EXPLANATIONS, MECHANISMS, AND CAUSES: TOWARDS EXPLANATORY PLURALISM IN INTERNATIONAL STUDIES

Jack Donnelly

Josef Korbel School of International Studies
University of Denver

jack.donnelly@du.edu

Paper prepared for presentation at
Columbia University, School of International and Public Affairs
September 24, 2019

Table of Contents

How one designs and carries out scientific research is intimately connected with how one understands scientific explanation – which shapes what one can (and cannot) know and understand.

A common view in International Studies[1] is that there are multiple "styles" (methods) of social scientific research but only "one logic of inference;"[2] that all scientific explanations have the same character or structure. In the philosophy of science today, however, it is generally agreed that different disciplines, and even different branches of the same discipline, employ different explanatory logics.[3] For example, Physics is centrally concerned with natural laws. Laws, however, play little role in Biology, which focuses instead on mechanisms.[4] And neither laws nor mechanisms play much of a role in International Studies.

I focus here on the contrast between associational explanations that predict outcomes and mechanismic explanations that show how an outcome was produced. Although International Studies has largely ignored mechanisms, "practically the whole sweep of natural phenomena as described by the physical, life, and social sciences, are the product of mechanisms."[5] Therefore, if science aims to understand how the world works (which seems to me obviously the case) then the social sciences should actively inquire into *how* the world works; that is, give central attention to mechanisms.

As Mario Bunge puts it, "if we wish to understand a real thing, be it natural, social, biosocial, or artificial, we must find out how it works. … real things and their changes are explained by unveiling their mechanisms."[6] This, I argue, implies a radically different kind of scientific research from that which has dominated mainstream International Studies in recent decades.

In depicting the social-scientific mainstream, I focus on Gary King, Robert Keohane, and Sidney Verba's *Designing Social Inquiry: Scientific Inference in Qualitative Research* (KKV), the predominant research design text of the past quarter century;[7] Henry Brady and David Collier's edited volume *Rethinking Social Inquiry: Diverse Tools, Shared Standards*[8] (B&C[9]), the leading mainstream alternative;[10] and Gary Goertz's *Multimethod Research, Causal Mechanisms, and Case Studies: An Integrated Approach*,[11] which is, in my view, the best recent work on multimethod research.

1 I use the term "International Studies" rather than, for example, "IR and Comparative Politics" in an attempt to avoid not only disciplinary bias (especially in the label Comparative Politics) but also any particular (in this case binary) categorization of the subject matter studied.

2 (King, Keohane, and Verba 1994, 3).

3

4

5 (Glennan 2017, 48. Cf. 3).

6 (Bunge 1997, 410. Cf. 455).

7 [Cite}. KKV has ??? cites [comparison]

8 (Brady and Collier 2010).

9 More precisely, by B&C I will mean the chapters in R*ethinking Social Inquiry* that Brady on Collier wrote or co-wrote (which I will cite by chapter).

10 ???

11 (Goertz 2017).

These texts demonstrate that International Studies has in recent decades made considerable progress towards *methodological* pluralism.[12] They also, however, show that it still clings to *explanatory* monism. I argue for a deep explanatory pluralism open to the full range of questions required to understand how our complex and varied social world works.

1. VARIETIES OF EXPLANATIONS

To explain, in ordinary language, is "to make plain or intelligible … give an account of in order to bring about understanding," "to be the reason or rationale for, to account for; to be the cause of."[13] As Brady and Collier (B&C[14]) put it, an explanation is "a statement about why an outcome has occurred."[15] This roughly corresponds to the ordinary-language sense of explain as

"Why?" questions, however, have many different types of answers. Although explanations usually can be formulated as "Because …" there are many kinds of "because."

1.1. Laws

A scientific law (or law of nature) is expressed by a "statement that a particular phenomenon always occurs if certain conditions be present."[16] For example, Newton's law of gravitation says that the attraction between two particles is directly proportional to the product of their masses and inversely proportional to the square of the distance between their centers.

Nomological-deductive (covering law) explanations show that something is an instance of a law.[17] Why do the children of Adam and Esther have a one in four chance of inheriting Tay-Sachs disease? Because of Mendel's (Third) Law of Dominance, which states that a dominant allele always masks a recessive allele (and thus a recessive Mendelian trait, such as Tay-Sachs, will only be expressed in a recessive homozygous phenotype, which has a one in four chance of being realized in the offspring of heterozygous parents).

This kind of "why" answer states an invariant relationship between things in the world. Why? *Because.* (This is the way the world is.)

Nomological-deductive explanations are not uncommon in the physical sciences. An influential view in mid-twentieth-century philosophy of science held that all scientific explanations took this form.[18] In fact, however, they are rare in Biology and other sciences that deal with irreducibly complex phenomena.[19]

[12] I do not address the issue of bridging the divide between quantitative and qualitative research, which has been at the core of the discussion of multimethod research since KKV, both because this is no longer a matter of controversy (in principle at least) and, more importantly, because it is largely irrelevant to the logical form of scientific explanations.

[13] *Oxford English Dictionary*.

[14] See n. ???.

[15] (Seawright and Collier 2010, 329. Cf. 325). "To explain an event is to give an account of why it happened." (Elster 1989, 3).

[16] *Oxford English Dictionary*.

[17]

[18]

[19] (Giere 1999) presents a particularly lively and forceful argument against the idea of law-based sciences. See also ???.

1.2. Associations

Associational explanations, which are the most common form of scientific explanation in International Studies, provide "because of what" answers to "Why?" questions. If X (i.e., X_1, X_2 … X_n) both precedes Y and is sufficiently closely associated with it then we are warranted in considering X to be in some degree responsible for ("a cause of") Y. For example, statistical associations (e.g., regression models) explore the probability of Y given X. Set-theoretic approaches (e.g., QCA) look for Xs that are necessary or sufficient conditions for Y.[20]

The underlying idea is, roughly, that the things for which X variables stand have certain properties or powers that bring about changes in the world that result in Y.[21] Although those powers usually cannot be observed, their effects can – allowing us to work back from effect to infer cause. Y is explained, in part, by the prior presence of X, which (inferential techniques warrant us believing) influences the production of Y.

"Smoking causes lung cancer." Although about ninety percent of cases of lung cancer are associated with smoking, most smokers do not get lung cancer. And some nonsmokers do. Those who have smoked, though, are much more likely to develop the disease than those who haven't. And if a smoker does have lung cancer we can say that it probably is because she smoked – although many other things were also causes of her disease (and it is possible that she might have developed lung cancer had she not smoked).

Laws explain by reference to a universal pattern sufficient to predict an outcome (in the conditions in which the law applies). For example, having two copies of "the cystic fibrosis gene" assures getting the disease.[22] And one can calculate the probability of any particular child inheriting cystic fibrosis from the relevant genetic information about her parents.

Associations explain by contingent (empirical) relations.[23] Their predictions apply to populations not individuals – and even then only within a certain margin of error. And, at least in the social sciences, associations provide only part of "the explanation" (e.g., regression equations always have an error term).

1.3. Mechanisms and Processes

Mechanismic explanations take still a different form. They reference processes that produce outcomes. Why? Because that's how the world works.

[20] On the similarities between statistical and set-theoretic approaches, see (Paine 2016). (Mahoney and Goertz 2006) emphasizes their differences.

[21] The alternative is that causality is a relationship between variables; an analytic device rather than a force in the world. As Nobel-prize-winning economist James Heckman puts it, "causality is a property of a model." (quoted in Collier, Brady, and Seawright 2010b, 6). "A model is a set of possible counterfactual worlds constructed under some rules. … A model is in the mind. As a consequence, causality is in the mind." (Heckman 2005, 2).

In that case, though, it is hard to understand the mainstream social scientific focus on, even obsession with, causality. Causes, thus understood, really don't explain very much or very deeply (because of their tenuous connection to the world and how it works). And even if this is what causation "really means," this makes "causal explanation" but one type of scientific explanation – recall that scientific laws are not causes in this sense – and thus my central argument of the diversity of scientific explanations is not touched. Cf. §§ ???.

[22] Things are quite a bit more complicated – see, for example, ??? – but this is a reasonably good lay approximation.

