ones. In other cases, depending on the context, interview opportunities might instead favor White men in suits and ties.

Oral histories require that the interviewer not feature prominently but rather establish a suitable starting point and ensure a good flow in the narrative. In oral histories, the questions asked are open-ended and the narrative is structured chronologically, thereby allowing the perceived causal connections to emerge. The best oral histories also engage the relevant documentary record, if available, in an iterative fashion—as the interviewer returns to reengage the subject and solicit responses to facts that the documents have surfaced.

Focus groups mirror, in a fine-grained way, group dynamics, but group narratives—including shared subjective causal accounts—may be over-represented. Focus groups can also reveal tensions and disagreements within the group; possible contestation of the dominant group narrative; and they can be productive if there is disagreement on certain issues. In interviews, the role of the interviewer is more active and opportunities arise for manipulation of information to fit preconceived causal expectations. Interviewers should be careful not to ‘lead the witness,’ but to ask questions in a way that would elicit disconfirming information or information that would support rival causal accounts. Overall, to trace and identify potentially causal processes, it is crucial to learn and practice active listening: listen and interpret what is being said, process and respond.

Finally, ethnographic approaches, such as observation, provide the opportunity to study phenomena that would otherwise remain hidden but at the disadvantage that the researcher is affecting what is being researched. In qualitative research, it is crucial to balance reflexivity and strive for objectivity and reflect on how one's own identity influences the research experience. This is also why the triangulation of information with several sources and deep immersion in the context are of crucial importance.

There are obvious weaknesses in tracing causality that are inherent to interviewing. We may prime our interviewees to establish relations that do not exist. We may misunderstand or ignore hidden group dynamics that may be shaping the kinds of answers we get. We may strive to find and give too much prominence to explanatory factors that in fact co-exist with others. Recent interpretivist work has placed ethics to the forefront and in ways that are important also to better understand causality. In her book, late Fujii (2018) introduced the concept of relational interviewing, based on an interpretivist, rather than a positivist methodology. Her interpretivist understanding is clear from early on:

The world is what people make of it. The meanings they give to 'money,' 'race,' or 'witchcraft,' for example, constitute the very existence of these concepts. But for shared understandings about the worth of money, people would not work two jobs, play the stock market or rob the corner store. Without historically situated understandings of what it means to be 'Black' and 'White' in America, poor Irish immigrants would not have worked so hard at becoming White (Roediger, 2007), and protesters today would not organize around the claim that 'Black Lives Matter.' (Fujii, 2018: 27)

How can qualitative researchers know whether certain statements support or shed light on underlying causal relations? There are tools that we can use to understand what is really going on—that is for our study to understand where causality lies. In this context, metadata are crucial and are defined as 'spoken and unspoken expressions about people's interior thoughts and feelings' (Fujii, 2010: 232). Along similar lines, Fujii also helps us to understand how to use the surrounding context to better assess the reliability of an answer. Fujii argues that, often, more information than the direct responses to questions can come to the fore during interview research. Based on her own research in post-genocide Rwanda in 2004, Fujii identifies four types of metadata, namely rumors, invented and untruthful statements, denial, evasions, and silences. Metadata is important, Fujii argues, because it provides valuable information about current political and social landscapes which, in turn, influence testimonies and should therefore inform conclusions the researcher may draw based on the testimonies. One important example is the role that silence plays: 'those attending to metadata know that words can hide just as much as silences can reveal' (Fujii, 2010: 239). She gave metadata the same status as other data: 'metadata are not ancillary or extraneous parts of the dataset, they *are* data' (emphasis in original, Fujii, 2010: 240).

Along similar lines, Fujii helps us explore if and how non-ethnographic studies can benefit from ethnographic observations (i.e. data), drawing on five stories from her own fieldwork in Rwanda, Bosnia and the United States, as well as other anecdotes from travels (Fujii, 2015). She argues that ethnographic, namely observation-data can provide researchers with additional context on the research and their own positionality in the field. Yet, according to Fujii, political science research is lacking theories on the potential of 'accidental ethnography'—by which she means observations or instances when the researcher is not conducting interviews or other forms of research (Jennings, 2019). Such moments, unlike other forms of data collection in the field, are unplanned, and emerge from the daily life around the researcher. Fujii proposes these moments should be documented and analyzed systematically. The method of turning 'accidental ethnography' into data works as follows: first, the researcher takes notes of accidental moments throughout the field research period without prejudging how they may, later on, relate to the actual research (explicit consciousness). Second, the researcher writes down such

observations as they occurred. Third and last, the researcher reflects on the observations in an effort to link accidental moments and what they convey about the social and/or political context to the overall research questions.

