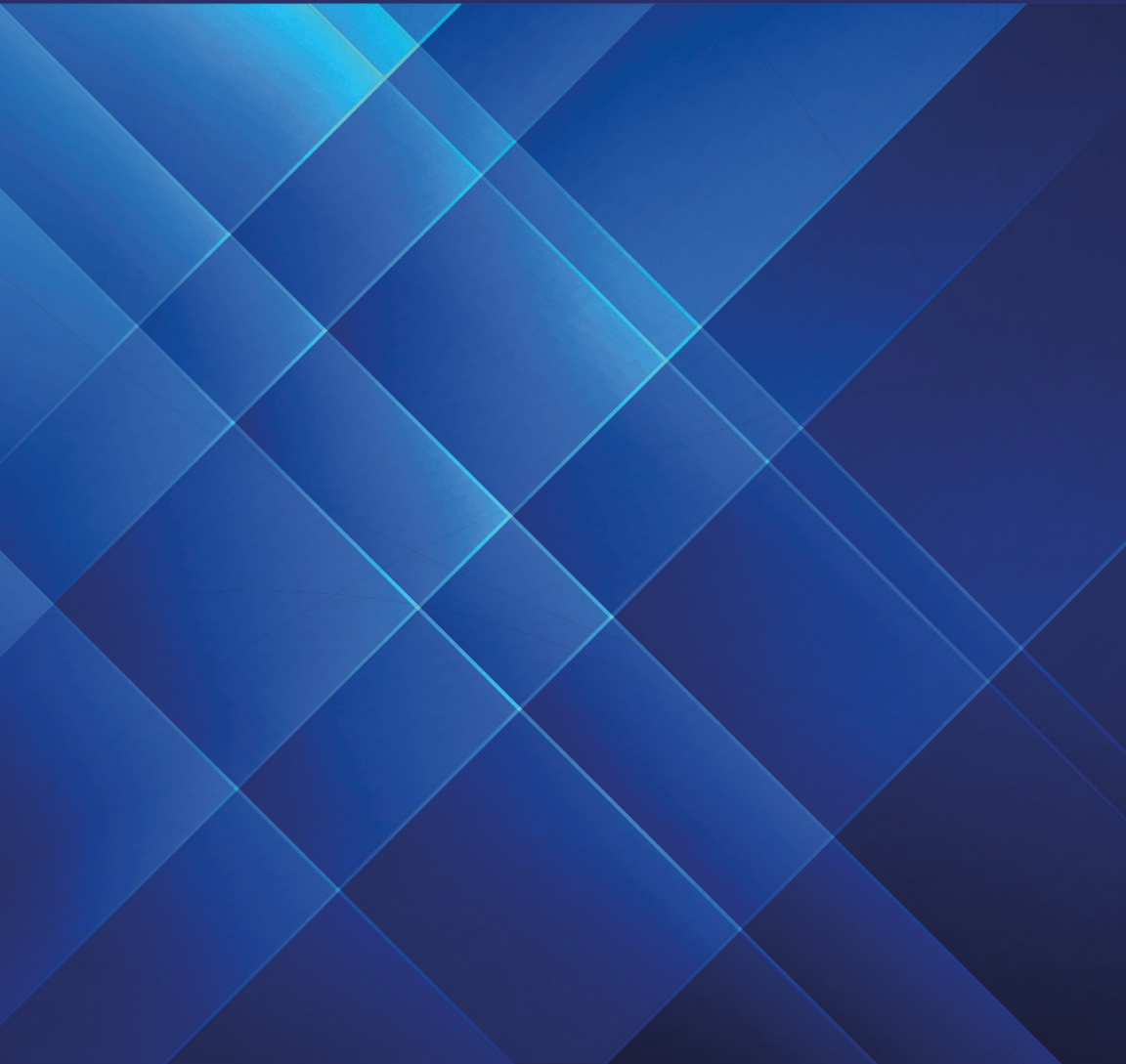# METHODS OF THE POLICY PROCESS

Edited by CHRISTOPHER M. WEIBLE
and SAMUEL WORKMAN

"More than a supplemental guide to *Theories of the Policy Process*: The book offers a fantastic overview and inspiring insights into gathering and analyzing data about policy processes. Thereby, the book not only provides the necessary tools for empirically-grounded research, but also for advancing theories of the policy process."