[23] Although these empirical regularities are often stated probabilistically and referred to as law-like regularities, they are not probabilistic laws– which state that something *always* occurs, with a fixed probability.

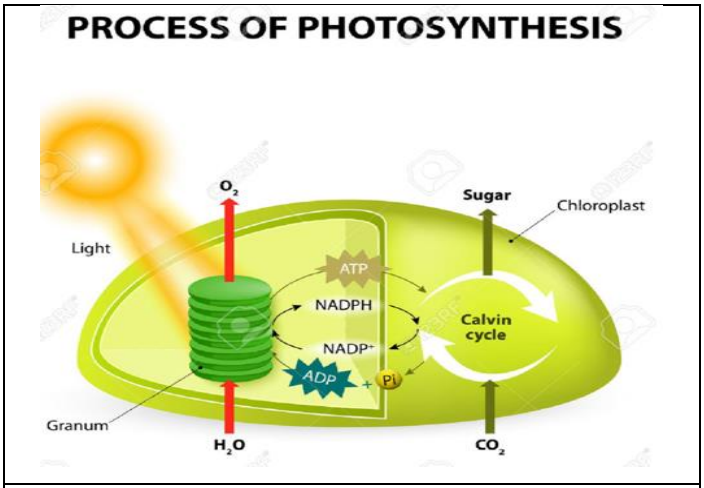
In ordinary language, a mechanism is "a system of mutually adapted parts working together in a machine or in a manner analogous to that of a machine" or "an ordered sequence of events involved in a biological, chemical or physical process."[24] This sense is also standard in the philosophy of Biology.[25]

The most cited definition is that of Peter Machamer, Lindley Darden, and Carl Craver (known in the field as MDC): "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions."[26] William Bechtel and Adele Abrahamsen similarly define a mechanism as "a structure performing a function in virtue of its component parts, component operations, and their organization."[27]

Phyllis Illari and Jon Williamson, looking at "mechanisms across the sciences," note that

> All mechanistic explanations begin with (a) the identification of a phenomenon or some phenomena to be explained, (b) proceed by decomposition into the entities and activities relevant to the phenomenon, and (c) give the organization of entities and activities by which they produce the phenomenon. … There is consensus that … mechanisms have two distinct kinds of constituents. We have 'entities', 'parts' and 'component parts' used for the bits and pieces of the mechanism, and 'activities', 'interactions' and 'component operations' for what those bits and pieces do.[28]

Consider a textbook example: photosynthesis.



Source: https://www.123rf.com/photo_41854862_stock-vector-photosynthesis-diagram-of-the-process-of-photosynthesis-showing-the-light-reactions-and-the-calvin-c.html

[24] *Oxford English Dictionary*.

[25] Here I take (Machamer, Darden, and Craver 2000), (Bechtel and Abrahamsen 2005), (Illari and Williamson 2012), and (Glennan 2017, ch. 1, 2, 8) as representative of what is sometimes called "the new mechanical philosophy." For a broader overview, see (Glennan and Illari 2017). I have also found (Bunge 1997) to be especially helpful (despite – or perhaps because of – his somewhat idiosyncratic systems perspective).

[26] (Machamer, Darden, and Craver 2000, 3). This seminal article has almost 2,500 Google Scholar citations. (For purposes of comparison, this is about a fifth more than (Brady and Collier 2010).)

[27] (Bechtel and Abrahamsen 2005, 423). Compare (Glennan 2017, 17).

[28] (Illari and Williamson 2012, 123, 125). Cf. (Steel 2008, 40-42), (Bechtel 2016, 705-706), (Glennan 2017, 19-20).

Photosynthesis in plants takes place within organelles called chloroplasts. Light energy is processed in the granum, initiating an electron transport chain that reduces NADP to NADPH and creates an energy gradient that is used by ATP synthase to produce ATP. ATP then participates in the Calvin cycle, along with carbon dioxide taken from the atmosphere, to produce a sugar.

I will use the formula "productively organized entities and activities."[29] Entities and activities are the interdependent elements of mechanisms.[30] Their organization into a process makes them "parts" of a mechanism.[31] What mechanisms "do" is produce particular phenomena.[32] Studying how these productive processes operate is the essence of mechanismic analysis. "Scientists aim to give a detailed account of how the phenomenon is produced by entities and activities. … Mechanistic explanation succeeds when the mechanism discovered and described is the mechanism responsible for the phenomenon."[33]

Mechanismic explanations are a type of processual explanation, understanding "process" in the ordinary-language sense of "a continuous and regular action or succession of actions occurring or performed in a definite manner, and having a particular result or outcome."[34] Agent-based models[35] and narrative explanations[36] are other types of processual explanations.

Mechanismic/processual explanations show *how* a process produces a result. Demonstrating the operation of a productive process *is* the explanation.

1.4. Other Types of Explanations

Although the differences between laws, associations, and mechanisms clearly establish the fundamental diversity of scientific explanations, it probably is worth noting several other types of explanations that are common in International Studies

Interactional explanations explain by the structure of the interaction. Consider game-theoretic models such as Prisoners' Dilemma. Although predictive, such explanations involve "more" than associations. Usually, though, do not rely on mechanisms (most game-theoretic explanations are predictive "as if" accounts rather than descriptive accounts of actual productive processes) or laws of nature (for example, "irrational" action is more or less common).

Relational explanations identify "emergent" effects that arise from the arrangement of the elements of irreducible wholes. Common framings include complex systems, networks, and fields – which

[29] Darden (2008, 965, Table 1) adds that biological mechanisms have spatial arrangement characterized by localization, structure, orientation, connectivity, and compartmentalization, components that exhibit temporal order, rate, frequency, and duration, and typically are located in hierarchies and series. These elaborations point to productive strategies for exploring mechanisms that, unfortunately, lie beyond our scope here.

[30] (Glennan 2017, 20-22, 29-36), (Darden 2008, 961-964), (Machamer 2004, 28-30, 32-34).

[31] "Mechanisms behave as they do because of the organized activities and interactions of their parts." (Glennan 2017, 23). Cf. (Bechtel and Abrahamsen 2005, 430; 2016, 719), (Illari and Williamson 2012, 127).

[32] Although a phenomenon may be produced by different mechanisms (see §6) a particular mechanism produces a particular phenomenon (or set of phenomena). One mechanism, one phenomenon.

[33] (Illari and Williamson 2012, 123). Cf. (Bechtel and Abrahamsen 2005, 430).

[34] *Oxford English Dictionary*. For similar scientific definitions, see ???. Although "mechanism" and "process" can sometimes be used interchangeably, I take "mechanism" to refer to a *regular* (systemic; structured) "process." This usage is at least as well established in ordinary language, does not waste a perfectly good word, and is compatible with both processual and non-processual ontologies.

[35]

[36] (Suganami 2008) offers a strong brief argument for narrative explanations.

generate system effects, network effects, and field effects. Why does information diffuse so much more rapidly today than in the past? Because of the structure of communications networks. Why do certain kinds of men (and women) tend to rise to positions of power in France? Because the field of power in the French state advantages those with certain kinds of capital and dispositions.[37]

Intentional explanations are common in both scientific and lay contexts, especially when human action is involved. Why did you do that? Because I wanted to help him.

Explanations by preferences, habits, dispositions, emotional states, "conventions" (understood broadly to include customs, norms, roles, rules, and laws), and "instincts" also have a similar form. An action is explained by reference to an internal state of the actor, whether that state is individual or social in origin and whether it is conscious, subconscious, dispositional, or "programmed." "The nature" of the actor does the explaining.

We also often explain an entity or activity as being *functional*. For example, "fitness" is a largely functional notion in evolutionary accounts. Designed or purposive entities or activities are commonly explained functionally. (Why does that part have that form? Because of its role in that system. Why did you do that *that* way? Because that works well in these circumstances.)

2. Explanatory Diversity in Diverse Sciences

Because I have not tried to present a comprehensive typology of explanations, it does not matter that, for example, many people are skeptical of the scientific credentials of functional or narrative explanations or that the distinction drawn above between relations and processes is obscure. My point was to demonstrate beyond any reasonable doubt that diverse types of explanations are regularly employed in the social sciences.

