I

Introduction

Theories of Institutions

1.1 justification for the book

The human condition teems with institutions. “To be a member of a community or society,” Jack Knight argues, “is to live within” them.1 Yet scholarly interest in this unavoidable aspect of our lives, which dates at least as far back as Aristotle’s study of politeia,2 ebbs and flows. The growth and accumulation of knowledge often proceed by punctuated periods of differentiation and consolidation. Spurts of rapid growth feature the opening up of new problems, puzzles, and approaches, often elaborated and researched at great length with little concern for how different parts fit together. These periods of growth often give way to consolidation and synthesis. The last four decades have witnessed a veritable explosion across the social sciences of work subscribing to “institutional theory” or operating under cognate monikers such as “(new) institutionalism,” showing that institutions matter across wide swaths of social life. Yet, there is surprisingly little consensus about key questions concerning origins, persistence, and how and when institutions change. More problematically, less conversation takes place across the many involved fields than we would expect and hope for on these important issues. In reading several decades of literature, we even find some indication of different approaches operating in increasingly separate orbits. This profile – a huge, maturing, interdisciplinary literature driving in different directions – cries out for an overview, critique, and synthesis that is broadly accessible and invites engagement across disciplines, subfields, and levels of expertise. We hope that Theories of Institutions advances these aims.

Our method involves identification of core emphases within different institutional traditions, critical analysis of important and representative work, and

1 Knight 1992, 1.  2 Barkanov 2013.
Introduction: Theories of Institutions

synthesis of existing theories grouped around key variables with disciplinary connotations: temporality (history), sociality (sociology), (in)efficiency (economics), and power (political science). We focus especially on different approaches to answering common questions surrounding the origins, persistence, and change of institutions. We hope to speak to advanced undergraduate and graduate students, teachers and scholars working in the various substantive fields of our home discipline of political science (including American, comparative, international politics, public policy, and political theory), in the allied disciplines of history, sociology, and economics, and in related social scientific, policy, and professional fields such as organizational studies.

1.2 WHAT ARE INSTITUTIONS?

1.2.1 Motivation

It is conventional to decry the many definitions of institutions within and across the social sciences. Heclo characterizes the search for definitional consistency among theories of institutions as an exercise in frustration, contemplating no fewer than twenty-one offerings. A decade-and-a-half on, we have gathered about eighty definitions, which appear in Appendix 1 and which the word cloud in Figure 1.1 summarizes. Recent interventions in organization studies urge us to abandon the term, alleging that it has become little more than a slippery branding device.

More sympathetically, Blondel has argued that appropriate definitions are problem-, context-, and discipline-dependent, and that pursuit of a single definition will never satisfy every scholar or serve every use. We agree. Institutions are so multifaceted, and the real-world problems and explanatory concerns with respect to which they arise are so vast, that any search for a consensus definition proves at best quixotic and at worst counterproductive. With that said, tolerating definitional proliferation provides the ancillary benefit of freeing us up to offer our own understanding without feeling that we are diminishing the valuable work of others or pigeonholing our readers. This also allows us to avoid falling into the trap whereby “social scientists relate to the theoretical concept of institutions as ordinary people relate to some established institutions: they take the meaning for granted and proceed to make use of it.”

---

5 We do not imply that these are the only emphases of the four disciplines.
1.2 What Are Institutions?

1.2.1.1 Our Definition

In defining institutions we set a standard of fruitfulness given the tasks that we identify in this book: to capture patterns of distinctiveness of the various theoretical traditions we engage, to identify their strengths and weaknesses, to locate blind spots and fertile terrain, and to highlight promising avenues for more integrated institutional theory and empirical research. In this spirit, we define institutions as *intertemporal social arrangements that shape human relations in support of particular values*. Let us unpack this definition.

“*Intertemporal.*” However we slice and package time, institutions bridge consecutive periods. They bear the imprint of their birth and life history, even as they move into and through subsequent moments. Grafstein conceives of them as aggregates of “time slices” of individuals, which bundle together human elements from a single point and then proceed to move forward in time, like a “lengthened [human] shadow.”

