Rethinking causal explanation in interpretive international studies

Ludvig Norman
Stockholm University, Sweden

Abstract
This article develops a model for causal explanations amenable to interpretive International Relations (IR) research. A growing field of scholars has turned toward causal inquiry while stressing the importance of shared understandings, identities, and social practices for their explanations. This move has considerable potential to strengthen the contributions of interpretive approaches to IR. However, the article identifies shortcomings in the causal models on which this research is based which work to limit this potential. The article provides a detailed discussion of these limitations and offers an alternative model of causal explanations for interpretive IR. The proposed model builds on a clear differentiation between constitutive and causal analysis and supplies an explicit argument for how they can be combined to generate causal explanations. This paves the way for a more well-defined notion of causal explanation than has commonly been the case in interpretive IR. In doing so, it also offers a more coherent and detailed account of the points at which interpretive explanations intersect with more mainstream approaches and where they differ. Finally, the paper outlines an application of the model through a discussion on an updated form of interpretive process tracing (IPT).

Keywords
causal explanation, interpretive process tracing, constitutive explanation, methodology, interpretive international studies, counterfactual explanation

Introduction
This article offers a new framework for thinking about causal explanations from an interpretive perspective. A growing body of interpretive research in International Relations (IR) has worked to define what it means to produce causal explanations while remaining with
assumptions that emphasize the importance of norms, shared understandings, collective identities, and social practices for international politics. The interpretive turn toward causal analysis carries great promise in terms of exploiting the full potential of interpretive perspectives on central issues in IR. However, prevalent approaches to causal explanation in this field have also been beset with problems that have thus far curbed the potential of this turn. In particular, arguments for how to apply causal arguments in interpretive work have often been characterized by a lack of clarity regarding what the mobilization of causal language in these studies might imply. I argue that the absence of a well-defined causal theory at the base of these approaches comes with problematic consequences for both research design and for our abilities to evaluate causal arguments.

The article elaborates on two interconnected problems with existing approaches to causal reasoning in interpretive IR. I start of from the observation that interpretive work often relies on overly broad notions of cause that work to obscure rather than highlight the causal aspects of explanations. This is discussed here as a consequence of blurring the conceptual boundaries between two distinct types of explanation, constitutive and causal ones. Constitutive explanations capture structural conditions of particular social systems and speak to the latent dispositions and causal capacities of such systems. Causality captures the moving parts of explanations by tracking the relations between broadly defined events and direct our attention to processes and mechanisms. These different explanatory types thus perform different roles in making sense of the world. Defining what those roles might be and how they can be combined in a coherent way is an important part of chiselling out a more well-defined approach to causal explanations in interpretive work.

A second and related sticking point has to do with how interpretive conceptions of causation have evolved on the basis of, what I take to be, a partially misdirected critique of the standard regularity models of causation used in much mainstream and quantitative work in IR. My argument is not that interpretive scholars need to accept assumptions that are often associated to these models. Rather, I demonstrate how the basic counterfactual form on which many causal models are based is neither inherently tied to specific views on generalization, that interpretive scholars find problematic, nor do they necessitate a commitment to particular statistical applications. Thus, rather than to proceed through a more wholesale rejection and reinvention, a common strategy in existing interpretive work, my argument serves to identify which aspects of these conventional notions of causality that interpretive scholars can put to good use.

Building on these arguments, the article outlines a model for causal explanations amenable to interpretive international studies. The model relies on a distinct understanding of how cause and constitution can be nested in each other to explain specific outcomes in an interpretive framework. A central tenet of this argument is that while causal and constitutive explanations are different in type, they are similar in that they share a counterfactual structure.1 I elaborate on the implications of this perspective on cause and constitution and define their respective roles in producing causal explanations. I further discuss how the model can accommodate interpretive approaches in terms of their concern for contextual specificity and causal complexity. The potential usefulness of the framework is then illustrated by how it can be applied through an updated approach to interpretive process tracing (IPT).
The argument proceeds through a largely internal critique of causal inquiry in interpretive international studies in so far as it recognizes the important contributions of interpretive research to key issues in IR. Such contributions include, for instance, a more multifaceted view of the inner workings of international organizations, an understanding of the relations between socially produced identities and the foreign and security policies of states as well as a renewed understanding of international diplomacy as held together by social practices. The argument thus works within the assumption that there are considerable benefits associated with the application of interpretive forms of analysis in international studies.

While the article draws on philosophical discussions on causation, I do not pretend to provide any comprehensive analysis on the wide-ranging topic of causation in the social sciences nor do I engage at any depth in debates on ontology. The slightly more modest aim of the article is to work through a set of problems associated to the application of causal reasoning in interpretive IR and to elaborate a more useful and well-defined framework for causal explanations for this field. The aim is thus to contribute to causal theorization for scholars who recognize the importance of retaining an interpretive sensibility in the study of international politics and who recognize the importance of producing causal explanations in as clear and transparent way as possible.

### The causal turn in interpretive IR

Interpretive IR scholars have shown how studying the systems of meaning in which agents find themselves provides new ways to understand how states and international organizations behave, what characterizes their relations, and the varied forms of power they wield. Interpretive international studies emerged in contrast to long-time dominating rationalist and structuralist perspectives in IR where human and state behavior were largely understood as driven by narrowly defined instrumental calculation or overshadowed by the structurally defined material imperatives of the international system of states (cf. Lynch, 2014). Interpretive scholars have made important contributions by bringing a broader set of theories and analytical tools to the study of key issues in IR. This includes research on the inner life of international organizations (IOs), for instance by studying how bureaucratic rationales shape IOs’ interpretations of the world and contribute to the influence they project in the international system (Barnett and Finnemore, 2004; Nair, 2020). Interpretive studies of peace building operations have similarly highlighted how peace-making cultures among actors charged with organizing operations contribute to their success or failure (Autesserre, 2010). By approaching international diplomacy through the study of social practices, interpretive scholars have helped redefine what drives the agents that engage in it and have identified relationally defined resources that shape diplomatic conduct (Adler-Nissen and Tsinovoi, 2019; McCourt, 2016; Pouliot, 2016). Others have shown how deeply held collective identities condition states’ engagements in military interventions (Finnemore, 2003; Hansen, 2006; Hopf, 2005). A broad field of scholars have also provided insights to how foreign and security policies are intimately bound up with such identities (Campbell, 1998; Fierke, 2015; Guzzini, 2013; Tannenwald, 2007; Weldes, 1999). Interpretive international studies have thus demonstrated, in various ways, how the careful study of intersubjective meanings,
collective identities, social practices, and institutionalized shared understandings helps address core questions in IR.

This is a research field that has conventionally relied on observational studies built around the researcher’s deep immersion in particular social contexts, for instance in the secretariats of international organizations or in the context of ongoing peacekeeping operations. Scholars have relied on explanations that focus on how social systems condition agents to speak and act in particular ways, what we may refer to as constitutive explanations (Bevir and Blakely, 2018; Wendt, 1998; cf. McCourt, 2016). Constitutive explanations can take two principal forms. They may explain the dispositions of a social system as a whole with reference to its parts and how those parts are organized (Cummins, 2000; Ylikoski, 2013). For instance, they might focus on the collection of taken for granted practices that help explain the character of particular aspects of the international system as a whole (cf. Pouliot, 2016). They can also run in the opposite direction, explaining dispositions of individual or collective agents with reference to the social system in which those agents are situated (Wendt, 1998, 1999; cf. Hopf, 2005).³ Importantly, constitutive explanations do not rely on references to processes, triggers, or mechanisms. It is a form of explanation that identifies the relation between an explanans, that which explains, and an explanandum, that which is explained, where the two elements of the explanation are synchronous, in the sense that they are not temporally separate and cannot be regarded as separate existences. Such explanations thus have a different structure from a common understanding of causal claims. The understanding of causality, which also underlies the argument of this article, implies the spatial and temporal separateness of explanans and explanandum and understands causal explanations as mapping relations between broadly defined events (Chernoff, 2005: 118; Hausman, 2005; Schaffer, 2005; Ylikoski, 2013).