It simply is not true, as KKV claim, that "all good [scientific] research can be understood – indeed, is best understood – to derive from the same underlying logic of inference."[38]

Part of the explanation for this (obviously incorrect) insistence on a singular logic of inference, or in B&C's terms "shared standards," seems to be a prescriptive vision of "science" as a fundamentally unitary enterprise. For example, B&C define "scientific" as "a normative view of the theoretical, methodological, and empirical goals of research."[39] And they advocate a narrow prescriptive account: "the overarching goals of valid descriptive and causal inference and of building theory are central to the idea of shared standards."[40]

Many mid-twentieth-century philosophers of science championed "the unity of science."[41] Today, however, fundamental differences between different types of scientific explanations are almost universally acknowledged.[42] And contemporary philosophy of science typically seeks to provide a

37

38 (King, Keohane, and Verba 1994, 4).

39 (Seawright and Collier 2010, 348).

40 (Seawright and Collier 2010, 349). Although this may seem uncontroversial to most mainstream social scientists, in §§4 and 7 I show that causal inference is only one type of explanation and in §9 I argue that theory need not be the organizing framework of scientific research.

41

42

philosophical account of scientific "best practice" (rather than advocate a prescriptive vision of science).[43]

In this light, consider KKK's claim that scientific research has "four characteristics. 1. The goal is inference. … 2. The procedures are public. … 3. The conclusions are uncertain. … 4. The content is the method."[44] Inference through rigorously applied public procedures can take many forms. KKV's narrow vision of causal explanation[45] rests not on the nature of the enterprise of science but on a particular prescriptive vision of how to carry out scientific research. And that vision is inconsistent with the practice of natural and social scientists.

I focus below on the contrast between associational and mechanismic explanations not only for reasons of space and ease of exposition as well as because their striking differences nicely illustrate my broader argument for explanatory pluralism but also for substantive reasons. Associational explanations dominate self-consciously social-scientific International Studies. Mechanisms, however, have received short shrift. And when "causal mechanisms" have been addressed, they have typically been treated associationally (see §5) or interactionally, not mechanismically. I argue instead for giving mechanismic explanations a central place in scientific International Studies.

3. PREDICTIVE ("WHAT") VS. PROCESSUAL ("HOW") EXPLANATIONS

Associations and mechanisms are examples of two broad classes of scientific explanation: predictive ("what") explanations and processual ("how") explanations.

Some explanations, including nomological-deductive, associational, and game-theoretic explanations, predict outcomes. They tell us what we can expect to happen, because of what.

Other explanations, such as mechanismic and narrative explanations and agent-based models, chart the operation of productive processes. They show us how the world works to produce an outcome.

Associational (correlational and set theoretic) explanations establish relations of probability, necessity, or sufficiency between variables (or the states of the world that they represent). Explanation is understood as prediction; being able to say, with more or less confidence, that if X then Y. Such explanations tell us "what" – not "how."

For that we need processual explanations. For example, a mechanismic account of photosynthesis does not merely show that the production of carbohydrate is invariably associated with water, carbon dioxide, and energy from light – or even that it follows a particular formula. It tells us *how* the reaction occurs.

Some mechanists argue that "what" answers to "Why?" questions are not really explanations. "Predictivists believe that any good predictive model is explanatory. Mechanists believe that, in addition, explanation requires knowledge of a mechanism. … To explain a phenomenon, one has to know the mechanism that produces it."[46] "For explanatory purposes the mechanism is what matters. It provides understanding, whereas prediction at most offers control."[47]

[43]

[44] (King, Keohane, and Verba 1994, 7-9).

[45]

[46] (Craver and Kaplan 2011, 269).

[47] (Elster 1989, 10).

Knowing that given X we can expect Y, however, certainly is knowledge – and knowledge that can facilitate effective action in the world. Therefore, it seems to me inappropriately prescriptive to insist that causal associations do not explain.

Nonetheless, associational "what" explanations, compared to mechanismic "how" explanations, are weak and incomplete. Mechanismic explanations, in addition to identifying causes and predicting outcomes, tell us how those outcomes are produced. They are the "gold standard" of explanation.

This implies reversing KKV's prioritization of regression models (and other associational estimates of causal effects) over process tracing.[48] Showing the operation of a mechanism provides *stronger* (better, deeper, more revealing) explanations. Or, if this formulation seems overstated, there is no good reason, in general or in principle, to prefer associational "because of what" explanations over mechanismic "how" explanations.

4. CAUSATION AND EXPLANATION

Brady and Collier subtitle their book "Diverse Tools, Shared Standards," echoing (in a slightly more expansive way) KKV's "one logic" formulation. Furthermore, they advocate a unitary "statistical" model of explanation.[49] Although this is a bit broader than KKV's regression-based model it is no less monistic. And it reflects the general tendency of mainstream methodologists to understand explanation as (associational) causal inference.[50]

4.1. Humean Causes, Causal Effects, and Counter-Factual Causality

Associational explanations, not surprisingly, typically employ an associational conception of causation that is usually traced back to Hume, who famously defined a cause as something prior in time to and constantly conjoined with a type of phenomenon (effect).[51]

Associational research in International Studies usually aims to determine "causal effects," understood as "the difference between the two values of the dependent variable that arise according to whether an independent variable assumes one of two specific values."[52] We compare two situations that are the same except for the presence or absence of X. If Y is present when X is present but absent when X is absent, and if X precedes Y, then we are likely to be warranted in saying that X causes Y.

[48] ???. For now, I will treat "process tracing" as analysis of the operation of a mechanism, relying on the ordinary-language sense of the phrase "process tracing." However, see §13.

[49]

[50] A classic example in IR is (Hollis and Smith 1990), which distinguishes explanation, which appeals to causes, from understanding, which appeals to reasons and other forms of non-causal "explanation" (in the ordinary-language sense). Recall the *OED* definition of explain: "to make plain or intelligible … give an account of in order to bring about understanding."

[51]

[52] (Seawright and Collier 2010, 316). Cf. (King, Keohane, and Verba 1994, 81-82), (Brady 2010, 74), (Imbens and Rubin 2015, 6, 21), (Goertz 2017, 75).

In mainstream social-scientific International Studies, a "counter-factual" or "potential outcomes" definition of causality[53] (causal effect) predominates.[54] Counter-factual causality involves "the difference between what actually happened and what would have happened if some prior circumstance(s) had been different in a particular way."[55] In estimating causal effects, we compare not two facts but one fact with a "counter-factual" (i.e., imagined) alternative; an actual outcome with a "potential" (i.e., fictional) outcome.

This is necessary when, as is usually the case in the social sciences, we cannot in fact compare two cases that are otherwise the same. For example, if we want to know why Ann voted for Trump, we cannot compare her voting for Trump with her voting for Clinton and then work back to find what differed in the two cases.

Here I will note, but not pursue, the strangeness of the idea that scientific explanation depends not on knowledge of the world but on "reasoning about phenomena that did not occur."[56] Neither will I pursue the problems of establishing that cases, whether actual or hypothetical, are otherwise the same – which have been extensively explored in the methodology literature. I focus instead on the underlying associational account of causation.

4.2. Causal Inference

A constant conjunction account of causation is most plausible where the conjunctions really are constant. For example, gravity in a certain sense *is* the invariant relationship between mass and attraction. Why did the glass you dropped fall? Because of gravity.

Things become more problematic when the association is only more or less regular. We then must *infer* causation; use what is known (effect) to establish what is not known (cause).

Furthermore, and no less importantly, the inability to say "how" becomes increasingly troubling as the constancy of the conjunction declines. With a law (*constant* conjunction) we usually can "live with" lack of knowledge of the mechanism (although it may impede our ability to act effectively). When we have an r^2 of 0.5 – when, at best, we barely have half an idea of what is going on – lack of knowledge of mechanisms is a much more serious explanatory shortcoming (although we often are forced to live with it).

These problems are further exacerbated in counter-factual accounts, in which we cannot observe half of the crucial data. This generates the so-called fundamental problem of causal inference: "causal inference implicitly depends on a comparison with something that did not occur."[57]

This "fundamental problem," however, arises from the counter-factual definition of causation, which makes crucial evidence unobservable *in principle*.[58] In an associational account that is not counter-factual, any problem of inference is the ordinary problem of acquiring knowledge of

53 (Morgan and Winship 2015 [2007}) #13057} and (Imbens and Rubin 2015) provide extensive technical discussions (and the first chapter of each book offers a brief nontechnical overview of the approach.)