On what these views have in common: theory, process, concepts, and contexts

Regardless of their views on causality, we find that all three approaches we have reviewed share at least four common features: a keen interest in theorizing, process, thick concepts and context. Much of the qualitatively oriented literature is interested in the process, how an *explanandum*—the thing to be explained—is accounted for by an *explanans*—the ‘story’ that explains the object that is being explained. The process is the concatenation of conditions, perspectives, views, turning points that collectively may plausibly explain why Y happened. The objective of qualitative approaches interested in causality is usually to describe and explain such a process: either by developing or testing theories or by exploring some specific part of the process. Unlike in quantitative work, the starting point of this kind of research is rarely a correlation—with the exception of mixed-method work that starts with a large-*n* analysis to identify the relevant cases for closer examination—but rather a puzzling fact. And the objective is to identify, test or develop a theory that can help us understand the puzzle. Consider, for example, why one of the most gender-neutral countries in the world (Sweden) has had a hard time retaining women in its armed forces.² Starting from such a puzzle, one may then start exploring its causes, identifying relevant theories to test. In this respect, theory plays a central role in qualitative research interested in causality.

In this respect, the process connecting the independent variable to the dependent variable—sometimes referred to as causal mechanism—is of crucial relevance and this is also what makes positivist qualitative research different from quantitatively-oriented one. When talking about quantitative research, Hedström and Swedberg (1998: 324) note how ‘it has contributed to a variable-centered form of analysis that pays scant attention to the processes that are likely to explain observed outcomes.’ By contrast, qualitative research puts processes at its core and does not see causal mechanisms as an “add-on” to make sense of an interesting correlation. Nor does it pick a causal mechanism “off the shelf,” borrowed from a comparable context and transferred to another. On the other hand, if the close qualitative investigation reveals causal mechanisms that are widely evident elsewhere, such contributions to more generalized patterns should be welcome.

In seeking agreement between quantitative and qualitative researchers, Bove, Ruffa, and Ruggeri (2020: 33), recently defined causal mechanisms as ‘further specification of theories’. The three authors struggled to find the right level of analysis at which a mechanism would play. Johnson provided some guidance when writing: ‘mechanisms operate at an analytical level below that of a more encompassing theory, they increase the theory’s credibility by rendering more fine-grained explanations’ (Johnson 2002: 230–31 cit. in Checkel 2008: 115). At the same time, Bove, Ruffa and Ruggeri followed Gerring’s (2010: 1499) note of caution against a ‘mechanismic’ view—that is a mechanism-centered perspective. A mechanismic view of a mechanism would risk having a mechanism artificially juxtaposed to connect independent and dependent variable and *de facto* too detached from the theory itself. While finding an appropriate level for a mechanism, we find that the focus on the process and the importance of theory are of fundamental importance for the research endeavor of any kind. Gerring (2008) proposed a definition of mechanism that is widely acknowledged as valid across the spectrum: ‘a causal mechanism is the pathway or process by which an effect is produced or a purpose is accomplished; mechanisms are thus relational and processual concepts, not reducible to an intervening variable.’

²See for instance: http://resmilitaris.net/ressources/10259/88/res_militaris_article_jonsson___osterberg_effective_recruitment_of_women.pdf; <https://app.dimensions.ai/details/grant/grant.8662122>

Moreover, paying attention to the pace and timing in which processes unfold is also of crucial importance. For instance, when Ruffa started observing the processes of peacekeepers' behavior, it looked like certain military organizations had a certain type of military culture and, when they deployed, their military culture 'deployed' with them: Once deployed, the contingent started operating in line with its pre-existing military culture. Observing the consistency of the troops' behavior before and during deployment was crucial to develop her theory. To be fair, similar dynamics can also be leveraged quantitatively but are more assumed and derived than qualitatively developed: for instance, Moncrief (2017) finds that a peacekeeping mission may carry its own norms and socializing processes that either constrain or facilitate the emergence and endurance of sexual exploitation and abuse or SEA. Quantitatively, he draws from a dataset of SEA allegations between 2007 and 2014, as well as the first publicly available data from the United Nations that identify the nationalities of alleged perpetrators. He finds evidence suggesting that 'SEA is associated with breakdown in the peacekeeping mission's disciplinary structures, and argue that the lower levels of command are the most likely sites of this breakdown' (Moncrief, 2017: 716). He also finds a positive but non-statistically significant association between SEA, the proportion of a mission's troops sent by militaries with previous histories of widespread or systematic sexual violence during wartime. While otherwise compelling, his argument cannot capture timing and micro-level processes that could be captured qualitatively.