Constitutive explanations remain an important staple in interpretive work in IR. However, a growing literature has sought to find ways to add a more clearly defined causal element to these studies while remaining with key assumptions of an interpretive research tradition (Banta, 2012; Guzzini, 2013, 2017; Jackson, 2006, 2017; Klotz and Lynch, 2007; Lebow, 2009, 2014, 2015; Lynch, 2014; O’Mahoney, 2015; Pouliot, 2014, 2016; Schwartz-Shea and Yanow, 2013; Tannenwald, 2007, 2015; Zehfuss, 2002).⁴ The ambition to move beyond constitutive explanations and to explore what an interpretive approach to causal explanations might consist in is laudable, and one which I share. In particular, such a move deals head on with the problem that while interpretive approaches are apt at capturing habitual action and the reproduction of norms, identities, and social practices, they have limited abilities to account for change (Hopf, 2018; Stappert, 2020). Many interpretive scholars are, after all, also interested in why things happen, making the move toward causal inquiry a crucial one.

However, in exploring what causal explanations might mean in an interpretive framework, existing research have ventured in directions that, I argue, limit the potential of this turn. First, discussions have often worked to blur rather than to clarify the boundaries between constitutive and causal explanations. I elaborate on the problems that this generates and suggest an alternative perspective that identifies why it is important to keep them analytically separate as well as how they can be combined to generate causal explanations. Second, interpretive scholars have tended to use covering law and regularity
Theories of causation as their main foil when developing their causal models. I demonstrate how aspects of this critique is misplaced and based on partly incorrect assumptions, held by interpretive and noninterpretive scholars alike. In particular, these assumptions concern the view that there is a necessary relation between, on the one hand, particular and widely used conceptions of cause and, on the other, a commitment to particular views on contextual specificity and generalization. I challenge these assumptions and show how elements of a conventional and widely embraced statement of counterfactual causality in IR and political research more generally can be employed in interpretive studies. These two moves help pave the way for the development of a more clearly defined model for causal inquiry in interpretive international studies.

**Blurring causal and constitutive explanations**

Interpretive scholars tend to argue that a redefinition of causation is necessary to ensure that causal explanations remain sensitive to the contextually specific meanings that imbue how actors perceive the world and understand themselves in it. This is, in principle, a reasonable position. To proceed without such an interpretive sensitivity would indeed be self-defeating from an interpretive perspective. However, the specifics of how causality has been incorporated in interpretive work has tended to blur the lines between two distinct types of explanations: constitutive explanations and causal ones. The implication of this conceptual move is an understanding of cause where the differentiation between dispositions and the events and processes that produce particular outcomes dissolves. Klotz and Lynch argue in this vein that an interpretive account of causality should be understood as the “effects of structure,” how people reproduce dominant practices as a result of specific social, historical and spatial conditions (Klotz and Lynch, 2007: 15). Smith (2004) makes this move in introducing the notion of “generative causality.” It occurs “when an actor is provided with a capacity to do certain things but not others—making it more probable, but not inevitable, that the actor will do those things, and less likely, though not impossible, that the actor will do something else entirely” (Smith, 2004: 46–47). Lebow’s notion of “inefficient causation” is based on a similar perspective, where constitutive and causal reasoning overlap, where there is no strict binary between the two (Lebow, 2014: 60). Instead, Lebow argues that the causal import of intersubjective frameworks emerges as “once entrenched [they make] other courses of action increasingly less appealing and even unacceptable” (Ibid., 78; cf. Lebow, 2009). Finnemore asserts that socially produced beliefs “constitute certain behavioral possibilities and, in that sense, causes them” (Finnemore, 2003: 15 italics added; cf. Dessler and Owen, 2005). These assumptions also resound with more recent discussions on the causal aspects of interpretive forms of process analysis in dispositional terms (Guzzini, 2017).

Contrary to what some analysts have argued, I believe there are good reasons to retain a category for constitutive explanations (cf. Banta, 2012; Dessler and Owen, 2005; Grynaviski, 2013; Kurki, 2008). However, it needs to be held separate from the category of cause. Many research problems that interpretivists are interested in are associated to the constitutive explanatory type. It speaks to dispositions and causal capacities that are inherent to a specific social system, a set of practices, a culture, collective identities, or a
collection of taken for granted norms. As will be argued below, constitutive explanations play a crucial role in the model for causal explanations developed in this article. They are of essential importance for interpretive studies but they are not of the causal type.

A principal problem with an unclear differentiation between cause and constitution is that dispositions and capacities are made synonymous with process through which one or several events produce a particular outcome. Notions like dispositions and cognitive frames or capacities do not help us zero in on the causal aspects of an explanation (cf. Lebow, 2014: 60–61). A similar point can be made in relation to Guzzini’s discussion on power as cause where the language of dispositions and latent abilities is employed to underpin a reconceptualization of what cause may mean from an interpretivist position. The argument is exemplified by the relation between the capability to read and reading (Guzzini, 2017). A constitutive explanation of reading can indeed refer to someone’s capability to read; however, it is not enough for a causal explanation of why someone is reading in some specific instance to cite the person’s capability or disposition to do so.

Similarly, to supply a causal explanation of state action, for instance, a particular state’s decision to engage in a military intervention, it is not enough to refer to the identities that predisposed its leaders to act in a particular way (cf. Hopf, 2005). Such an analysis can again be informative, but a causal explanation of why a state uses military force in a particular situation requires that we also account for the antecedent chain of events that triggered that activity at that particular time. Such events will, from an interpretive perspective, only be discernible through an interpretation of the social context in which the activity occurs. That is, events might only be understandable as causally effective if preceded by a constitutive explanation through which widely held identities among political elites are mapped. It might also depend on examining the meanings that agents assign to their relations to others, specific states or organizations, or the meanings that agents assign to the use of force in the international sphere in general. Nonetheless, accounting for such dispositions for state action is different from accounting for the actual process that activate latent dispositions and produce specific actions in the particular case.

Why would this matter? Most importantly, if no analytical differentiation is made between broader systems of meanings in which agents are embedded, on the one hand, and the series of events that are interpreted and acted upon by those agents, on the other, we lose sight of the mechanisms through which dispositions are translated to action. We also make it more difficult to evaluate if our causal accounts of these events are reasonable or not.

In this reading, it is indeed problematic to take the rout chosen by interpretive perspectives that conceptualize social systems and intersubjective meanings as having causal effects. I thus differ from analysts who have suggested that discursive analyses which identify conditions of possibility or enabling contexts in social structures are “just as causal” as accounts that invoke mechanisms or the relations between events (Kurki and Suganami, 2012: 413; cf. Banta, 2012). Once in place, a set practices or dominating discourses, much like cognitive frames of reference or general dispositions, do not trigger processes. Constitutive and causal explanations imply different types of methods for substantiating their respective claims. They also imply different criteria according to which such claims can be evaluated. Finally, when combined they
require explicit arguments regarding how the different types of explanations are nested in one another.

Reexamining the specter of covering laws

The development of distinctly interpretive notions of causation has largely proceeded in contra-distinction to causal theories that understand cause in terms of general probabilistic relationships, regularities, or related to covering laws. If social and political outcomes are shaped by contextual social meanings and enacted through contextually specific practices, other notions of causation than those found in Hume and Hempel, it is argued, need to come into play (Guzzini, 2013, 2017; Jackson, 2006, 2017; Klotz and Lynch, 2007; Lebow, 2014, 2015; Pouliot, 2014, 2016). Klotz and Lynch tie their concept of “conditional causality” explicitly to the goal of “context dependent generalizations” (Klotz and Lynch, 2007). Jackson relies on the Weberian idea “adequate causality” for similar reasons, as an alternative to a nomothetic conception of causation, with its focus on testing the strength of correlations across a large number of cases (Jackson, 2006 cf. Jackson, 2017). Guzzini (2013) and Pouliot (2014) discuss causality by emphasizing how establishing causal relations from an interpretive perspective needs to proceed through the interpretation of “local” social contexts. The concept of inefficient causation is also devised as an explicit counter position to the law-like statements of regularity theories (Lebow, 2014: 69). I agree that the study of how social institutions shape particular outcomes is, from an interpretive perspective, necessarily one which will emphasize contextual specificity over attempts to identify regularities that are as general as possible. The question is to what extent this necessitates a reconceptualization of what causation is. I argue that rather than to develop new and sometimes idiosyncratic concepts of causation, interpretive scholars can employ similar understandings of causation to those that underlie commonly used counterfactual causal models in mainstream IR.