By contrast, much noninstitutional human interaction resembles a spot transaction: episodic, limited in duration, atomized. We bump into someone on the street, we speak with a customer.

---

8 Grafstein 1988, 590; Emerson 1841.
service representative, we swap glass beads for dried fish: these are all events, transactions, interactions, or mere “happenings.”9 They generally occur, as it happens, within some institutional fabric, where for example protagonists can meaningfully seek and grant forgiveness, start with expectations about “good service,” share a sense of fair exchange, or try not to interfere with the participants. Institutions may surround the happenings, and indeed the events typically embed themselves in larger institutional, social, cultural, economic, historical, and other contexts. But the actions themselves are ad hoc and – one hopes in some of these examples – do not extend beyond a single period. Institutions, conversely, exist diachronically – in and through time. This feature stands as one root of our conception of institutions.

“A social.” A second root idea is that institutions stand as social facts in Durkheim’s (1938) sense. We note, first, that humans are necessary but not sufficient conditions for social facts. A physical arrangement of bricks and mortar once known as a church may outlive the Rapture, but religion, and hence a “church,” cannot. Second, institutions are artifactual: humans conjure them. Searle puts institutional reality on a linguistic foundation, and expresses its generative aspect beautifully in his 2009 book Making the Social World. Language gives us “the capacity to create a new reality by representing that reality as existing. We create private property, money, marriage, and a thousand other phenomena” in this way.10 In declaring “I promise,” we call forth an institution that did not and could not exist in the absence of such a speech act. Third, institutions, as we conceive them, are collective. We emphasize the word “social” rather than “human” because we want to rule out purely individual beliefs, values, neuroses, habits, and other attributes, even if many individuals hold, embody, or enact them.11 “Private” property is an inherently social construct since its notions of inclusion and exclusion exist only relationally. Institutions must be intersubjectively shared, even while we might hold private beliefs about them.12 Fourth, institutions are immaterial, living between the ears and existing as discrete ideas, inarticulate premises, or general attitudes in people’s minds. Institutions may and often do find material expression, but they must exist invisibly in the first instance. Fifth, a social fact exists “in its own right independent of individual manifestations.”13 Numerous individuals may prefer eating meat or vegetables, or drinking or not drinking alcoholic beverages. When this condition occurs only as the sum of individual tastes, we do not have a social fact but an aggregate of microlevel facts.14 Note that commonality, the simple sharing of attributes, does not imply sociality – it can be arrived at through nonsocial processes, as with an evolved physical feature such as the nose. Noses are “shared” in the sense that pretty much

1.2 What Are Institutions?

everybody has one, as *The Nose Book* reminds us. The trait of having a nose correlates highly across individuals, but it is not a “social fact” in Durkheim’s sense. A nose does not depend for existence on any mental representation, let alone a common one. In general, traits arrived at independently of some interaction between people, however indirect and mediated, are not social, while those constructed by networked minds are.

“ARRANGEMENTS.” This word implies multiplicity – more than one element stands to be arrayed – as well as unity. The motto of the American Republic, *E Pluribus Unum* (“out of many, one”), is an essentially institutional construction. It also connotes pattern, regularity, (relative) invariance, or order. Institutions reduce entropy in the world, at least as an initial proposition, though they may unleash new energy or become, themselves, more entropic over time. They are conservative. Relative to an imaginary institution-free baseline, they organize, structure, and channel human thoughts and activities. To institute something, from the Latin *instituere*, means “to set up” or “to establish,” words that connote some kind of patterning. A “random arrangement,” in our sense of the word, would be a contradiction in terms.

We are agnostic as to what is being structured and ordered, be it tastes, expectations, meanings, identities, time, power, policies, practices, property rights, transactions, rules of the road, or any number of other things of interest to social scientists. We also hold no brief for any sources or manifestations of order. Arrangements can be grounded in charismatic, traditional, or legal authority, can prescribe or proscribe behavior, can be supported by coercion or consent, and can vary in any other substantive way. But any sense of institutions must connote that they organize the social world at some level. They organize in particular ways. Institutions define the range of choices, sometimes forcefully, but more typically by highlighting particular paths, providing salience to some routes while decentering others, and suggesting what is focal and what is not. Institutions perform an enormously important cognitive service by preventing us from having to think of everything at once. Whitehead observed that “Civilization advances by extending the number of important operations which we can perform without thinking about them.” If this is true, then institutionalization stands at the heart of any *mission civilisatrice*.