54 It is endorsed whole-heartedly by KKV (King, Keohane, and Verba 1994, 76-82) and with modest reservations by B&C. (Seawright and Collier 2010#322-323, 340-341), (Brady 2010, 71, 81-82), (Collier, Brady, and Seawright 2010c, 163-164). It is also implicitly adopted by Goertz. ???

55 (Seawright and Collier 2010, 322-323).

56 (Seawright and Collier 2010, 322).

57 (Collier, Seawright, and Munck 2010, 38). (Actually, this dependence is explicit, not implicit.) Cf. n. 52.

58 As B&C (Seawright and Collier 2010, 317 [emphasis added]) put it, "*given a counter-factual definition of causation*, the problem is … "

empirically unobservable entities or processes that are known by their effects or manifestations. *This* problem is no different for causation than for, for example, life, consciousness, power, authority, or solidarity (and thus cannot be the fundamental problem *of causal inference*).

Furthermore, and crucial for my purposes here, mechanismic explanations face no problem of causal inference – because there is no causal *inference*; that is, no "reasoning from something known or assumed to something else which follows from it;" no "forming of a conclusion from data or premisses, either by inductive or deductive methods."[59] Causation is established by tracing a process of production. That is what causation *means* (in this context).

Mechanismic explanations do not infer causes. They reveal productive processes of causation. Causal inferences are required only when we lack knowledge of the mechanism of production. And that lack of knowledge of mechanisms is the source of associational problems of causal inference.

4.3. Associational Causes vs. Mechanismic Causation

Mechanisms are "causal" in a perfectly good ordinary-language sense. (They produce (cause) an outcome (effect).) In fact, a mechanism *considered as a whole* is "a cause" in the Humean sense.

Mechanismic explanations, however, which show the operation of productively organized entities and activities, depend on a radically different conception of causation.[60]

A mechanism, in Humean terms, involves *necessary connections* (not just more or less constant conjunctions). A particular sequence of activities, by entities of particular sorts, produces a particular outcome. Necessarily.

Associational explanations, as we saw above, seek to identify *causal effects*. Mechanismic explanations seek to understand the operation of *productive causal processes*.

Associational explanations establish *causal relevance*; grounds for believing that particular "things" are part of the story of causation.[61] Mechanismic explanations establish *causal efficacy* by identifying productive (causal) processes.[62]

Associational approaches identify *causes*; the entities and activities that precede (and can be inferred to be responsible for) an outcome. An associational explanation catalogues causes; tells us what causes what.

Mechanismic approaches identify processes of *causation*. A mechanismic explanation documents the operation of a productive process; shows us how a process produces ("causes") an outcome.

We thus see, again, that mechanismic explanations are, in principle, "better" – broader, deeper, and more inclusive. Good knowledge of a process of causation encompasses good knowledge of the causes involved. Knowledge of causes alone, however, may tell us little about the process of causation.

Often, however, the best we can do is establish causal relevance; acquire knowledge of (associational) causes that leaves us far short of understanding processes of causation. The resulting

59 *Oxford English Dictionary*.

60 (Glennan 2017, ch. 6) provides a useful overview, from a mechanismic perspective, of alternative accounts of causation.

61 For example, B&C (Seawright and Collier 2010, 318) define a cause as "a factor that helps to bring about the occurrence of an outcome." (I take the term "causal relevance" from (Glennan 2017, 150-151).)

62 (Glennan 2017, 154-155) contrasts "causal relevance" to "causal production."

research focus on associations, however, is a pragmatic adjustment to the contingent limits of our epistemic reach, not demanded by the nature of scientific explanation.

4.4. Causes and Explanations

The social-scientific mainstream in International Studies has largely reduced explanation to (associational) causal inference. KKV are typically clear and forceful: "We regard arguments in the literature about 'noncausal explanation' as confusing terminology; in virtually all cases, these arguments are really about causal explanation or are internally inconsistent."[63] B&C do consider the possibility of non-causal explanations.[64] In the end, though, not only do they adopt an agnostic view and fail to encourage pursuing non-causal (e.g., mechanismic) explanations, they remain narrowly focused on causal inference, taking for granted that social science aspires to "improving and evaluating [causal] inference."[65] So does Goertz. "Multimethod work involves cross-case causal inference AND within-case causal inference. Multimethod in this book means complementary causal inference methodologies."[66]

Few people without extensive professional training, however, would agree that causation "is" a predictive association among variables, let alone that scientific explanation "is" fundamentally a matter of identifying "causes" thus understood. And this commonsense understanding is supported by scientific practice in disciplines such as Physics (which gives considerable attention to both laws and fields) and Biology (which is particularly concerned with mechanisms).

My argument is that it is long past time for International Studies to return to the idea that explanations provide understanding, which comes in many forms. Associational causal analysis is but one way to generate one type of scientific knowledge. And it deserves no special pride of place.

Mario Bunge nicely makes a very similar point. "Whereas every social cause has (by definition) a social effect, not every social change results from a social cause. The methodological consequence is obvious: not every correct explanation in social science is of the causal type."[67]

Mechanists insist that mechanismic explanations should be at least as central to International Studies as (associational) causal explanations. And if mechanisms are "real systems in nature"[68] – "pieces of the furniture of the real world"[69] – they have a special claim on our scientific attention (especially in contrast to the imagined values of variables that lie at the heart of counter-factual explanations).

5. MECHANISMS ARE NOT INTERVENING VARIABLES

KKV largely dismiss mechanisms.[70] B&C, however, add "causal process observations" to standard "data set observations."[71] And Goertz places "causal mechanisms" at the heart of multimethod

63 (King, Keohane, and Verba 1994, 75 n. 1). (Brady 2010, ???) assembles and discusses several similar passages.

64 (Brady 2010, ???).

65 (Collier, Brady, and Seawright 2010a, 130).

66 (Goertz 2017, 5).

67 (Bunge 1997, 434).

68 (Bechtel and Abrahamsen 2005, 424-425).

69 (Bunge 1997, 414).

70

71 See especially (Brady 2010). Cf. (Seawright and Collier 2010, 318). B&C, however, only advocate including *observations* derived from the examination of cases into broadly "statistical" efforts to assess causal effects. They also explicitly define a causal mechanism as "a link or connection in a causal process. In the relationship between a given independent

research, both in the title of his book and in its content. In fact, Goertz "argues for taking a causal mechanism view of research."[72]

This openness to mechanisms, however, is more apparent than real. Some accounts in the social sciences do understand mechanisms as productively organized entities and activities. (For example, John Goldthorpe defines a mechanism as "some process existing in time and space … that actually generates the causal effect of X on Y and, in so doing, produces the statistical relationship that is empirically in evidence."[73]) Goertz, however, accurately reflects most mainstream multimethod research in treating "causal mechanisms" as either associational intervening variables or game-theoretic models – both of which take *mechanisms* out of "causal mechanisms."

Consider Goertz's "standard diagram used to conceptualize causal mechanisms."[74]

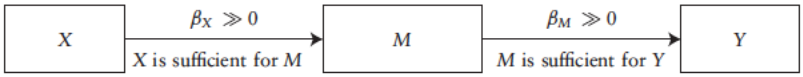


Figure 2.1: Conceptualizing causal mechanisms.

Source: Goertz 2017, 31

This is a textbook representation of an *intervening variable*. (Mechanismic philosopher Stuart Glennan presents a nearly identical diagram to illustrate what he calls "bare causal explanations [that] simply assert the existence of causal dependencies, arising from some unspecified mechanism."[75]) B&C even explicitly define a causal mechanism as an intervening variable.[76]

Furthermore, both of Goertz's interpretations of this "standard diagram" strip any sense of mechanism from "causal mechanisms." Neither the positive association of M and Y nor M's sufficiency for Y tells us anything about how (the mechanism by which) M produces Y.