Contrary to what both interpretive and noninterpretive scholars tend to argue, the foundations of such models are not inherently tied to any specific assumptions regarding the possibility of generalizing findings across social contexts. Such assumptions are more accurately tied to the statistical application of those models and to an epistemological position that social scientific research should, to the extent possible, approximate experimental scientific ideals. Interpretive scholars are warranted to question such ideals when applied to IR and the social sciences more broadly. However, in doing so they need not close the door on understandings of cause that have no necessary connections to those ideals. A more fruitful way to proceed is to differentiate fundamental statements of causation from particular experimental and statistical applications. This will help scholars distinguish between aspects of causal models that emerge in tension with interpretive assumptions and which aspects could help clarify what causal explanations from an interpretive perspective may consist in.

To illustrate this point, it is useful to consider a well-known and influential account theorizing causal inference from a statistical point of view that has been, and remains, widely used in mainstream IR and elsewhere in the social sciences: the Rubin causal model (RCM) (Holland, 1986). The model builds on a simple statement of counterfactual causality which, in contrast to the model as a whole, could be used to underpin an
interpretive model for causal explanations. The RCM, however, uses an experimental template where any attempt to measure the effect of a particular cause requires gauging the contrast between two observable causes, the treatment $t$ and some other cause, the control, $c$ on a unit $u$. In this model, the dependent variable $Y$ measures the response to the cause, and comes in two variants $Y_t$ and $Y_c$ where $Y_t$ means the value of the response (dependent) variable if exposed to the treatment, $t$, and $Y_c$ the value of the response variable when exposed to the control $c$. $Y_t(u)$ is the value of the response that we get if the unit $u$ were exposed to the treatment, $t$ and $Y_c(u)$ is the value of the response if the same unit were exposed to the other cause, the control, $c$ (ibid. 946). This means that in its most basic formulation the model is concerned with measuring the effect of $t$ relative to $c$ on a specific unit, $u$ measured by $Y$. This is expressed by Holland through the basic formula

$$Y_t(u) - Y_c(u)$$

The causal effect is the difference between $Y_t(u)$ and $Y_c(u)$. It is thus at its core a counterfactual causal statement. Getting at the difference between the two elements of this statement is, however, frustrated by the problem of not being able to observe the treatment $t$ and the non-treatment, the control $c$, in the same unit. According to the RCM, it is not possible to establish, for instance, whether a hypothesized determinant of armed conflict, peaceful interstate relations or IO behavior have a particular effect if we are not able to compare cases where the cause is present to cases where that particular condition is absent.

It is important to underline that this problem, commonly referred to as the “fundamental problem of causal inference,” is a consequence of assumptions related to the experimental ideal and does not flow naturally from the basic causal statement itself (cf. Johnson, 2006). For those who hold these assumptions, however, the problem is addressed by moving beyond the confines of the particular unit and measure the effect of the treatment across a larger sample of a well-defined population of units $U$ where each unit is potentially exposable to either $t$ or $c$. The statistical solution to the problem, hence, “replaces the impossible-to-observe causal effect of $t$ on a specific unit with the possible-to-estimate average causal effect of $t$ over a population of units” (Holland, 1986: 947). So, the objective of the RCM as a whole, but not of the basic causal formula, is to understand the average causal effect, $T$, of a given cause (or treatment) on a population of units $U$. This is then calculated as the expected value of the difference between $Y_t(u)$ and $Y_c(u)$ over all the $u$’s in a population of units $U$ (or a representative sample thereof which is more commonly the case) expressed as

$$T = E(Y_t) - E(Y_c)$$

This motivates the reformulation of a general causal statement such as the difference that $X$ makes for $Y$ to a focus on average causal effects. The central point here is that the general causal statement at the basis of this model, the difference that $t$ makes for $u$ is quite amenable to an interpretive model for providing singular causal explanations, while the statistical application of this statement is less so.

The interpretive criticism that cause is defined in these causal models in terms of regularities is thus only partially correct. A more accurate criticism from an interpretive
perspective would be that the statement \( Y_t(u) - Y_c(u) \) tends to be misinterpreted as being *intrinsically* linked to an experimental template, which then requires the calculation of the average effect of the cause over a representative sample of a population. It is this move that brings regularity theories to the fore and which underpins the questionable position that casual inquiry in the social sciences is unified by a common probabilistic notion of causality (Gerring, 2005). Interpretive as well as non-interpretive scholars thus tend to incorrectly interpret the model as a whole, and its association to generalization, as a necessary outgrowth of the basic formula of causation on which it builds.

The modelling of causation on the experimental template is thus tied to a set of specific assumptions that interpretive scholars are not bound to accept. If those assumptions are relaxed, the statement \( Y_t(u) - Y_c(u) \) could be read simply as a statement regarding the difference that the cause makes for the outcome. Supplying theoretically informed accounts of such relations conforms to widely held intuitions regarding causal explanations and does not by any necessity actualize “the fundamental problem of causal inference,” nor its statistical solution.6

The statistical application of the general statement of causation at the base of the model effectively makes questions of generalization across a specific population of units intrinsic to the question of causal dependencies, as it defines the strength of such dependencies in terms of quantified averages. The calculation of average causal effect is the same as calculating to what degree the relation between cause and effect holds across a large number of cases, that is, to what extent it holds more generally. By implication, this also necessitates a range of assumptions to hold true, such as a high degree of unit homogeneity as well as assumptions regarding the independence of causal effects across units. It does not, however, as Holland also demonstrates, conform to the Humean notion of constant conjunction (Holland, 1986: 950). Relatedly, contrary to what interpretive scholars tend to argue, it is also doubtful if the assumption of covering laws are important, or even relevant for the model.

This further supports the notion that discussions regarding the possibilities of generalizing findings can be held separate from discussions on causation. The implication for interpretive scholars is that it clears the ground for a conceptualization of causation unrestricted by various additional assumptions that do not, strictly speaking, help clarify what a causal explanation consists in. An interpretive position does indeed provide good reasons to move beyond the understanding of causation in terms of the calculation of average causal effect as the exclusive notion of what causation can be and how it can be gauged. Ontological assumptions underpinning much interpretive work provide good reasons to question the possibility of universal and context independent social scientific knowledge. However, it is less clear to what extent causation needs to be reinvented and what the appropriate reference points for such a reinvention should be. Treating the question of causation as separate from its statistical application as well as distinct from discussions on generalization prevents us from throwing out the baby with the bathwater.

**A causal model for interpretive international studies**

The analytical moves argued for so far include first a clear differentiation between constitutive and causal relations and a notion of cause that conforms to a broadly defined
notion of causal explanations in counterfactual terms. I build on these notions and develop a model for causal explanations of individual events for interpretive research based on existing theories of “singular causal explanations” (Woodward, 1984, 2003; Hausman, 2005; Schaffer, 2005; Ylikoski, 2013 cf. Lewis, 1986). The model offers a way to clarify the component parts of interpretive causal analysis and helps define the respective roles of cause and constitution in generating causal explanations. In particular, the model builds on the notion that causal explanations from an interpretive perspective are necessarily nested in constitutive ones. The discussion helps demonstrate how the model can accommodate concerns regarding interpretive sensitivity to contextually defined social systems and their causal complexities.

**Contrastive explanations**

At the heart of the model discussed here lies a contrastive understanding of both causation and constitution where both the *explanans*, that which explains, and the *explanandum*, that which is explained, are defined in relation to some well-defined alternative. Thus, both constitutive and causal explanations are understood in counterfactual terms. Constitutive explanations from this perspective have the same basic structure as causal ones but differ in the type of relations that are mapped between *explanans* and *explanandum* (Ylikoski, 2013). Singular causal explanations capture relations between events (broadly defined) and constitutive explanations capture relations between the characteristics of some social whole and some aspect or part of that whole.