Timing is also relevant to the question of 'when and where to begin and stop in constructing and testing explanations' through process tracing (Bennett and Checkel, 2015: 12). Evangelista (2015), for example, found that explanations for key Soviet decisions that contributed to the end of the Cold War would imply a consensual decision process if the researcher stopped at a certain point, but a conflictual one if the research continued beyond the decision itself to reveal opposition after the fact. Whether the decision was conflictual or consensual bears directly on which theory of political change exhibits more explanatory power (Evangelista, 1991).

As much as theory and process matter, as crucial are the concepts that contribute to the theory. One important peculiarity of qualitative research is that it works with thick concepts. This means that the research endeavor may entail developing and constructing new concepts or challenging hidden assumptions and dilemmas (see for instance: Belloni, 2007). Qualitative researchers work with concepts displaying high levels of conceptual validity, which are often not as well captured by standardized measures typical of datasets. When Bove *et al.* (2020) collaborated to develop the concept of *military involvement in politics* they struggled precisely with that issue. In fact, the conceptual problem is an old one and led to much confusion about, for example, the nature of the early Soviet regime after the Red Army was formed (Evangelista, 1988b). For quantitative studies, military involvement in politics can be operationalized according to whether a defense minister has a military background or not. That is an easy but not a highly valid measure from a conceptual standpoint: it will depend quite heavily on contextual factors, such as whether the country has conscription or a volunteer system. For instance, we judge contemporary South Korea to have a low level of military involvement in politics. Yet most of its 'civilian' defense ministers have had a 'military' background in the sense that they served in the army. Due to living in a country with an exceptionally high level of threat and a widespread conscript system, every male citizen is obligated to compulsory military service.

Conclusions and ways forward

There are several limitations and opportunities inherent to exploring causality qualitatively. In terms of limitations, the first is whether the causal process one observes is merely in the eyes of the beholder. This is the risk of 'cherry picking' evidence to develop an argument. The second is whether the search for supporting evidence makes the researcher blind to potential competing explanations. The third is whether one modifies one's own explanation in light of what the research discovers. This is a natural part of historical research, but social scientists should be self-

conscious about it and not insist that they have separated their theory development from their theory testing when the process is in fact more iterative—as it is in much scientific research (George and McKeown, 1985: 38). The fourth limitation has to do with the lack of transparency in qualitative research, which is still a perplexing matter and one that is difficult to address (Moravcsik, 2014). The ‘best practices’ of process tracing (Bennett and Checkel, 2015) are designed to cope with many of these limitations. At every step of the proposed causal chain, for example, researchers should explicitly entertain competing explanations and posit counterfactual developments that would undermine their preferred account. Bennet and Checkel (2015: 19) offer this advice:

Evidence that is contrary to the process-tracing predictions of an explanation lowers the likelihood that the explanation is true. It may therefore need to be modified if it is to become convincing once again. This modification may be a trivial one involving a substitutable and logically equivalent step in the hypothesized process, or it could be a more fundamental change to the explanation. The bigger the modification, the more important it is to generate and test new observable implications to guard against ‘just so’ stories that explain away anomalies one at a time.

The conjunctural process posited in Skocpol’s theory of social revolutions illustrates this point well. Given that her universe of cases is small, because her definition of a social revolution is rather precise, one wonders how well Skocpol’s theory travels to other cases. The 1979 Islamic Revolution in Iran, which took place as her book was published, seems to fit her definition of a social revolution, yet there was no widespread peasant rebellion. In a subsequent ‘modification’ of the process she developed from her other cases, Skocpol (1994: 281) suggested that the mass urban demonstrations in Iran played a comparable role to the peasant rebellions in the cases from her book. This hardly seems a ‘trivial’ modification, but Skocpol’s insight was confirmed by more recent work by Chenoweth and Stephan (2011: ch. 4) on civil resistance, including their case study of Iran. They did essentially what Bennett and Checkel suggested by generating and testing observable implications of mass, nonviolent protests in urban settings and finding that they yielded (more often than armed guerrilla insurgencies) successful revolutions.