A "causal mechanism" thus understood is a black box. A mechanismic explanation, however, "open[s] up the black box and show[s] the nuts and bolts, the cogs and wheels of the internal machinery."[77] "The basic idea of a mechanism-based explanation is quite simple: … proper explanations should detail the cogs and wheels of the causal process through which the outcome to be explained was brought about."[78]

Goertz, though, explicitly endorses a black box reading. "The M box of the causal mechanism figure is often called the 'black box.' This is because the causal mechanism is not known."[79] We

variable and a given dependent variable, a causal mechanism posits additional variables, sometimes called intervening variables." (Seawright and Collier 2010, 317).) The detailed critique of Goertz that I offer in what follows thus applies equally to B&C.

72 (Goertz 2017, 167).

73 (Goldthorpe 2001, 9).

74 (Goertz 2017, 31). Goertz (oddly) never defines "causal mechanism," presents a typology, or otherwise tries to clarify and delimit what "causal mechanism" means.

75 (Glennan 2017, 226).

76 See (Seawright and Collier 2010, 317) quoted in n. 71. Cf. (Mahoney 2001, 578). "It is common in political science work that utilizes mechanismic thinking to conflate mechanism with intervening variable." (Falleti and Lynch 2009, 1147).

77 (Elster 1983, 24). Cf. (Falleti and Lynch 2009, 1147).

78 (Hedström and Ylikoski 2010, 50). Cf. (Mayntz 2004), (Falleti and Lynch 2009).

79 (Goertz 2017, 39. Cf. 54).

thus have a "causal mechanism" in which "*the causal mechanism is not known*"! (Roughly: "Something seems to be going on here, but damned if I know what.") And this is more or less true of all of Goertz's (numerous) examples.

The first example in Goertz's book is based on Jon Pevehouse's study of the impact of democratic IGOs on democratic transitions.

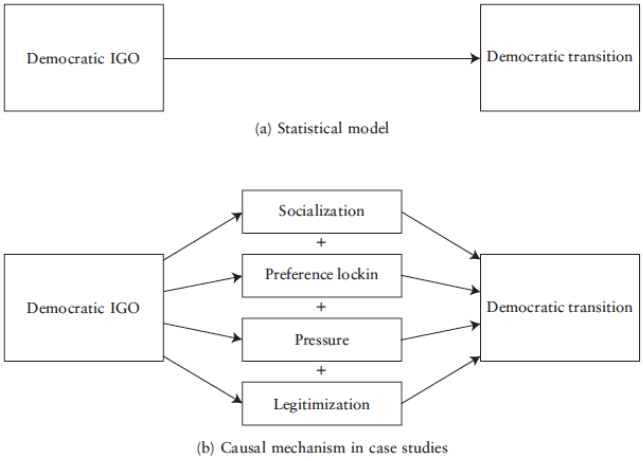


Figure 1.2: Causal mechanisms and statistical multimethod research: democratic IGOs and democratic stability. Source: based on Pevehouse (2005), table 5.1.

Source: Goertz 2017, 7, Fig. 1.2

The "causal mechanism" model on the bottom of the figure has the same form as the "statistical model" at the top. It simply adds an intervening variable – or, rather, four intervening variables. And these independent variables are modeled as separate black boxes that independently influence the outcome. These "parts of the mechanism" are so loosely related that the mechanism as a whole is not even modeled as having an effect. This jumble of elements tat best suggests where we might look for mechanisms.[80]

Or consider the first model in Goertz's chapter "Causal Mechanisms."

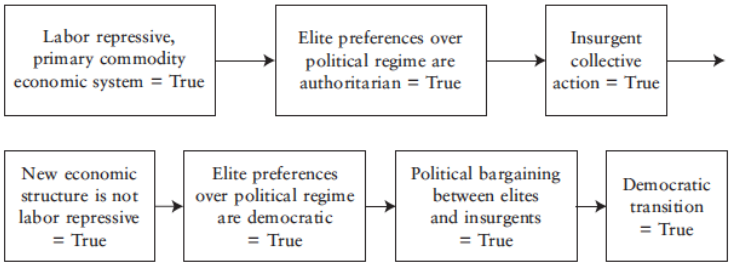


Figure 2.2: Causal mechanisms as causal chains: Wood's analysis of the democratic transition in El Salvador. Source: based on Waldner (2015).

Source: Goertz 2017, 32

[80] Although the labels might represent either processes or outcomes, the model clearly treats them as outcomes (effects) that cause later outcomes.

This simply says that if, starting at X, *a*, *b*, *c*, *d*, and *e* happen, then Y will result. Each box includes only the presence (= True) of a phenomenon; *an effect – not the mechanism that produces it.* This set of black boxes at most gestures in the direction of mechanisms.[81] It is, as the title of the figure accurately puts it, a causal chain – which is *not* a mechanism.

Finally, consider Goertz's treatment of John Snow's classic mid-nineteenth-century epidemiological work on the transmission of cholera. Although Snow established a connection with contaminated water, he knew only that "something" in *that* water (but not *this* water) "somehow" transmitted cholera. This was an associational, not mechanismic, explanation. As Goertz puts it, "it is critical to understand that Snow could not really get at the causal mechanism of cholera. This would require virus theory, Pasteur, and developments in biochemistry."[82]

Nonetheless, Goertz not only calls Snow's association of cholera and "bad" water a "causal mechanism[83] but identifies a general "causal mechanism X = 1."[84] This is a value of a variable (or a state of the world represented by that value); an effect not a mechanism.

To "interpret" a mechanism as a set or string of causal variables turns it into a Humean cause.[85] This eliminates both the organization of the entities and the activities that they engage in – *which make a mechanism out of a pile of parts*. The productive process that "is" the mechanism is black-boxed rather than explored, leaving the emergent properties that arise from the process unaddressed.

6. EQUIFINALITY AND THE CHOICE OF OBJECTS OF RESEARCH

Space prohibits a similar critique of game-theoretic treatments of "mechanisms." This and the following sections move on to illustrate the varying epistemic reaches of mechanismic and associational research, focusing on the three principal stages of mechanismic explanations: set-up conditions, termination conditions, and the processes that produce the latter from the former.[86]

Associational approaches view causality or explanation as a relation among variables. Therefore, the standard research strategy is to treat an outcome of interest as a dependent variable (effect) and then generate, organize, and analyze data to identify its "causes" and estimate their "causal effects."

From a mechanismic perspective, this gives insufficient attention to the choice of objects of scientific inquiry. One simply assumes that because something is of social importance or interest we can and should investigate "its causes" – and that those investigations need not be based on a plausible understanding of the operation of productive processes in the world.[87]

This strategy may work for objects that have one set of causes (or one process by which they are produced) – for example, cholera. But cancer is a very different thing. Although searching for "the causes" of "cancer" has generated some useful knowledge "of cancer," we now know that many

[81] Similarly, in Chapter 5, "Interactive Causal Mechanisms," Goertz focus on the models X_1 *AND* $X_2 \rightarrow Y$ and X_1 *OR* $X_2 \rightarrow Y$ – which obviously are models of associations, not productive processes.

[82] (Goertz 2017, 18, n. 6).

[83] "Snow is a major figure in the history of epidemiology, sometimes considered the founding father for his various innovative methods for studying the causal mechanism whereby cholera is transmitted." (Goertz 2017. Cf. 20, 21).

[84] (Goertz 2017, 21).

[85] This nicely illustrates the grip of a Humean conception of causation. If all scientific explanations involve associational causes, "causal mechanisms" *should* be reduced to associational causes (intervening variables).

[86] Cf. §1.3.

[87] Or, to the extent that knowledge of mechanisms is thus brought into the explanation, it is not simply a causal explanation. I briefly address multimodal explanations in ???.

different things, with many kinds of causes, "are cancers" – and that real progress in understanding (and combatting) "cancer" requires looking at different types of abnormalities, in different organs, produced in different ways, in different groups of people.

Most of the objects of investigation in International Studies are more like cancer, or even disease, than cholera. War, democratization, human rights violations, ethnic conflict, poverty, and inequality are abstract categories that aggregate varied (although in some ways similar) "things" that are produced in varied ways.