The fundamental notion here is that to provide an explanation in general, constitutive or causal, is to demonstrate that if the conditions cited in the explanans would change, so would the explanandum. The contrastive notion of causal and constitutive explanations means that for both these explanatory types, the relations between the *explanans* and the *explanandum* are elucidated through contrasts with some specified alternative. In Woodward’s words, an explanation of why something happened, what he refers to as singular causal explanations, is always also the answer to the question “what-if-things-would-have-been-different?” (Woodward, 1984). This position is tied to a well-established position among an epistemologically diverse set of perspectives that view the specification of alternative realities as intimately associated to what it means to supply an explanation (Fearon, 1991; Grynaviski, 2013; Lebow, 2015; Lewis, 1986; Schaffer, 2005). Such causal statements rely on contrasts being specified on both sides of the explanation so that

\[
c \text{ rather than } c^* \text{ causes } e \text{ rather than } e^*
\]

This formalized way of thinking about explanations might seem far removed from how interpretive scholars generally understand their analyses and may appear overly reductive and simplistic. Interpretive IR scholars have, as demonstrated above, been suspicious of similar causal statements, associating them to neopositivist perspectives. However, the contrastive notion of causal explanations does not necessitate a commitment to probabilistic or regularity theories of causal explanations nor a commitment to making broad generalizing claims. If understood in terms of a singular causal explanation, it does not
require analysts to look for explanatory leverage beyond the confines of the specific context studied:

Singular causal explanations wear the source of their explanatory efficacy on their face—they explain not because they tacitly invoke a “hidden” law or statement of sufficient (or probabilifying) conditions, but because they identify conditions such that changes in these conditions would make a difference for whether the explanandum-phenomenon or some specified alternatives would ensue. (Woodward, 2003: 216–217)

The extent to which an explanation is relevant and informative is thus not bound up with the identification of general statements of neither a deterministic or probabilistic kind. The question of how well an explanation performs in a particular context is different from the question regarding the degree to which they can be generalized to other contexts. It just asks us to specify, for each relation that we do seek to establish, the “contrast focus” that make that particular relation intelligible.

It is important to note that many interpretive scholars, even when staying away from causal explanations, employ contrasts as analytical tools. That is, the use of contrasts is not foreign to interpretive work. Indeed, many constitutive explanations in this field exhibit a contrastive structure, albeit implicitly. They explain by identifying the aspects of a particular social context which make a difference for the explanandum. This is, to take an example, true for Smith’s constitutive argument which, it can be argued, resonates with much interpretive scholarship. He states that social wholes, in his case widely embraced notions of nationhood, generates the “capacity to do certain things but not others” (Smith, 2004: 46–47). This is a constitutive argument, but also clearly a contrastive one: a particular notion of nationhood, \( n \), rather than some other notion of nationhood, \( n^* \) is associated to dispositions to engage in actions \( a \) rather than some alternative actions \( a^* \). A similar structure applies to Autesserre’s constitutive analysis of failed peacebuilding in the Congo (Autesserre, 2010). She persuasively demonstrates how a particular culture of peacebuilding, dominant among NGOs and UN agencies deployed to intervene in the conflict, was expressed in strategies that ultimately contributed to the intervention’s failure. In particular, Autesserre argues, a culture that conditioned particular interpretations of what was unfolding on the ground made these outside actors blind to the central dynamics that motored the continuation of widespread violence in the country. Had they instead been embedded in a culture that allowed them to be more attentive to the local knowledge of these conflict dynamics, chances are that their strategies would have been shaped differently. These arguments thus imply that if the socially produced dispositions of agents had been different, if the notion of nationhood or the peacebuilding culture among NGOs and UN agencies had been different, then the dispositions for acting in one way rather than another would also have been different. That is, the causal capacities of the social system would have been different. This structure conforms to standard arguments made by interpretive scholars and undergirds many of the contributions made in interpretive IR. They derive their explanatory power partly from the way in which they invoke contrasts. It is in supplying these contrasts that an explanation becomes intelligible as such. Explicit references to contrasts push researchers to be explicit about what a rea-
sonable alternative might be to what is observed in a particular context, and what difference this makes for a particular outcome.

Causal explanations and social context

An important rationale among interpretive scholars for developing their own notions of causality is to enable the study of local and highly contextualized social processes. As discussed above, this does not require us to rethink causation. It does, however, require us to work out the more specific relations between constitutive and causal explanations. I suggest that we can do this by seeing constitutive explanations as establishing the relevant components of a particular causal field on which causal relations play out (cf. Mackie, 1965). Constitutive explanations help make sense of how events are interpreted and acted upon by socially situated agents and why causal processes are triggered and unfold in particular ways.

In discussions on causation, causal fields are often exemplified in terms of physical background conditions which need to be present for a specific causal process to produce itself. For interpretive approaches, mapping such conditions or causal capacities will instead focus on the social institutions in which agents find themselves, whether understood as dominating cultures, social practices, identities, norms, or systems of meaning. However, a well-worn example from philosophical discussions on causation serves as a useful analogy here. The example proceeds by unpacking the causal claim that the occurrence of a fire in a particular house is explained by the occurrence of a short circuit (Mackie, 1965; Schaffer, 2005; Woodward, 1984, 2003).

The example helps demonstrate how a seemingly simple causal explanation is dependent on a complex set of conditions, constituting the causal properties of the house, including for instance, the presence of flammable material, the presence of a ventilation channel supplying oxygen, and the absence of a functioning sprinkler system. According to Mackie, the causal field is the house in its entirety, thought of in terms of the configuration of these conditions, a set of causal capacities, which in the event of a short-circuit will help explain the outbreak of fire. Establishing the features of the causal field is analogous to answering questions regarding constitutive relations. This means identifying latent dispositions and theorizing the causal relations that would play out given a particular set of events in a certain context (Ylikoski, 2013).

How does this reasoning apply to the social world? For social scientists, and for interpretive scholars in particular, causal explanations of social action require identification of the causal capacities of the social system in which agents act. Social wholes, such as organizations, systems of rules, norms, and taken-for-granted social practices, can be thought of as assemblages of causal capacities that will shape the effects that particular events will produce. The mapping of such capacities requires constitutive analyses of the social institutions and relations in which agents are embedded. This implies the use of interpretive methods through which researchers immerse themselves in a social setting to gain an “insider” perspective and develop a sense of the weight and significance that actors assign to particular events. For instance, the deeply institutionalized norm that identified the use of nuclear weapons as increasingly illegitimate during the post-war period, what Tannenwald (2007) has referred to as “the nuclear taboo,” helped constitute
what can be understood as a causal field upon which a set of more specific causal processes played out. When decision-makers were faced with problem situations and events where the use of nuclear weapons could have been expected, they still refrained from using them, and instead chose other courses of actions, due to, Tannenwald argues, the taboo. To exemplify in somewhat closer detail, we can turn to Barnett and Finnemore’s (2004) study of the organizational cultures that informed decisions taken by the UN’s department of peacekeeping operations (DPKO) in relation to the Rwandan genocide. Their arguments are not structured in line with any specified casual model and as such, similar to Tannewald, they remain fuzzy about the specific causal components of their account. However, their analysis proceeds implicitly through a logic similar to the model developed here. Their account supplies a constitutive analysis of an organizational culture within this IO and demonstrates how we can make sense of certain decisions in light of this culture. A key example from this account includes decisions by the UN Secretariat to withhold information from the UN Security Council about the widespread and highly organized mass killings that were reported by personnel on the ground in Rwanda and to order its personnel to not interfere (ibid., 147). Like the example with the house where fire is caused by a short-circuit, the component parts of a specific organizational culture establish the causal field that helps explain how particular events will instantiate themselves in the lives of the agents that are studied and how those events will produce particular outcomes. A specific set of shared understandings and collective identities constituting the social fields in which agents find themselves will condition them to perceive, interpret, and act on particular events in different ways. Offering a constitutive explanation, in terms of defining the causal field, supplies the conditions for offering a causal one.