A worrying development in political science has been the tendency to construct a hierarchy whereby the (inferior) role of positivist qualitative research is to support (superior) quantitative research. In this respect, the role of qualitative work is merely to unveil causal mechanisms, or worse yet, to suggest a plausible mechanism to link correlated variables. We contend, by contrast, that qualitative scholars have an important role to play that is in no way secondary or subordinate. Consider, for example, the common claim that experiments, particularly the ‘double-blind randomized control trial constitutes the “gold standard” of causal analysis’ (Blankshain and Stigler, 2020). For policy-oriented qualitative research, this claim is hardly self-evident. Impressively designed experimental studies have, for example, shed valuable light on US public opinion regarding such issues as the use of nuclear weapons (Press, Sagan, and Valentino, 2013; Sagan and Valentino, 2017; but see Carpenter and Montgomery, 2020) and humanitarian justifications for military force (Maxey, 2019; Maxey and Kreps, 2018). Yet, absent convincing evidence that US policymakers actually pay attention to the public when they carry out their decisions on the use of force, these elegant experiments lack policy relevance. A convincing process-tracing test of a hypothesized causal link between public attitudes and leaders’ decisions, on the other hand, makes those experimental findings meaningful (Maxey, *forthcoming*). By contrast, a process-tracing exercise that demonstrated the absence of influence of popular opinion on elite decision-making, or reverses causation through media ‘priming’ and elite cues, would suggest that experimenters’ time might be better spent elsewhere.

Unfortunately, at this time of great polarization, there seems to be an increasing division between qualitative and quantitative research, with relatively few interactions among them

both in terms of professional associations and potential outlets for publication. To illustrate, within the field of peace and conflict research, the yearly association of the Peace Science society gathers peace scholars of quantitative orientation, while the Conflict Research society has a solid representation of qualitative peace scholars. Notwithstanding considerable efforts to change the frame and open up professional associations and outlets to more diverse approaches, much remains to be done to bridge the gap. Our paper has suggested that qualitative researchers working in positivist, post-positivist, and interpretivist traditions have a lot to offer in terms of theoretical, conceptual, and empirical sophistication. Young scholars, as they pursue their professional development, should search for their own foundations both ontologically and epistemologically, find a comfortable place on the spectrum and be allowed to negotiate and renegotiate their own identity as researchers as they progress with their research.

Notwithstanding the lively methodological debate on qualitative causality, there is a relatively little practice-oriented reflection on the main differences among different views of causality and how those may shape the outlook of the research process. While drawbacks remain—such as the potential for missing confounders and omitted variables as well as overemphasizing negligible patterns—we advocate for a pragmatic view on causality in qualitative research: one that acknowledges complexity but strives to navigate the world in terms of patterns; one that strives to strike a balance between empirical richness and theoretical parsimony, a balance long recognized as necessary for policy-relevant research (Snyder, 1984: 85). We see that balance lying somewhere on the spectrum. Causality is rarely linear and often hard to establish. Yet striving in our research to capture loosely causal patterns gives us a direction of travel and pushes us to develop parsimonious and general theories. We oppose Manichean divisions between positivist, post-positivist and interpretivist scholars and have sought to demonstrate considerable overlap and blurring of boundaries between the various approaches. For instance, there seem ample opportunities for qualitative positivist-oriented researchers to conduct more collaborative work with both post-positivist and quantitative scholars and to think in more creative ways about mixed-methods (Bryman, 2007; Barnes and Weller, 2017; Koivu and Hinze, 2017).

Finally, we should not forget the key role that concepts and theories still play in the research endeavor. Sartori (1970) importantly advocated for the importance of mid-range concepts that should have enough attributes to capture key properties but also be light enough to be able to travel. Qualitative researchers employ valuable tools to craft and develop mid-range theories. And the same applies to theories: nuanced, well-crafted, and original theories are well suited to be studied by qualitative researchers. At an age of increased method-specific specialization, what really matters is to keep engaging in a joint conversation with ‘the other side’—whatever that might be—and strive to produce new, creative, parsimonious theories that can make our research journey meaningful and policy-relevant.