“SHAPE HUMAN RELATIONS.” Sharp-eyed readers will recognize elements from economic historian Douglass North’s classic definition of institutions as “humanly devised constraints that shape human interaction” in our own conception. We intend “social” to cover the “humanly devised” part of North’s definition, further noting that “devised” implies a purposiveness or intentionality that might serve North’s aims but proves too limiting for our

---

2. Schotter 1981, 139.  
15. The Nose Book reminds us.  
17. Schotter 1981, 139.  
broader purposes. Similarly, North’s notion of institutions as “constraints” captures only one face of the coin of institutional consequentiality, the other being enablement.\textsuperscript{23} We have toyed with the idea that institutions always involve some constraint, but this idea falls afoul of purely enabling rules such as that in capitalist economies “labor may sell itself to the highest bidder.” Such a rule might squeeze a capitalist, but this is an outcome conditioned on extra-institutional factors (such as competition for scarce labor) and not embodied in the rule itself. In the end, we agree with Searle that to see “institutional facts as essentially constraining . . . is a very big mistake.”\textsuperscript{24} We intend the verb “shape” to mean that institutions make a difference in the world. While this smuggles some causality into our definition, a putative institution without any effect at all would not be much of one. A mere behavioral regularity is not an institution if it derives from purely individual sources. We substitute “relations” for North’s “interaction” for the same reason that we eschewed his term “devised”: it tilts a bit too far toward behavior and agency than is ideal for our purposes. That said, it might be too limiting in that it rules out individual effects of the kind that stand at the center of social psychology. One might substitute the phrase “the human condition” to capture such phenomena, but we retain relational language because it seems more central to our social concerns.

“In support of particular values.” Inclusion of values as part of the definition of institutions may be controversial. For example, what do we make of pure coordination rules? Nevertheless, Stinchcombe (1968) has persuaded us that institutions generally operate non-neutrally with respect to values.\textsuperscript{25} The values we imagine as supported by institutions may include neutrality (e.g., a neutral civil service or lottery draft system) but that is simply a way of saying that in some circumstances, neutrality is a value. Our point is that institutions are never valueless, value-free, or value-innocent and that this property arises not empirically but is baked into what they are.

1.2.2 Other Semantic and Ontological Comments

Let us briefly discuss how we see institutions relating to other terms which frequently arise in the literature. First, we use the term interchangeably with rules\textsuperscript{26} and for stylistic purposes may substitute other terms as well. Second, we treat organizations “as a special subset of institutions, involving deliberate coordination,”\textsuperscript{27} though for presentational purposes we sometimes interchange

\textsuperscript{23} For a good discussion, see Clemens and Cook 1999, 445–446.

\textsuperscript{24} Searle 2009, 105.

\textsuperscript{25} See also Schattschneider 1960, 71, portraying organizations – a kind of institution in our conception, though not in others’ (see Section 1.2.2 at footnote 28) – as embodying the “mobilization of bias.”

\textsuperscript{26} As do Ostrom 1986; March, Schulz, and Zhou 2000; and others. Indeed, Figure 1.1 illustrates the centrality of this institutions-as-rules conception.

\textsuperscript{27} Hodgson 1998, 180.
1.2 What Are Institutions?

the terms, as is often done in common usage. Thus, in our own empirical work, we refer to organizations such as the European Commission, Council, Parliament and Court of Justice as European Union institutions. This approach is quite distinct from the one taken by most organizational theorists, who treat organizations and institutions as ontologically distinct and, indeed, causally related. On that view, “institutions are sets of norms which apply across a variety of specific organizations. Organizations are structures of social relationship, social actors arranged in positions and roles; usually, but not always, deliberately arranged and designed to achieve some end. Institutions provide normative environments shaping the activities of organizations.”