Equifinality – multiple paths from "the same" starting point to "the same" outcome – suggests that we investigate not termination conditions (dependent variables) but the mechanism(s) by which states of the world are produced. When there are multiple paths to an outcome, the particular path followed matters decisively *for the purposes of explanation*. To perhaps overstate the point, what makes a thing "the thing that it is" is less its "essential characteristics" than how it came to be what it is and how it operates in the world – especially when we are trying to understand the causes of outcomes (rather than the effects of those outcomes).

Associational approaches treat "an effect" as a single kind of thing. Mechanismic approaches treat "the same effect" produced by different mechanisms as different "things."

Equifinality at the very least requires rejecting claims for the epistemic priority of causal inference. Such claims, to the extent that they are not driven by a mistaken understanding of science or a prescriptive definition of causation, rest on bets that the world is relatively simple and homogeneous (allowing us to find relatively powerful and widely applicable lawlike regularities) and that we can learn most of what we need to know about "things" without knowing how they work. Those bets, however, have not paid off.

For example, decades of work on the associated causes of war have not produced much epistemic progress – not, I am suggesting, because we have found only correlates, rather than causes, but because war is not *a* thing (with many causes). Similarly, regression models continue to tell us little of interest about the causes of human rights violations – because human rights are not a "thing" (and respect for and violation of those diverse things come about in many varied ways).

Types or classes of things are not themselves things that have causes. The associated "causes" of type or class x – the "things" with which x is causally associated – do not cause x.

Selecting dependent variables without regard to processes of production usually is a poor way to organize research. And when a stream of associational research produces multiple "significant" findings that not only fail to provide a plausible account of how the world works but remain scattered insights that continue to resist integration (or even aggregation), we probably should step back and ask whether we are looking at the wrong things – because of a narrow prescriptive understanding of explanation (or causes).

At the very least, trying to identify and understand the operation of mechanisms is no less promising a research strategy than looking for causal effects. And which "works better" is an empirical question.

7. COMBINING MECHANISMIC AND ASSOCIATIONAL RESEARCH

This, of course, does not mean abandoning, let alone rejecting, associational research. It does, however, mean putting it back in its place, namely, as one way to learn certain things about some parts of the world.

When we have little knowledge of mechanisms – when we don't know, even in broad outline, "how" – just figuring out what is connected to what can be a significant epistemic achievement. (Associational approaches have particular promise for exploring what later I will call "causal thickets."[88]) The resulting knowledge, however, should be understood as fundamentally heuristic. And we should stop undervaluing, or even disparaging, heuristic knowledge. Knowledge that points the way to other knowledge can contribute greatly to scientific progress. Rather than generate not very revealing "explanations," we often would do better simply to contribute to the growth of knowledge about a topic.[89]

Consider the following methodological maxim: "Faced with an interesting regression model, ask what mechanism(s) might produce this result in the world." Given the difficulty of identifying mechanisms and unpacking how they work, associational suggestions about what to look for and where can be an immensely valuable epistemic contribution.

Or, because much more is involved in producing an outcome than any regression model can capture,[90] identifying outliers may be the greatest contribution of an associational analysis. (Personally, I usually find seeing what is way off the line in scatter-plots more informative than the coefficients of the explanatory variables.)

Associational approaches can also help to discipline the development of other kinds of explanations. Is what we know about (associated) causes consistent with a proposed explanation? If so, that counts in its favor. If not, we have identified potentially productive lines of inquiry for further developing our knowledge of causes or causation (or both).

This suggests extending the idea of triangulation beyond methods (and sources of data) to the questions we ask – in sharp contrast to KKV's claim that "the issue of triangulation … is not the use of different logics or methods …but the triangulation of diverse *data sources* trained on the same problem."[91] Similarly, B&C define triangulation as a "research procedure that employs empirical evidence derived from more than one method or from more than one type of data. Triangulation can strengthen the validity of both descriptive and causal inference."[92] And although triangulation is so central to Goertz that he organizes his book around a "research triad" of between-case causal inferences, within-case causal inferences, and causal mechanisms, he too focuses almost exclusively on causal inference.[93]

Such arguments are even plausible only if all scientific explanations have a single (causal inference) form – which, we have seen, they do not. A good understanding of the complex social world is most likely to be obtained by the concerted application of different kinds of explanations that in

88 See §???.

89 Extending the point further, the sharp distinction between description and explanation – and the prioritization of explanation – in mainstream approaches [cites] breaks down once we accept multiple types of explanation. Some statements are "descriptive" in one mode of explanation but "explanatory" in another. (For example, causal associations are "merely descriptive" in mechanismic explanations.) And it is at best unhelpful and misleading to sort all knowledge into these (or any other) two categories. (For example, heuristic knowledge is neither descriptive nor explanatory nor part way between the two. It is a qualitatively different kind of knowledge.)

90 We also know that adding more variables to the model usually undermines, rather than enhances, its explanatory power. See, for example, Christopher Achen's (2002, 446-447) Rule of Three and Philip Schrodt's (2014, 287-289, 294-295) discussion of the "deadly sins" of "greed" and "gluttony" in quantitative research.

91 (King, Keohane, and Verba 2010, 122).

92 (Seawright and Collier 2010, 356).

93 Cf. n. ???.

different ways get at different dimensions of "what is out there." In addition to multimethod research, we need explanatory pluralism in International Studies.

8. TREATMENTS, MECHANISMS, AND MULTIFINALITY

A mechanism begins operating at a particular point, with what MDC call set-up conditions. Where we should start a mechanismic explanation thus is determined by the mechanism (or sub-mechanism or stage) in question.

Associational research, by contrast, starts with the dependent variable, which is then linked to independent variables in various ways,

One standard strategy is to work back to causes – or, more precisely, jump back to independent variables. For example, regression modelers typically try out numerous possibilities to see "what works."

Another standard strategy is to identify a "treatment" variable and then try to estimate its causal effect (on a previously specified outcome). This, however, is unlikely to tell us much of explanatory value unless the chosen treatment happens to be central to a productive mechanism or sub-mechanism leading to the outcome. And such associational linkages are especially problematic because *multifinality* is pervasive in International Studies; "the same causes" have different effects depending on when and in what context they operate (in which productive process).

For example, the Goertz/Pevehouse example above assumes that there is a democratic IGO "mechanism" of democratization. The fact that one happens to be interested in the impact of democratic IGOs, however, does not necessarily justify organizing analysis around them. Mechanists would suggest instead studying IGOs in the context of processes that are known to facilitate democratization. The goal would be to try to determine which characteristic activities of democratic IGOs have what kinds of effects in which mechanisms (operating in which contexts).

When equifinality or multifinality is present – which in International Studies is usually – we rarely will be able to understand "an effect" simply by looking at "its causes." Similarly, we rarely will be able to determine what "a cause" causes without looking at *how* it causes. In International Studies there is little reason to believe that in general (or as a matter of principle) we can understand causation, let alone the world, by looking at the associations of independent and dependent variables.

Just looking for mechanisms, of course, does not guarantee that we will be able to link the "right" set-up conditions with the "right" termination conditions in the "right" way. It does, however, keep us focused on the search for patterns that result from the way the world works. And such patterns are at least as worthy of social scientific attention as patterns in our observational data about disconnected and context-free states of the world or values of variables.

9. THEORIES AND MECHANISMS

[At this point, the paper gets *very* rough.]

One might argue that many of the problems I have identified can be at least ameliorated by embedding causal explanations in good theories. Most obviously, a good theory should provide

useful guidance in selecting independent and dependent variables. And the central role of theory is not merely acknowledged but emphasized in mainstream discussions of research design.[94]

B&C, however, define a theory as "the conceptual and explanatory understandings that are an essential point of departure in conducting research, and that in turn are revised in light of research."[95] KKV similarly define a theory as "a reasoned and precise speculation about the answer to a research question, including a statement about why the proposed answer is correct."[96] Although in one sense admirably open, such definitions are hopelessly vague. And they make pretty much any reasonably rigorous explanation a theory.

As Kenneth Waltz famously put it, "theories explain laws"[97] (by which he meant regularities[98]). But causal associations, mechanisms, systems, fields, and game-theoretic models also explain regularities – making most social-scientific explanations "theories."