Funding. Chiara Ruffa gratefully acknowledges the financial support from the Royal Swedish Academy on History, Letters and Antiquities.

Acknowledgements. The authors thank Jeff Checkel and two anonymous reviewers for their excellent and constructive suggestions on earlier version of this manuscript and Chiara Tulp for her excellent research assistance.

References

- Autesserre S** (2010) *The Trouble with the Congo: Local Violence and the Failure of International Peacebuilding*. New York: Cambridge University Press.
- Barnes J and Weller N** (2017) Case studies and analytic transparency in causal-oriented mixed-methods research. *PS: Political Science & Politics* 50, 1019–1022. doi: 10.1017/S1049096517001202.
- Beck N** (2006) Is causal-process observation an oxymoron? *Political Analysis* 14, 347–352. doi: 10.1093/pan/mpj015.
- Belloni R** (2007) The trouble with humanitarianism. *Review of International Studies* 33, 451–474.
- Bennett A and Checkel JT**, eds. (2015) *Process Tracing: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press.

- Bennett A and Elman C** (2006) Complex causal relations and case study methods: the example of path dependence. *Political Analysis* 14, 250–267.
- Blankshain JD and Stigler AL** (2020) Applying method to madness: a user's guide to causal inference in policy analysis. *The Strategist* 3, 1–14.
- Blatter J and Haverland M** (2012) *Designing Case Studies: Explanatory Approaches in Small-N Research*. ECPR Research Methods. London: Palgrave Macmillan UK.
- Bode I and Karlsrud J** (2018) Implementation in Practice: the use of force to protect civilians in United Nations Peacekeeping. *European Journal of International Relations* 25(2), 458–485.
- Bove V, Rivera M and Ruffa C** (2020) Beyond coups: terrorism and military involvement in politics. *European Journal of International Relations* 26(1), 63–88. doi: 10.1177/1354066119866499.
- Bove V, Ruffa C and Ruggeri A** (2020) *Composing Peace. Mission Composition in UN Peacekeeping*. Oxford: Oxford University Press.
- Brady HE, Collier D and Seawright J** (2006) Toward a pluralistic vision of methodology. *Political Analysis* 14, 353–368.
- Bray Z** (2008) Ethnographic approaches. In Della Porta D and Keating M (eds), *Approaches and Methodologies in the Social Sciences: A Pluralist Perspective*. Cambridge: Cambridge University Press, pp. 296–315.
- Bryman A** (2007) Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research* 1, 8–8. doi: 10.1177/1558689806290531.
- Campbell SP** (2018) *Global Governance and Local Peace: Accountability and Performance in International Peacebuilding*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108290630>.
- Carpenter C**. The stopping power of norms: saturation bombing, civilian immunity, and U.S. attitudes toward the laws of war. *International Security* 45(2), 140–169.
- Checkel JT** (2008) Process tracing. In Klotz A and Prakash D (eds). *Qualitative Methods in International Relations: A Pluralist Guide, ECPR Research Methods*. New York: Palgrave Macmillan UK, pp. 114–127.
- Chenoweth E and Stephan MJ** (2011) *Why Civil Resistance Works The Strategic Logic of Nonviolent Conflict*. New York: Columbia University Press.