Pairing cause and constitution in this way helps address a perennial problem for causal analyses, namely the inherent indeterminacy in any attempt to clearly separate causes from background conditions. For the example of the short-circuit and the outbreak of fire as well as for causal explanations in the social sciences, there is a fundamental indeterminacy of what we take to be a cause and what forms part of the causal field of background conditions. Singular causal explanations do not supply solid reasons why the short-circuit rather than any other of the factors necessary for the fire (presence of oxygen, flammable materials, absence of functioning sprinklers) should be cited as the primary cause. We can imagine a range of different scenarios where we identify the presence or absence of factors as that which made the difference to the outbreak of the fire. Sticking with the notion of causal explanations as tracing the relations between (broadly defined) spatiotemporally separate events helps narrow down the candidates, especially in a stylized setting as that of the short-circuit and the fire. Thinking about causal relations as a contextually specific relation between events helps differentiate causes from background conditions and helps differentiate constitutive explanations from causal ones.

Nevertheless, in a messy social world, indeterminacy in identifying causes will invariably linger. The choice of identifying some aspect of a particular situation as a cause and some other aspect as part of the background causal field is not easily resolved. Selecting the cause stems, as Schaffer argues, “from background expectations which
generate different causal inquiries” (Schaffer, 2016: 52). Building on this notion, the indeterminacy may be at least partially addressed by identifying “abnormal” events in the context of our expectations regarding the “normal functioning” of some specific aspect of the social world (Hart and Honoré, 1985: 34; Schaffer, 2005). Returning again to a stylized example like the eruption of fire, under normal circumstances, citing the presence of oxygen as a cause for a fire would seem rather banal and we would perhaps rather, cite the short-circuit, as the short-circuit departs from what we would expect from the normal functioning of the house. However, if we posit an environment, for instance a laboratory where experiments are conducted in an environment free from oxygen, and a fire erupts in this environment, we might be inclined to cite the introduction of oxygen as a cause rather than a condition, in the same way we did to the short-circuit in the prior example. The deviance from the normal functioning of the house or the laboratory informs us of what aspects of the situation it is reasonable to cite as a cause. But how can we then translate this notion of normal functioning to the social world?

One way to do this is to study how social conventions in a particular setting are structured, and how such conventions contain information regarding understandings among the agents we study of what constitutes normality and abnormality (cf. Hart and Honoré, 1985: 37; Chernoff, 2005). Interpretive methods are particularly well-suited to establish the shape and form of such conventions as these methods have been specifically developed to make sense of how research subjects view the world. Barnett and Finnemore’s account of IO behavior is quite explicit about this aspect. Indeed, they highlight the departure of the IOs they study from their expected normal functioning by explicitly referring to these instances as “pathologies.” Those pathologies are identified by the researchers through a combination of an interpretation of the agents’ interpretations of themselves and their organization while also exposing those understandings to broader, external, expectations that we might have on the conduct of such organizations. Knowing about the shared understandings and the meanings that agents assign to specific events supply vital information for causally explaining their actions.

The interpretive assessment of “abnormality” as a tool to pin-point cause also helps address difficult questions regarding how to understand ‘events’. It allows us to treat particular actions as causally effective events. An interpretive assessment of abnormality also allows us to appreciate how absences can play a role for causal explanations, especially in interpretive studies. The causal effect of absences, of something missing or not being present, might seem counterintuitive in a theory of causal explanations that rely on the notions of causes as events. However, the contextual analysis of conventions help demonstrate how certain absences can be treated as, for example, instances of negligence and failure to act in line with social expectations, and thus as events rather than “nothings” (Hart and Honoré, 1985: 38; Schaffer, 2005; cf. Moore, 2009). In discussions on causation in the law, concepts like “neglect” and “failure to act” point in this direction. These are concepts associated to institutionalized expectations of appropriate conduct in particular social situations. Absences as causes come into clearer view through the explicit mobilization of contrasts at the basis of explanations. The use of contrasts demands an estimation of the “normal functioning” of social situations and an estimation of to what extent agents’ actions or non-actions were perceived as divergences from that state of normality.
Social relationships place demands on international actors such as representatives of states, IOs and INGOs to engage in particular practices. Absence, in terms of not doing things, abstaining, not being present, non-participation are all events that, given the causal field where they occur, can have wide-ranging consequences. We may, for instance, consider the possible impact of an influential actor not partaking at an international summit or in high-level peace talks. Not turning up, not voting, not speaking, not funding all have causal significance in that they can trigger social processes. Crucially however, estimating the causal significance of such events require our prior understanding of the practices that constitute social relationships, whether between individual agents, between states or some other corporate actors like international organizations. Such relations can for instance be characterized by mutual distrust, perceptions of cultural superiority or inferiority, and perceptions of differing social status, which will have implications for how events are perceived and which reactions they trigger. By understanding the socially produced causal field in which events acquire meanings among specific sets of agents, the “as-if” objectivity of such events becomes discernible. This makes it possible to study how events move processes that produce particular outcomes and thus serves as the basis for giving a causal explanation from an interpretive perspective.

Causal complexity

A possible criticism of the framework for causal explanations argued for here might be that it appears too simplistic. It reduces complex and multifaceted social processes to the causal relations between a small number of factors. Indeed, Lebow identifies this as one of the more important reasons for turning from causal analyses that relates a specific X to a specific Y, to inefficient causation which seeks to capture layered and intricate causal webs that provide an understanding of particular outcomes (Lebow, 2014: 69; see also Guzzini, 2017). However, from the perspective argued for here, such a criticism would only be partly accurate. Like all attempts at making some aspect of the social world intelligible, interpretive explanations, however complex, are also necessarily simplifications. Not only does explanation imply a simplifying exercise, but explanations draw their force from simplification. For instance, to argue, as Barnett and Finnemore (2004) do, that a culture of bureaucratic rationality explains certain types of failures of international organizations also relies on considerable simplification. Among a range of possible things that might help explain such organizational pathologies, the internalization of a specific bureaucratic culture among its officials is identified as the most salient one.

To decide on how complex an explanation needs to be is strictly speaking an analytical choice that has to do with how we apply causal models, including, it should be said, statistical ones. There are no set criteria for identifying the ideal degree of complexity or simplicity for an explanation to qualify as such. By consequence, there are no reasons why we should build our basic notion of causality on a predetermined notion of how complex an explanation needs to be to be useful. Such choices are invariably theory driven and pragmatic as well as, to some degree, informed by the specific aspect of the social world for which an explanation is sought (Humphreys, 2017).
It is thus not required by an explanation to cite the exhaustive list of conditions that were necessary to produce a particular outcome. What explaining does require is “accounting for the contrast between the actual situation and one distinctive alternative” (Woodward, 1984: 246; see also Lewis, 1986: 217). The same applies to constitutive explanations. Constitutive explanations do not require us to supply a complete mapping of a particular social system (as if this would be possible). Rather, such explanations acquire their explanatory force by identifying the specific aspect of the system which makes a difference for dispositions of the agents we are interested in. For either type of explanation, there is no a priori set level of complexity that analysts must conform to. The model discussed here only creates a minimal baseline for constructing each component part of a configuration of causal and constitutive explanations.

**Interpretive process tracing: an application**

In this last section, I elaborate on one possible application of the causal model to illustrate its potential benefits for interpretive IR. I use the framework to further specify the component parts of interpretive process tracing (IPT). As I have argued elsewhere, IPT emphasizes interpretation as a means to enter the conceptual worlds of those we study but also works to account for how changes in intersubjective worlds play into our causal explanations of well-specified social and political outcomes (Norman, 2015, 2016 cf. Geertz, 1973: 24). IPT is thus closely related to the type of highly contextualized process studies proposed by Pouliot (2014) and Guzzini (2013, 2017) but adds to this a more clearly defined framework for how constitutive explanations contribute to causal explanations. This enables a sharper differentiation of dispositions of particular social systems from causal analyses that map relations between events. As such, it makes the processes studied visible and supports more careful attendance to the temporal aspects of causal process accounts. IPT borrows its emphasis on temporality, sequence, and self-reinforcing dynamics from historical institutionalism (HI) but departs from the often-strict materialism and the largely rationalistic view of human action underlying such accounts (Cappoccia and Kelemen, 2007; Pierson, 2004). Thus, while IPT is concerned with causal explanations, with events and the relations between events, it places causal explanations in interpretively generated constitutive explanations.