Behind such definitions, it seems to me, lurks a plausible and attractive "common sense" understanding of theories as *explanatory structures* with a certain epistemic heft. Their coherence, rigor, and scope aggregate and integrate bodies of knowledge and guide both further research and narrower explanatory endeavors, establishing, as B&C emphasize, a recursive interplay between the structure of our knowledge and the world we are seeking to understand. Understanding "theory" in this sense, I want to return to the contrast between associational and mechanismic approaches.

Theories are central to associational research *because* the research is (only) associational. Associations, by themselves, do not aggregate into explanatory structures. And their range of application is unclear. "Theory" provides coherence to the findings of associational research.

In mechanismic research, however, the mechanism organizes knowledge and guides research. We start with a whole that is then broken down into parts. Those wholes are real things in the world, not abstract epistemic structures, that are represented by models, which often takes a pictorial form that depicts the relations of the elements of the mechanism. Theories typically have no important explanatory role in mechanismic research (although theories may play a heuristic role and mechanisms may be parts of theories).

Glennan thus contrasts a "laws and theories" vision of science, which associational researchers have adopted – more accurately, they have adapted it to cover law-like regularities, producing a "variables and theories" vision – with a "mechanisms and models" approach.[99] In the next section, I will look in a bit more detail at models of mechanisms. The key point here it to, once more, underscore the fundamental epistemic diversity of the sciences – in this case by questioning the fetish for "theory" among much of the social-scientific mainstream in International Studies. Theories are central to associational research but not to mechanismic research.

The other point I want to make here is that the theories employed in associational research programs often (usually?) are not simply (associational) causal theories. Typically, they add mechanisms (or other processes), intentions, or relations, often in ad hoc ways, to produce what I

[94] ???

[95] (Seawright and Collier 2010, 354).

[96] (King, Keohane, and Verba 1994, 19).

[97] (Waltz 1979, 6).

[98] Cf. n. 23.

[99] (Glennan 2017, 7-8).

will call multimodal explanatory structures that help us to understand how the world is and how it works. [need quotes]

I am not criticizing multimodal explanations. Quite the contrary, I have argued that they are a good thing. My point here is to emphasize, again, in a different way, that associational explanations are but one way to carry out scientific research to generate one type of knowledge about some parts of the world.

10. MODELS OF MECHANISMS

[into paragraph or two on meaning of "models']

10.1. Schemas, Sketches, and Black, Grey, and Glass Boxes

Representations or "models" of mechanisms – understood in the broad generic sense of "a simplified or idealized description or conception of a particular system, situation, or process" or "a conceptual or mental representation of something"[100]– take several forms.

MDC identify a mechanism *schema*[101] as a simplified (abstracted) representation, often in the form of a diagram,[102] that leaves out "unimportant" features. Although a schema may be made more or less detailed – for example, in the photosynthesis diagram above we could look inside the granum – it aims to be fundamentally accurate. A mechanism schema claims both that these are the actual entities and activities that produce the outcome and that no "important" entities, activities, or organizational features are left out.

A *sketch* of a mechanism has "missing pieces, black boxes, which we do not yet know how to fill in;"[103] "missing components that are sought *as the search for the mechanism proceeds*."[104] This, it seems to me, is the best way to understand Snow.

Rather than a simple dichotomy, though, I think that we can profitably think of a continuum of models of mechanisms composed of what Darden calls glass boxes (in which the components are clearly specified and supported by strong evidence), black boxes (in which the mechanisms are largely unknown or highly speculative), and grey boxes (which contain incomplete or contested knowledge of the mechanism).[105] What we might call a maximal schema is composed entirely of glass boxes. (Consider photosynthesis.) A minimal sketch is composed entirely of black boxes. (This is one way to read the causal chain Goertz's Figure 2.2.)

Progress in understanding a mechanism thus can be understood as better identifying the parts and processes of the mechanism, reducing the number of black boxes in the model, and lightening the shades of its grey boxes.

100 *Oxford English Dictionary*.

101 MDC ???. "A mechanism *schema* is a truncated abstract description of a mechanism that we know how to fill with more specific descriptions of component entities and activities." (Darden 2013, 23).

102

103 (Machamer, Darden, and Craver 2000, 18).

104 (Darden 2013, 23 [emphasis added]). Cf. (Bechtel 2011, 537).

105 (Darden 2008, 966-967)

10.2. Perspectives and Causal Thickets

Even this, though, seems too demanding for many models of mechanisms in International Studies. They involve instead what William Wimsatt calls perspectives and causal thickets.

Perspectives are partial cuts into complex problems that require "the use of two or more perspectives for their solution."[106] "Sometimes problems appear to be big enough, or generally enough stated (e.g., the mind-body problem), that they seem to be intrinsically multi-perspectival."[107] Sometimes our knowledge supports only multiple, and sometimes even contradictory, perspectives.

Causal thickets arise when there are unclear or incomplete boundaries between perspectives – which Wimsatt suggests is the normal situation in the social sciences.[108] In causal thickets we have "an unusually large proportion of conceptual issues, methodological arguments, and boundary disputes. Some of these disputes are likely to indicate sources of genuine disagreement, but this can't be determined when so many things are up for grabs."[109]

Causal thickets are more the result of the complexity of the objects of investigation than of the immaturity of the science. As Wimsatt puts it "on a priori grounds, considering the possible connectivities of causal networks, shouldn't causal thickets be the norm?"[110] We inhabit a world in which there are "regularities at all levels, and mechanisms tying them together, and perspectives that give cross sectional cuts on the phenomena for a range of problems. And then there are some things that are just too multiply-connected" to fit neatly defined models.[111] (Physicists have it easy, being able to deal with relatively simple and isolated entities and activities.)

What I suggest calling *pathways into causal thickets* are suggested lines of inquiry that do not (yet?) involve plausible sketches of mechanisms or well-elaborated perspectives.[112] They do, however, point in the direction of a mechanism. Goertz's models of democratic transitions might be seen as pathways into the causal thicket of democratic transitions or even sketches of perspectives on particular types of transition mechanisms.

10.3. Complexity Limitations on Mechanismic Explanations

[this subsection seems to me especially weak]

The complex of social phenomena, however, pose serious problems for mechanismic explanations. Faced with long or complex causal chains, it usually will be asking too much to try to work through the entire mechanism. And if our model includes black and grey boxes over which there is considerable contention, as is common in International Studies, we *cannot* work through the whole process.

Successful mechanismic explanations typically exceed our epistemic reach. And to the extent that our knowledge of process is incomplete, those black or grey boxes become a lot like intervening variables.

106 (Wimsatt 2007, 237).

107 (Wimsatt 2007, 238).

108 (Wimsatt 2007, 239).

109 (Wimsatt 2007, 239).

110 (Wimsatt 2007, 240).

111 (Wimsatt 2007, 240).

112 Pathways point to "where" we might find mechanisms. They are not mechanisms. Unfortunately, Goertz does not respect this distinction. ??? Cf. also [Mahoney???], ???

This, I want to argue, suggests the practical need for multimodal explanations; for research strategies that combine not just different methods but different kinds of questions – which often can partly compensate for (and be of heuristic value to) one another.

I also want to suggest, though, that causal thickets, perspectives, and glass, grey, and black boxes offer a promising framework for thinking about progressive multimodal social-scientific research. The epistemic ideal is a mechanism schema with all glass boxes. heuristic application of associational or game-theoretic approaches, however, may move us closer to that goal than t incomplete and speculative processual explanations. And there may be reason to hope that the concerted and recursive application of different modes of research can sharpen and strengthen one another, move us, over time, closer to our epistemic ideal.

Below I will look briefly at two examples of the convergence of mechanismic and other modes of explanation. First, though, I want to look at addressing problems of complexity through modularization and hierarchy.

11. MODULARITY AND LEVELS OF ORGANIZATION

At least since Herbert Simon's pathbreaking work on complexity, it has been widely thought that most social mechanisms are to some degree decomposable.[113] And to the extent that this is true, disassembling a mechanism into sub-mechanisms or stages may allow us to advance our understanding by working on smaller more manageable pieces.

It is an empirical question, though, what the modules or stages are. The problems of choosing an object of inquiry in mechanismic research thus may be more difficult than my initial discussion above suggested. The more incomplete our knowledge, the more difficult it likely is to be to cut the world "at its joints."