- Davenport C** (2013) Researching While Black: Why Conflict Research Needs More African Americans (Maybe). *Political Violence at a Glance*. <https://politicalviolenceataglance.org/2013/04/10/researching-while-black-why-conflict-research-needs-more-african-americans-maybe/>.
- Della Porta D and Keating M** (2008) How many approaches in the social sciences? An epistemological Introduction. In Della Porta D and Keating M (eds), *Approaches and Methodologies in the Social Sciences: A Pluralist Perspective*. Cambridge: Cambridge University Press, pp. 19–39.
- Evangelista M** (1986) Case Studies and Theories of the Arms Race. *Bulletin of Peace Proposals* 17(2), 197–206.
- Evangelista M** (1988a) *Innovation and the Arms Race: How the United States and the Soviet Union Develop New Weapons Technologies*. Ithaca, NY: Cornell University Press.
- Evangelista M** (1988b) "Military influence in Soviet politics: Red Militarism or National security consensus?". In Carlton D and Schaerf C (eds). *The Arms Race in the Era of Star Wars*. London: Macmillan.
- Evangelista M** (1991) Sources of moderation in soviet security policy. In Tetlock P, Jervis R, Husbands J, Stern P and Tilly C (eds). *Behavior, Society, and Nuclear War 2*. New York: Oxford University Press, pp. 315–316.
- Evangelista M** (1997) Domestic structure and international change. In Ikenberry GJ and Doyle M (eds). *New Thinking in International Relations Theory*. Boulder, CO: Westview.
- Evangelista M** (2015) Explaining the cold war's End: Process tracing All the Way down? In Bennett A and Checkel JT (eds). *Process Tracing: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press, pp. 329.
- Fujii LA** (2010) Shades of truth and lies: interpreting testimonies of War and violence. *Journal of Peace Research* 47(2), 231–241. doi: 10.1177/0022343309353097.
- Fujii LA** (2015) Five stories of accidental ethnography: turning unplanned moments in the field into data. *Qualitative Research* 15, 525–539. doi: 10.1177/1468794114548945.
- Fujii LA** (2018) *Interviewing in Social Science Research: A Relational Approach*. New York: Routledge.
- Geertz C** (1973) *The Interpretation of Cultures: Selected Essays*. New York: Basic Book.
- George AL and Bennett A** (2005) *Case Studies and Theory Development in the Social Sciences*. Cambridge, Massachusetts: London: MIT Press.
- George AL and McKeown T** (1985) Case studies and theories of organizational decision making. In Coulam R and Smith R (eds). *Advances in Information Processing in Organizations*, vol. 2. Greenwich, CT: JAI Press, pp. 21–58.
- Gerring J** (2008) The mechanismic worldview: thinking inside the Box. *British Journal of Political Science* 38, 161–179. doi: 10.1017/S0007123408000082.
- Gerring J** (2010) Causal mechanisms: yes, But.... *Comparative Political Studies* 43, 1499–1526. doi: 10.1177/0010414010376911
- Gerschenkron A** (1963) *Economic Backwardness in Historical Perspective*. Cambridge: Belknap Press of Harvard University Press.
- Goodin RE and Tilly C** (eds). (2006) *The Oxford Handbook of Contextual Political Analysis*. Oxford: Oxford University Press.