IPT thus differs from conventional approaches to process tracing, which in much recent work have moved further toward increasingly formalized Bayesian procedures for deductive hypothesis testing and for assigning and updating quantified probabilities for different steps in a process (Fairfield and Charman, 2017). Used in a pragmatic way, Bayesian assumptions can be employed informally to work through our theoretical expectations of the dynamics of a particular process (Beach and Brun Pedersen, 2013; George and Bennett, 2005). However, the formalization of such procedures, in terms of quantifying those expectations, works to radically minimize the type of abductively driven and context sensitive research designs that are necessary for retaining an interpretive sensibility.

Proponents of process tracing highlight its potential for including explorative and inductively generated evidence in the analysis of processes (Bennett and Checkel, 2014). The possibility to move back and forth between theory and insights developed throughout the
research process is indeed key for causal explanations from an interpretive perspective. However, there is a striking dearth in the process tracing literature on how to incorporate interpretive insights in causal process account. Hall’s (2013) call for process tracing methods that explicitly work to integrate interpretive methods have remained largely unanswered and Pouliot (2014) points to the neglect of interpretive perspectives in this literature. Most recently, Checkel (2020) has observed that despite the broad epistemological scope of process tracing, few attempts have been made to adapt it to an interpretive perspective (cf. Guzzini, 2013). The dominating trend has rather been to strengthen its deductive elements and the procedures for how to test *a priori* specified hypotheses (Schimmelfennig, 2014; Waldner, 2015). IPT speaks to this gap, instead making the interpretative exploration of social systems a key element of causal process analysis.

Similar to “mainstream” approaches to process tracing, the notion of social mechanisms plays an important role for IPT. To paraphrase Jon Elster, mechanisms help supply answers to the question “why did they do that?” and are as such at the core of supplying explanations (Elster, 1998 cf. O’Mahoney, 2015). However, in IPT the heavy lifting in terms of establishing causality is done by singular causal explanations combined with constitutive analysis as discussed above. Once specific causal patterns have been identified these patterns can be ordered and conceptualized as mechanisms.

Thus, mechanisms in IPT serve the dual functions of ordering causal relations in coherent causal story and crucially, tying the specifics of a particular process to broader social theoretical understandings of human conduct that can be made to “travel” to other, analogous social contexts. In that sense, mechanisms can be reformulated as abstractions, analytical tools that contribute to the theoretical toolbox that can be brought to bear on future studies. IPT thus treats mechanisms as a vehicle to order the set of causal relations identified in a particular case in an intelligible way and, second, as tools with which connections can be made between seemingly disparate settings and processes by paring down multifaceted and complex explanations to their bare bones. However, to be able to identify mechanisms we need to first establish the relations between events that form a particular process.

The primary aim of IPT is to produce causal accounts of specified outcomes. It is thus more directed than conventional interpretive mapping of social institutions, and it serves to establish causal processes that unfold over time. However, it does so by explicitly incorporating the study of such institutions at the basis of the approach. It thus depends on a configuration of causal and constitutive explanations, where the primary aim is to establish the causal relation between different points in the process leading up to the outcome.

By way of exemplifying, suppose we are interested in finding out why a particular civil conflict $Y_1$ broke out. We are interested in identifying the sequence of events that produced this specific outcome, in terms of the eruption of organized collective violence between two more or less well-defined social groups. Applying IPT, this will prompt an in-depth analysis of the social context in which the conflict was triggered, thus using constitutive analysis to define the causal field as it appears at a certain points in time. As part of this constitutive analysis, we will want to know how the agents involved in the conflict understood themselves, their actions and how they understood the situation they found themselves in. Indeed, this conforms to standard interpretive accounts in the sense...
that the aim is to enter the conceptual worlds of the subjects, to understand their practices, shared understandings, habits, and identities. In short, an analysis of the social system that constitutes the dispositions of agents. An in-depth understanding of the context will also help identify the appropriate contrast for the *explanans*, which means identifying the aspects of that particular context which departs from its “normal functioning.”

This assessment of normal functioning, which in this case may be the absence of violent conflictual dynamics, serves as our contrastive state, the counterfactual of the outcome. We are interested in the events that drove the agents involved to a situation characterized by all-out conflict and organized mass violence. This requires establishing the longer process, the causal history that produced this outcome. We might exemplify here with the type of studies suggested by Tilly to help explain the politics of collective violence. Such studies engage with the broader social processes through which boundaries are drawn between social groups, and through which meanings are gradually assigned to such boundaries (Tilly, 2017: 278–279).

IPT helps make the contrastive aspect of causal explanations concrete by using within-case comparisons. The preliminary identification of an event as the trigger for an outcome of interest may find further support through the identification of previous similar events that have occurred in the same social context but where such an outcome did not produce itself. That is, previous situations where there is evidence of the *explanans* but without the *explanandum*. This move has the potential of providing considerable explanatory leverage. Researchers may use such comparisons between different points in a longer process to demonstrate how changes in social background conditions make a certain set of causal relations possible. If the occurrence of collective violence is, as Tilly argues, dependent on the construction of social boundaries and the social assignment of meanings to such boundaries, tracing changes in such background conditions will be necessary to provide a causal explanation that relates a specific event to a particular outcome.

What this implies then is that the process tracing exercise proceeds by tracing broader changes in the causal field while simultaneously looking at how changes in the field reshapes the effects of the proximate cause X. The logic of this explanation depends on demonstrating how, for example, changes in collective self-understandings change the dispositions among particular agents. Changes in the social institutions in which agents find themselves may help demonstrate how similar events occurring at a point in time \( t \) and point in time \( t+1 \) are perceived differently and trigger different responses, producing different outcomes.

\[
\text{X rather than X}_1 \text{ caused Y rather than Y}_1 \quad \text{in causal field}^1_{t} \quad X \text{ rather than X}_1 \text{ caused Y}_1 \text{ rather than Y} \quad \text{in causal field}^1_{t+1}
\]

\[
\begin{array}{c}
\text{Causal field}^1_{t} \quad X \rightarrow Y \\
\end{array} \quad \rightarrow \quad \begin{array}{c}
\text{Causal field}^1_{t+1} \quad X \rightarrow Y_1 \\
\end{array}
\]
While a certain event might be accepted as part of the *normal* course of things and will not attract any particular attention at $t$, similar actions might be perceived as controversial, and *abnormal* and thus result in conflict at $t+1$. For the causal explanation of the conflict to make sense, there is a need to identify what in the context had changed, to produce these different reactions. As Tilly indicates, broad changes in self-understandings among a specific set of agents might appear as a crucial factor. This includes changes in shared stories about social boundaries as well as shifting relations across such boundaries (Tilly, 2017: 279). The interpretive twist here is that the identification of the causal “difference maker” is necessarily dependent on the researcher tapping into what the agents in this context themselves perceived as being the “normal functioning” of the social institutions they inhabit. This makes the research task one of identifying why they perceived an event as “abnormal” at $t+1$, and triggering some action, for instance the initiation of organized violence, when this had not been the case in the past.

IPT also directs attention to how individual events that constitute the process may reshape the causal field on which the subsequent steps play out. If we stick with the example of the outbreak of collective violence, we might identify a first set of actions which, while they may not have produced the outcome we are interested in, may have contributed to reconfigure broadly shared stories of groups and group relations in that particular context. That is, actions in a first instance, where conflict is not triggered, might still feed into collective perceptions of other agents as, for instance, belligerent, arrogant, unpredictable, and untrustworthy.

This conceptualization of causal processes is consonant with the notion that single causes can have multiple outcomes. In this case, a cause $X$ at $t$ produces $Y$ (no-conflict) due to the specific configuration of the causal field at that point in time. However, even if $X$ does not produce the outcome and does not trigger collective violence at this point it might still have an effect (even if marginal) on the broader causal field. Thus, $X$ might contribute to reshaping the background conditions for producing $Y_1$ (conflict) at a later point in time.