We recognize the depth of this tradition, going back at least to Meyer and Rowan (1977), which sees organizations as ritually isomorphic with their broader institutional environments. But this is a case where we think flexibility is possible, and common usage does not interfere too much with the greater precision that certain treatments may demand.

Third, we need to acknowledge the vast literature on “norms,” which are sometimes treated as distinct from institutions, sometimes as synonymous with them, and sometimes as instances of them. We hold the last understanding, treating norms quite simply as informal institutions, that is, those that are not written down. Writtenness has very special qualities in the post-Neolithic human condition, casting what Hannah Arendt called a “sacred halo” around inscribed rules. Searle postulates that writtenness confers stability and endurance on status functions.

Of more practical concern for institutional analysts, formal rules can be observed empirically. While the literature has thus privileged them, a growing body of work examines informal institutions. We see this especially in comparative politics but also in a range of other fields. As we develop further in Chapter 6, there are very good reasons to take informal institutions, as such, seriously.

One final ontological point suggests itself. Because we embrace Searle’s view that language is the fundamental institution which generates all of the countless others, and we recognize the ubiquity of the Westphalian state over the earth’s landmass (excepting Antarctica), we hold that in the human condition, setting aside the proverbial child raised by wolves, there can be no institutional tabula rasa. Indeed, our reading and analysis lay bare a world that is profusely institutionalized, and our Searlian reading of the linguistic underpinnings of

---

22 In comparative politics, see, e.g., Stacey and Rittberger 2003; Helmke and Levitsky 2004; Tsai 2006; Farrell and Héritier 2007; Darden 2008; Grzymala-Busse 2010. Beyond it, see, e.g., Chacar, Celo, and Hesterly 2018.
institutions implies a boundless capacity for more. While we suspect that no reader of this book will wish to dispute this argument, drawing out its implications spotlights a promising research frontier, as we discuss in Chapter 6.

1.3 THEORIES OF INSTITUTIONS

1.3.1 Metatheory

In Section 1.2 we elaborated a concept of institutions embodying qualitative diversity, asserted that the quantity of existing institutions is vast, and argued that the possibilities for elaborating new ones are limitless. It thus should come as no surprise that there is no single theory of institutions, but that there are many distinct (though generally not mutually exclusive) theories – hence the book’s title. We do not intend by this only to study institutions as outcomes, though the phrase can be read to imply this. A longer and slightly more accurate title might be: “theories of institutions and their effects.” The first part of such a phrase does indeed engage institutions as outcomes, recognizing their artifactual nature and asking how and why they originate, endure, and develop or die. Here, institutions stand as dependent variables. At the same time, our notion of a richly populated institutional status quo implies that rules are always already present, at least as background conditions, and very commonly, in social scientific theories, as independent variables.\(^{35}\) In causal language, all statements placing institutions on either the left (effect) or right (cause) side of an explanatory equation stand as theories of institutions as we understand them.

However, there are clearly cases where it makes sense to think of the constitutive components of institutions.\(^{36}\) A parliament exists by virtue of the fact that it legislates, and a political party by the fact that it aggregates interests in pursuit of public office. The medical profession exists and can be recognized by the elaborate rules and procedures for practicing medicine and enforcing standards of practice for doctors. The game of baseball can be recognized by multiple facts, including that a batter continues to bat until she becomes an out, gets on base, or is removed from the game. Three strikes and the batter is out, four balls and she is entitled to first base. There is no limit on the number of foul balls that can be hit, unless one is caught before touching the ground or results from a two-strike bunt attempt. And so on. At the international level, a state exists not so much in virtue of its ability to satisfy functional or empirical requirements such as control over borders, maintenance of order, or provision of public goods, but rather in virtue of its admission to the club by other states, that is,

\(^{35}\) Field 1979, 1981.