- Hedström P and Swedberg R** (eds). (1998) *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge, USA: Cambridge University Press, pp. 1–31.
- Hopf T** (1998) The promise of constructivism in international relations theory. *International Security* **23**, 171–200.
- Hopf T** (2013) Common-Sense constructivism and hegemony in world politics. *International Organization* **67**, 317–354. doi: 10.1017/S0020818313000040
- Jennings KM** (2019) Conditional protection? Sex, gender, and discourse in UN peacekeeping. *International Studies Quarterly* **63**, 30–42. doi: 10.1093/isq/sqy048.
- Kapizewski D, MacLean LM and Read BL** (2015) *Field Research in Political Science: Practices and Principles*, Cambridge: Cambridge University Press, Chapters 5–10. [Available in the Library]. Cambridge, MA: Cambridge University Press.
- Knopf JW** (1998) *Domestic Society and International Cooperation: The Impact of Protest on U.S. Arms Control Policy*. Cambridge: Cambridge University Press.
- Koivu KL and Hinze AM** (2017) Cases of convenience? The divergence of theory from practice in case selection in qualitative and mixed-methods research. *PS: Political Science & Politics* **50**, 1023–1027. doi: 10.1017/S1049096517001214.
- Krause P** (2017) *Rebel Power: Why National Movements Compete, Fight, and Win*. Ithaca, NY: Cornell University Press.
- Lebow RN** (2015) Counterfactuals and security studies. *Security Studies* **24**(3), 403–412.
- Levy JS** (2015) Counterfactuals, causal inference, and historical analysis. *Security Studies* **24**(3), 378–402.
- Mahoney J and Goertz G** (2006) A tale of Two cultures: contrasting quantitative and qualitative research. *Political Analysis* **14**, 227–249.
- Mahoney J, Khairunnisa M and Nguyen C** (2016) Causality and time in historical institutionalism. In Fioretos O, Falletti TG and Sheingate A (eds). *The Oxford Handbook of Historical Institutionalism*. New York: Oxford University Press, pp. 71–88.
- Maxey S** (2019) The power of humanitarian narratives: a domestic coalition theory of justifications for military action. *Political Research Quarterly* **73**(3), 680–695.
- Maxey S** (n.d. Forthcoming). The Humanitarian Discourse of Force: How U.S. Presidents Justify Military Interventions.
- Maxey S and Kreps S** (2018) Mechanisms of morality: sources of support for humanitarian interventions. *Journal of Conflict Resolution* **62**, 1814–1842.
- McKeown T** (1986) The limitations of ‘structural’ theories of commercial policy. *International Organization* **40**, 43–64.
- Meyer S, Bidgood S and Potter WC** (2020) Death Dust: The Little-Known Story of U.S. and Soviet Pursuit of Radiological Weapons. *International Security* **45**(2), 51–94.
- Moncrief S** (2017) Military socialization, disciplinary culture, and sexual violence in UN peacekeeping operations. *Journal of Peace Research* **54**, 715–730. doi: 10.1177/0022343317716784.
- Moore Jr, B** (1966) *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.
- Moravcsik A** (2014) Trust, but verify: the transparency revolution and qualitative international relations. *Security Studies* **23**, 663–688. doi: 10.1080/09636412.2014.970846.
- Nielsen RA** (2016) Case selection via matching. *Sociological Methods & Research* **45**, 569–597. doi: 10.1177/0049124114547054.
- Njølstad O** (1990) Learning from history? Case studies and the limits to theory-building. In Gleditsch NP and Njølstad O (eds). *Arms Races: Technological and Political Dynamics*. London: Sage, pp. 220–246.
- Pouliot V** (2007) Pacification without collective identification: russia and the transatlantic security community in the post-cold war era. *Journal of Peace Research* **44**, 605–622. doi: 10.1177/0022343307080858.
- Press DG, Sagan SD and Valentino BA** (2013) Atomic aversion: experimental evidence on taboos, traditions, and the non-use of nuclear weapons. *American Political Science Review* **107**, 188–206.
- Risse-Kappen T** (1986) Applying arms race theory to NATO’s nuclear weapons deployments: a reply to Matthew Evangelista. *Bulletin of Peace Proposals* **17**, 207–213.
- Roediger DR** (2007) *The Wages of Whiteness: Race and the Making of the American Working Class*. Rev. ed. Haymarket Series. London and New York: Verso.
- Rogers EM** (1983) *Diffusion of Innovations*, 3rd edn, New York: Free Press.
- Ruffa C** (2018) *Military Cultures in Peace and Stability Operations*. Philadelphia, PA: University of Pennsylvania Press.
- Ruffa C** (2020) Case selection and case analysis. In Curini L and Franzese R (eds). *The SAGE Handbook of Research Methods in Political Science and International Relations*. London: SAGE Publications, pp. 1133–1147.
- Sagan SD and Valentino BA** (2017) Revisiting hiroshima in Iran: what Americans really think about using nuclear weapons and killing noncombatants. *International Security* **42**, 41–79.
- Salter MB and Mutlu CE** (2013) *Research Methods in Critical Security Studies: An Introduction*. Abingdon and New York: Taylor & Francis.
- Sartori G** (1970) Concept misformation in comparative politics. *The American Political Science Review* **64**, 763–772.
- Seawright J and Gerring J** (2008) Case selection techniques in case study research: a menu of qualitative and quantitative options. *Political Research Quarterly* **61**, 294–308.
- Senghaas D** (1972) *Rüstung und Militarismus*. Frankfurt: Suhrkamp Verlag.

- Senghaas D** (1990) Arms race dynamics and arms control. In Gleditsch NP and Njølstad O (eds). *Arms Races: Technological and Political Dynamics*. London: Sage, pp. 441–447.
- Skocpol T** (1979) *States and Social Revolutions: A Comparative Analysis of France, Russia and China*. Cambridge: Cambridge University Press.
- Skocpol T** (1994) *Social Revolutions in the Modern World*. Cambridge and New York: Cambridge University Press.
- Snyder J** (1984) Richness, rigor, and relevance in the study of soviet foreign policy. *International Security* **9**, 89–108.
- Talmadge C** (2015) *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes*. Ithaca, NY: Cornell University Press.
- von Billerbeck SBK** (2017) *Whose Peace?: Local Ownership and United Nations Peacekeeping*. New York: Oxford University Press.