This argument can be understood in agent-centric terms, which implies that agents, whether individuals, states, or organizations through social interaction, continuously update their understandings of each other, and by consequence their respective understandings of what type of social situation they find themselves in (cf. Wendt, 1999: 168). However, it can also be understood in more macro-institutional terms where actions and their interpretations feed into prevalent shared understandings of groups, and work to reshape collective identities. This points to yet another possibility for IPT used as a tool for causal analyses of the macro process through which social institutions change. Whatever the precise application of the IPT model, it relies on the clear delimitation of constitutive and causal explanations and benefits from the definition of such explanations in contrastive terms.

**Conclusions**

The turn to causal inquiry by interpretive IR scholars carries considerable potential in terms of identifying drivers and mechanisms of various aspects of international politics that have remained underexplored. The study of armed conflict, international diplomacy,
dynamics of IOs, the foreign and security policies of states are all areas in which interpretive approaches have made important contributions, especially by paying close attention to how socially embedded agents understand their worlds. To add a clearly defined causal element to interpretive studies further strengthens their ability to push research on these issues forward. IPT serves as one possible application of the causal model that could help researchers achieve such aims.

The causal model developed in this article offers a more clearly defined basis on which causal explanations can be generated. Key to the model is the redefinition of the relationship between constitutive and causal explanations and a clear argument for how the nesting of one in the other help address key research aims with which interpretive scholars are concerned, and with which the IR field could benefit from exploring in more depth. It deals with questions of causation and constitution in a manner that, I argue, is more straightforward and productive than has often been the case in this field. The model also highlights at which points interpretive notions of causality intersect with those widely used in mainstream IR and at which points they differ.

In the views of some it will push interpretive accounts too far away from notions of the social world as characterized by ambiguity, contingency, and fluidity. The arguments here also have had less to say about interpretive IR as a critical emancipatory project. It is not necessarily the case, however, that the understanding of interpretive research developed in this article is incompatible with such a project. A critical stance toward how power relations structure different aspects of the social world also, arguably, benefits from clearly defined casual arguments.

An added benefit of the model is its potential for conversation between interpretive research and a broader field of scholars in IR. There is a value in applying causal language in a way that is recognizable as such, also beyond the confines of the interpretive field. It allows for discussions on what a reasonable vantage point might look like from which causal explanatory claims can be assessed. Interpretive IR should be careful to not accept, wholesale, criteria for conducting research developed in traditions that rely on sometimes radically different assumptions and different scientific ideals. However, a more well-defined notion of causal explanations from an interpretive perspective offers a steadier platform from which its contributions to IR as a pluralistic discipline can be highlighted.

Acknowledgements
I would like to thank Richard Ned Lebow, Audie Klotz, Frederic Schaffer as well as the reviewers and EJIR editors for constructive comments on previous drafts. A special thanks to Stefano Guzzini for his profound influence on the ideas on which the article is based. Thanks also to panel participants at the 2017 Annual meeting of the American Political Science Association, the 2018 Annual meeting of the International Studies Association, and the Global and Regional Governance seminar at Stockholm University where previous drafts of the article were presented.

Funding
The author received no financial support for the research, authorship, and/or publication of this article.
Notes

1. For debates along these lines in the broader literature on causal explanations in IR and beyond see Fearon (1991), Grynaviski (2013), Lewis (1986), Lebow (2014, 2015), Sartorio (2005), Schaffer (2005), Woodward (1984, 2003) and Ylikoski (2013).

2. For seminal contributions to more general discussions on the philosophy of the social sciences and causality in IR see Chernoff (2005); Kurki (2008); Jackson (2011); Wight (2006).

3. This definition relies on a modification of Ylikoski (2013) who only see the first of these forms as possible.

4. Interpretive studies are conceptualized here as a broad tradition relying on variants of social constructivist and pragmatist positions (cf. Bevir and Blakely, 2018; Hacking, 1999; Kratochwil, 2007; Lynch, 2014).

5. For a contrasting view see Pouliot (2016: 14–15). Tannenwald’s (2015) notion of “constitutive process studies” to answer “how possible” questions also appears to differ from the position outlined here.

6. This point parallels McKeown’s discussions on the statistical worldview (McKeown, 1999) but differs from that of Fearon (1991) who frames his discussion of comparative as well as single case studies explicitly in statistical terms.

7. Grynaviski (2013) similarly suggests a contrastive view of causation for IR. However, in contrast to Grynaviski, the argument here does not revolve around a focus on different types of research questions. Here I follow Ylikoski (2013) who argues, contra Wendt’s (1999) similar position, that the semantic conventions of questions tell us very little regarding what type of explanation might be required to answer them.

8. Here it should be noted that Woodward does not exclude causal explanations that rely on deductive reasoning or rely on covering laws. He only states that this denotes a specific variety of causal explanations (Woodward, 2003: 20).

9. And, one might add, if \( n^* \) would have been the case, so would \( a^* \).

10. For a recent application of the IPT framework studying the emergence of ethnoreligious conflict, see Magcamit (2020).

11. Mechanisms is a hotly debated issue and there has yet to emerge a broad consensus on their ontological status or what role they play for social scientific explanations. It lies beyond the aims of this article to engage in those debates. See for instance, Elster (1998, 1999), Hedstrom and Ylikoski (2010), Little (1991), and Johnson (2006) and for interpretive approaches to IR, Guzzini (2013).

References

Adler-Nissen R and Tsinovoi A (2019) International misrecognition: the politics of humor and national identity in Israel’s public diplomacy. *European Journal of International Relations* 25(1): 3–29.

Autesserre S (2010) *The Trouble with the Congo: Local Violence and the Failure of International Peacebuilding*. Cambridge: Cambridge University Press.

Banta B (2012) Analysing discourse as a causal mechanism. *European Journal of International Relations* 19(2): 379–402.

Barnett M and Finnemore M (2004) *Rules for the World: International Organizations in Global Politics*. Ithaca: Cornell University Press.
Bevir M and Blakely J (2018) Interpretive Social Science: An Anti-Naturalist Approach. Oxford: Oxford University Press.

Campbell D (1998) Writing Security: United States Foreign Policy and the Politics of Identity. 2nd edn. Minneapolis: University of Minnesota Press.

Campbell D (2013) Process Tracing Methods: Foundations and Guidelines. Ann Arbor: University of Michigan Press.

Bennett A and Checkel JT (eds) (2014) Process Tracing: From Metaphor to Analytical Tool. Cambridge: Cambridge University Press.

Bevir M and Blakely J (2018) Interpretive Social Science: An Anti-Naturalist Approach. Oxford: Oxford University Press.

Campbell D (1998) Writing Security: United States Foreign Policy and the Politics of Identity. 2nd edn. Minneapolis: University of Minnesota Press.

Cappoccia G and Kelemen D (2007) The study of critical junctures: theory, narrative and counterfactuals in historical institutionalism. World Politics 59(3): 341–369.

Checkel JT (2020) Research transparency and open science: too much of a good thing? Paper presented at the Annual Convention of the American Political Science Association, 10–13 September, 2020. Available at: https://www.eui.eu/Documents/DepartmentsCentres/SPS/Profiles/Checkel/APSAAnnualConvention-CheckelPaper.0920.pdf (accessed 28 September 2020).

Chernoff F (2005) The Power of International Theory: Reforging the Link to Foreign Policy-Making through Scientific Inquiry. Oxon: Routledge.

Cummins R (2000) “How does it Work?” versus “What are the Laws?”: two conceptions of psychological explanation. In: Keil F and Wilson R (eds) Explanation and Cognition. Cambridge: MIT Press, 117–144.

Dessler D and Owen J (2005) Constructivism and the problem of explanation: a review article. Perspective on Politics 30(3): 597–610.

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

Fairfield T and Charman E (2017) Explicit Bayesian analysis for process tracing: guidelines, opportunities, and caveats. Political Analysis 25(3): 363–380.