Nonetheless, the project of mechanismic explanation keeps this necessity front and center. And much of the progress in our modelling will involve better identifying the mechanism, its component parts, and how they in fact operate.

Mechanisms also have a hierarchical organization that gives rise to another kind of modularity. In biological systems, for example, physical and chemical mechanisms continue to operate but are not an immediate object of investigation by most biologists most of the time (although biochemists do operate at the intersection of chemical and biological systems). Similarly, in social systems "lower-levels" mechanisms continue to operate but usually are not a central focus of attention of most social scientists (although social psychology is an important branch of research at the intersection of psychological and social systems). Different disciplines – and different approaches within a discipline – will identify and investigate different kinds of mechanisms even when faced with "the same" thing.

[need a much better transition to the next two paragraphs – both of which need to be better developed]

The semi-objective character of levels of organization – as Wimsatt puts it, ??? – provides a powerful argument against the sort of doctrinaire individualism often encountered in game-theoretic approaches. ??? This is little more than an arbitrary philosophical or methodological imposition,

[113] Simon cite. An assumption of what Simon calls near-decomposibility is probably too strong. A certain degree of modularity, however, does seem to be the norm.

based on no empirical evidence; an insistence that reality "is" (or ought to be approached) in a particular way.

Much the same is true of claims that social science must identify and work back to "micro-foundations," understood as preferences, dispositions, or other mental states of individual human beings. Whether such mental states are indeed foundational or in fact consequences of social conditions is in part an empirical question and in part a matter of our choice as to the level of organization being investigated. They are no more (or less) "foundational" that the biological, chemical, and physical "foundations" of social life – which social scientists rightly ignore most of the time. And in *social* science there is a pretty obvious reason why, prima facie, we should not prioritize atomistic individuals (even if we take seriously the as if stories often told about their preferences).

[awkward transition – at best]

To repeat, if we are going to investigate how the world works, we need to look, empirically, at how the world is in fact organized and how it works – at the mechanisms that in fact operate in the world.

12. Bringing Associational and Mechanismic Explanations Together: DAGs

[This section is too much of a mess to let it out into the world. The basic idea is that DAGs go a long way towards "mechanizing" associational analysis. By closing alternative causal pathways (controlling for confounders) we are likely to get something like a mechanism sketch, that can then be explored mechanismically – and then that knowledge can be fed back into our statistical modelling …]

13. Mechanisms, Case Studies, and Process Tracing

[This section also is too much of a mess to be presented here. The basic idea is something like the following. "Process tracing" in the mainstream methodology literature is a way of getting additional observations or evaluating (associational) causal explanations. More loosely, "process tracing" is sometimes understood as an unsystematic search for additional variables, causal forces, or information. I want to argue for literal process tracing – that is, for the virtues of both mechanismic and narrative explanations. In the case of mechanisms, we should look for regular or generalizable processes in our process tracing.]

14. Conclusion

[I haven't even tried to write this yet]

References

Achen, Christopher H. 2002. "Toward a New Political Methodology: Microfundations and ART." *Annual Review of Political Science* 5:423-450.

Bechtel, William. 2011. "Mechanism and Biological Explanation." *Philosophy of Science* 78 (4):533-557.

Bechtel, William. 2016. "Mechanists Must be Holists Too! Perspectives from Circadian Biology." *Journal of the History of Biology* 49 (4):705-731.

Bechtel, William. 2017. "Explicating Top-Down Causation Using Networks and Dynamics." *Philosophy of Science* 84 (2):253-274.

Bechtel, William and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36:421-441.

Brady, Henry E. 2010. "Doing Good and Doing Better: How Far Does the Quantitative Template Get Us." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Brady, Henry E. and David Collier, eds. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. 2nd ed. Lanham: Rowman & Littlefield.

Bunge, Mario. 1997. "Mechanism and Explanation." *Philosophy of the Social Sciences* 27 (4):410-465.

Collier, David, Henry E. Brady, and Jason Seawright. 2010a. "Critiques, Responses, and Trade-Offs: Drawing Together the Debate." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Collier, David, Henry E. Brady, and Jason Seawright. 2010b. "Introduction to Part II." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Collier, David, Henry E. Brady, and Jason Seawright. 2010c. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Collier, David, Jason Seawright, and Gerardo L Munck. 2010. "The Quest for Standards: King, Keohane, and Verba's *Designing Social Inquiry*." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Craver, Carl F. and David M. Kaplan. 2011. "Towards a Mechanistic Philosophy of Neuroscience." In *The Continuum Companion to the Philosophy of Science*, edited by Steven French and Juha Saattsi. London: Continuum.

Darden, Lindley. 2008. "Thinking Again about Biological Mechanisms." *Philosophy of Science* 75 (5):958-969.

Darden, Lindley. 2013. "Mechanisms Versus Causes in Biology and Medicine." In *Mechanism and Causality in Biology and Economics*, edited by Ksiang-Ke Chao, Szu-Ting Chen and Roberta L. Millstein. Dordrecht: Springer.

Elster, Jon. 1983. *Explaining Technical Change: A Case Study in the Philosophy of Science*. Cambridge: Cambridge University Press.

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

Falleti, Tulia G. and Julia F. Lynch. 2009. "Context and Causal Mechanisms in Political Analysis." *Comparative Political Studies* 42 (9):1143-1166.

Giere, Ronald N. 1999. *Science Without Laws*. Chicago: University of Chicago Press. sugan

Glennan, Stuart. 2017. *The New Mechanical Philosophy*. Oxford: Oxford University Oress.

Glennan, Stuart and Phyllis Illari. 2017. *The Routledge Handbook of Mechanisms and Mechanical Philosophy*: Routledge.

Goertz, Gary. 2017. *Multimethod Research, Causal Mechanisms, and Case Studies: An Integrated Approach*. Princeton: Princeton University Press.

Goldthorpe, John H. 2001. "Causation, Statistics, and Sociology." *European Sociological Review* 17 (1):1-20. 10.1093/esr/17.1.1

Heckman, James J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35 (1):1-97.

Hedström, Peter and Petri Ylikoski. 2010. "Causal Mechanisms in the Social Sciences." *Annual Review of Sociology* 36:49-76.

Hollis, Martin and Steve Smith. 1990. *Explaining and Understanding in International Relations*. Oxford: Clarendon Press.

Illari, Phyllis McKay and Jon Williamson. 2012. "What is a Mechanism? Thinking about Mechanisms Across the Sciences." *European Journal for Philosophy of Science* 2 (1):119-135.

Imbens, Guido W. and Donald B. Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge: Cambridge University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 2010. "The Importance of Research Design." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Machamer, Peter. 2004. "Activities and causation: The metaphysics and epistemology of mechanisms." *International studies in the philosophy of science* 18 (1):27-39.

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

Mahoney, James. 2001. "Review Essay: Beyond Correlational Analysis: Recent Innovations in Theory and Method." *Sociological Forum* 16 (3):575-593.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14 (3):227-249.

Mayntz, Renate. 2004. "Mechanisms in the Analysis of Social Macro-Phenomena." *Philosophy of the Social Sciences* 34 (2):237-259.

Morgan, Stephen L. and Christopher Winship. 2015 [2007}. *Counterfactuals and Causal Inference*. 2nd ed. Cambridge: Cambridge University Press.

Paine, Jack. 2016. "Set-Theoretic Comparative Methods: Less Distinctive than Claimed." *Comparative Political Studies* 49 (6):703-741.

Schrodt, Philip A. 2014. "Seven Deadly Sins of Contemporary Quantittive Analysis." *Journal of Peace Research* 51 (2):287-300.

Seawright, Jason and David Collier. 2010. "Glossary." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier. Lanham: Rowman & Littlefield.

Steel, Daniel. 2008. *Across the Boundaries: Extrapolation in Biology and Social Science*. Oxford: Oxford University Press.

Suganami, H. 2008. "Narrative Explanation and International Relations: Back to Basics." *Millennium - Journal of International Studies* 37 (2):327-356.

Waltz, Kenneth N. 1979. *Theory of International Politics*. New York: Random House.

Wimsatt, William C. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge: Harvard University Press.