1.3 Theories of Institutions

recognition. The operation of the institution of international legal recognition does not cause a state – it constitutes one. Such accounts, too, are theories of institutions. In general, two considerations guide our strategy in identifying, organizing, and analyzing theories of institutions. First, we want to be clear about what is constitutive and what is contingent. Second, we choose a strategy of minimalism, that is, we attempt to keep constitutive characteristics to a minimum, so that we can entertain, theoretically and empirically, a large number of interesting hypotheses. Complex definitions cannibalize potentially interesting causal relations. The number of hydrogen and oxygen atoms should be part of the definition of water, as well as perhaps its liquid state, but how fast water heats, what will float on it or sink in it, and whether water expands and contracts when subject to temperature changes should be free parameters. Similarly, when defining an institution we should remind ourselves that we are only trying to identify a relatively homogenous group of things, and for this not too many properties may be needed. Every time we add a property definer we remove a property that might be useful for explanation, at least within a certain range.

1.3.2 The “New Institutionalisms”

From the 1970s forward, a kaleidoscopic array of “new institutionalisms” emerged across the social sciences. One durable organizing logic postulates a triptych of rational choice, sociological, and historical variants. It was not until we began this book that we saw how centered on political science this organizing logic was. Rational choice institutionalism (RCI), emerging from the study of political economics as instituted at Cal Tech, Rochester, and Washington University, began by importing microeconomic assumptions and formal methodologies into the study of American political institutions and, eventually, comparative and international ones. While it eventually drew on actual “new institutionalists” in economics, such as Oliver Williamson and Douglass North, the resulting RCI identified political scientists interested in and drawing from economics, rather than scholars working in a truly interdisciplinary field. Similarly (i.e., as seen from political science), sociological institutionalism (SI) derived from the work of James March and Johan Olsen, the latter a political scientist and the former an influential representative of a strand of institutional theory in organizational studies/sociology (see Chapter 3).

Jackson and Rosberg 1982; Jackson 1990.
Goertz (2006), recommended by a reviewer, has enriched our understanding of how to proceed here.
Hall and Taylor 1996; see also Immergut 1998; Hall 2010; Peters 2011; Lowndes and Roberts 2013; Farrell 2018.
For within-camp surveys, see Weingast 2002 and Shepsle 2006.
The “new institutionalism in organizational analysis,” growing especially out of the work of sociologist John Meyer, has stood at the heart of that discipline’s debates about institutions. Yet, despite its label, this “SI,” the one that indigenous theorists might think of first, found minimal recognition by political science institutionalists before it began making inroads into international relations (IR), where it is still widely seen as a strand of the much broader constructivist school. Historical institutionalism (HI), finally, had little to do with the discipline of history, which mostly lacks an indigenous institution-theoretic tradition. Instead, it grouped historically inclined political scientists (and a few sociologists) who shared affinities for cross-national comparison, qualitative methods, the study of power, the possibility of endogenous preferences, and some degree of antipathy toward rational choice work. This overlaps substantially with a school of interdisciplinary scholars engaging in what they call comparative historical analysis. A number of excellent within-camp surveys of HI have been produced in the last two decades.

In the meantime, economists and sociologists were elaborating their own “new” and “neo” institutionalisms. Comparison of the old and new institutionalisms has formed a staple of the literature ever since, and is not a topic that we aim to revisit. To get a 30,000-foot vantage on disciplinary trends in institutional theory, Figure 1.2 displays developments in the use of the term “institutionalism” (top panel) and the phrase “institutional theory” (bottom panel) in articles in the JSTOR depository of social science journal articles from 1980–2020.

Three observations about these data suggest themselves. First, sociology emerges as the most “institutional” of the social sciences and history consistently the least so, while political science and economics have tracked each other in an intermediate position. Second, term “institutionalism” has generally been more popular than the phrase “institutional theory,” especially in political

45 Rowlinson and Hassard 2014. Even references to “new institutional history” (Spruyt 1994) end up citing political scientists and economists.
46 Steinmo, Thelen, and Longstreth 1992. HI was thus very much of its moment, opposing a rational choice movement that seemed ascendant in the subdiscipline of comparative politics, at the expense of more traditional work grounded in area studies. See Bates 1997 and Monroe 2005 and, on “economic imperialism,” Radnitzky and Bernholz 1987.
47 Mahoney and Rueschemeyer 2003.
48 Thelen 1999; Pierson and Skocpol 2002; Sanders 2006; Steinmo 2008; Fioretos, Falleti, and Sheingate 2016b.