Fearon J (1991) Counterfactuals and hypothesis testing in political science. World Politics 43(2): 169–195.

Fierke K (2015) Critical Approaches to International Security. Cambridge: Polity Press.

Finnemore M (2003) The Purpose of Intervention: Changing Beliefs about the Use of Force. Ithaca: Cornell University Press.

Geertz C (1973) The Interpretation of Cultures. New York: Basic Books.

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

Gerring J (2005) Causation: a unified framework for the social sciences. Journal of Theoretical Politics 17(2): 163–198.

Grynaviski E (2013) Contrasts, counterfactuals, and causes. European Journal of International Relations 19(4): 823–846.

Guzzini S (2013) The Return of Geopolitics in Europe? Social Mechanisms and Foreign Policy Identity Crises. Cambridge: Cambridge University Press.

Guzzini S (2017) Power as cause. Journal of International Relations and Development 20(4): 737–759.

Hacking I (1999) The Social Construction of What? Cambridge: Harvard University Press.

Hall P (2013) Tracing the progress of process tracing. European Political Science 12(1): 20–30.

Hansen L (2006) Security as Practice. Oxon: Routledge.

Hart HLA and Honoré T (1985) Causation in the Law. Oxford: Oxford University Press.

Hausman D (2005) Causal relata: tokens, types, or variables? Erkenntnis 63(1): 33–54.
Hedstrom P and Ylikoski P (2010) Causal mechanisms in the social sciences. The Annual Review of Sociology 36: 49–67.
Holland P (1986) Statistics and causal inference. Journal of the American Statistical Association 81(396): 945–960.
Hopf T (2005) Identity, legitimacy, and the use of military force: Russia’s Great Power identities and military intervention in Abkhazia. Review of International Studies 3: 225–243.
Hopf T (2018) Change in international practices. European Journal of International Relations 24(3): 687–711.
Humphreys A (2017) Causation, complexity, and the concert: the pragmatics of causal explanation in international relations. Journal of International Relations and Development 20(4): 717–736.
Jackson P (2006) Civilizing the Enemy: German Reconstruction and the Invention of the West. Ann Arbor: University of Michigan Press.
Jackson P (2011) The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the Study of World Politics. London: Routledge.
Jackson P (2017) Causal claims and causal explanation in international studies. Journal of International Relations and Development 20(4): 689–716.
Johnson J (2006) Consequences of positivism: a pragmatist assessment. Comparative Political Studies 39(2): 224–252.
Klotz A and Lynch C (2007) Strategies for Research in Constructivist International Relations. Oxon: Routledge.
Krausz F (2007) Evidence, inference and truth as problems of theory building in the social sciences. In: Lebow R and Lichbach M (eds) Theory and Evidence in Comparative Politics and International Relations, 25–54.
Kurki M (2008) Causation in International Relations: Reclaiming Causal Analysis. Cambridge: Cambridge University Press.
Kurki M and Suganami H (2012) Towards the politics of causal explanation: a reply to the critics of causal inquiry. International Theory 4(3): 400–429.
Lebow R (2009) Constitutive causality: imagined spaces and political practices. Millennium: Journal of International Studies 38(2): 211–239.
Lebow R (2014) Constructing Cause in International Relations. Cambridge: Cambridge University Press.
Lebow R (2015) Counterfactuals and security studies. Security Studies 24(3): 403–412.
Lewis D (1986) Philosophical Papers: Volume II. Oxford: Oxford University Press.
Little D (1991) Varieties of Social Explanations. Boulder: Westview.
Lynch C (2014) Interpreting International Politics. Oxon: Routledge.
Mackie J (1965) Causes and conditions. American Philosophical Quarterly 2(4): 245–264.
Magcamit MI (2020) Insecurities in imagined communities: manufacturing the ethnoreligious others as security threats. International Studies Quarterly 64(3): 684–698.
McCurt D (2016) Practice theory and relationalism. International Studies Quarterly 60(3): 475–485.
McKeown T (1999) Case studies and the statistical worldview. International Organization 53(1): 161–190.
Moore M (2009) Causation and Responsibility: An Essay in Law Morals and Metaphysics. Oxford: Oxford University Press.
Nair D (2020) Emotional labor and the power of international bureaucrats. International Studies Quarterly 64(3): 573–587.
Norman L (2015) Interpretive process tracing and causal explanations. Newsletter of the APSA Organized Section for Qualitative and Multi-Methods Research 13(2): 4–9.
Norman L (2016) The Mechanisms of Institutional Conflict in the European Union. Oxon: Routledge.
O’Mahoney J (2015) ‘Why Did They Do That?’ The methodology of reasons for action. *International Theory* 7(2): 231–262.

Pierson P (2004) *Politics in Time: History, Institutions, Social Analysis*. Princeton: Princeton University Press.

Pouliot V (2014) Practice tracing. In: Bennett A. and Checkel JT (eds.) *Process Tracing: From Metaphor to Analytical Tool*. Cambridge: Cambridge University Press, 237–259.

Pouliot V (2016) *International Pecking Orders*. Cambridge: Cambridge University Press.

Sartorio C (2005) Causes as difference-makers. *Philosophical Studies* 123(1–2): 71–96.

Schaffer J (2005) Contrastive causation. *The Philosophical Review* 114(3): 327–358.

Schaffer J (2016) Causal contextualism. In: Blaauw M (ed.) *Contrastivism in Philosophy*. London: Routledge, 35–63.

Schimmelfennig F (2014) Efficient process tracing: analyzing the causal mechanisms of European integration. In: Bennett A and Checkel JT (eds) *Process Tracing: From Metaphor to Analytical Tool*. Cambridge: Cambridge University Press, 98–125.

Schwartz-Shea P and Yanow D (2013) *Interpretive Research Design: Concepts and Processes*. London: Routledge.

Smith R (2004) The politics of identity and the tasks of political science. In: Shapiro I, et al (eds) *Problems and Methods in the Study of Politics*. Cambridge: Cambridge University Press, 42–66.

Stappert N (2020) Practice theory and change in international law: theorizing the development of legal meaning through interpretive practices of international criminal courts. *International Theory* 12(1): 33–58.

Tannenwald N (2007) *The Nuclear Taboo: The United States and the Non-Use of Nuclear Weapons since 1945*. Cambridge: Cambridge University Press.

Tannenwald N (2015) Process tracing and security studies. *Security Studies* 24(2): 219–227.

Tilly C (2017) The politics of collective violence. In: Castañeda E and Schneider L (eds) *Collective Violence, Contentious Politics, and Social Change*. London: Routledge, 275–292.

Waldner D (2015) Process tracing and qualitative causal inference. *Security Studies* 24(2): 239–250.

Weldes J (1999) *Constructing the National Interest: The United States and the Cuban Missile Crisis*. Minneapolis: University of Minnesota Press.

Wendt A (1998) On constitution and causation in international relations. *Review of International Studies* 24(5): 101–118.

Wendt A (1999) *Social Theory of International Politics*. Cambridge: Cambridge University Press.

Wight C (2006) *Agents, Structures and International Relations: Politics as Ontology*. Cambridge: Cambridge University Press.

Woodward J (1984) A theory of singular causal explanation. *Erkenntnis* 21(3): 231–262.

Woodward J (2003) *Making Things Happen*. Oxford: Oxford University Press.

Ylikoski P (2013) Causal and constitutive explanation compared. *Erkenntnis* 78(2): 277–297.

Zehfuss M (2002) *Constructivism in International Relations: The Politics of Reality*. Cambridge: Cambridge University Press.

**Author biography**

Ludvig Norman is Associate Professor at Stockholm University, Department of Political Science and Senior Fellow at the Institute of European Studies at UC Berkeley. He earned is PhD in political science from Uppsala University. His research focuses on the institutions of the European Union, democratic theory, and social science methodology. He is the author of the book *The Mechanisms of Institutional Conflict in the European Union*. Recent works have appeared in *Political Studies, European Journal of International Relations, European Journal of Social Theory, Journal of Common Market Studies* and *Cooperation and Conflict*. 