Florence Metz, *University of Twente, Netherlands*

"*Methods of the Policy Process* articulates best practices for applying policy process theories while emphasizing flexibility and creativity. It is an essential resource for scholars seeking to design and conduct rigorous, theory-driven research that better illuminates the complex dynamics of modern policymaking."

Elizabeth A. Koebele, *University of Nevada, USA*

"An eclectic methodological roadmap for new and experienced scholars. The volume edited by Christopher Weible and Samuel Workman presents a wide variety of research strategies, from different theoretical traditions in policy studies. It is an essential book that advances the discussion about methods and offers a major contribution to the global community of public policy academics."

Osmany Porto de Oliveira, *Federal University of São Paulo, Brazil*

Methods of the Policy Process

The increasingly global study of policy processes faces challenges with scholars applying theories in radically different national and cultural contexts. Questions frequently arise about how to conduct policy process research comparatively and among this global community of scholars. *Methods of the Policy Process* is the first book to remedy this situation, not by establishing an orthodoxy or imposing upon the policy process community a rigid way of conducting research but, instead, by allowing the leading researchers in the different theoretical traditions a space to share the means by which they put their research into action.

This edited volume serves as a companion volume and supplemental guide to the well-established *Theories of the Policy Process, 4th Edition*. *Methods of the Policy Process* acknowledges that growth and advancement in the study of the policy process is dependent not merely on conceptual and theoretical development, but also on developing and systematizing better methodological approaches to measurement and analysis. To maximize student engagement with the material, each chapter follows a similar framework: introduction of a given theory of the policy process, application of that theory (including best practices for research design, conceptualization, major data sources, data collection, and methodological approaches), critical assessment, future directions, and often online resources (including datasets, survey instruments, and interview and coding protocols). While the structure and focus of each chapter varies slightly according to the theoretical tradition being discussed, each chapter's central aim is to prepare readers to confidently undertake common methodological strategies themselves.

Methods of the Policy Process is especially beneficial to people new to the field, including students enrolled in policy process courses, as well as those without access to formal training. For scholars experienced in applying theories, this edited volume is a helpful reference to clarify best practices in research methods.

Christopher M. Weible is Professor in the School of Public Affairs at the University of Colorado Denver. His research focuses on conflict and collaboration in policy processes. He is the Editor of the *Theories of the Policy Process*, Co-Editor of *Policy & Politics*, and Co-Director of the Center for Policy and Democracy.

Samuel Workman is Professor of Political Science and Director of the Institute for Policy Research and Public Affairs in the Rockefeller School of Policy and Politics at West Virginia University. His work is focused on public policy, agenda setting, the bureaucracy, regulatory policy, and research methodology. He is Principal Investigator on the Education Regulations Project (EdRegs), funded by the National Science Foundation (NSF), and the author of *The Dynamics of Bureaucracy in the U.S. Government*. His work has appeared in the *Policy Studies Journal*, the *Journal of Public Administration Research and Theory*, *Policy and Politics*, *Interest Groups and Advocacy*, and *Towards Data Science*, among others. He is a founding Editor of *3Streams* with William Resh and Heath Brown. His popular writing focuses on data science and rural public policy, especially politics and public policy in Appalachia.

Methods of the Policy Process

Edited by
Christopher M. Weible and
Samuel Workman

Contents

Figures

Tables

Contributors

Frank R. Baumgartner is Richard J. Richardson Distinguished Professor of Political Science at UNC–Chapel Hill. His work has focused on agenda setting, lobbying, and issue-framing in the contexts of US politics as well as in comparative perspective. With Bryan Jones, Dr. Baumgartner created the Policy Agendas Project, and with many collaborators, he has worked to create a comparative network of agendas scholars.

Ramiro Berardo is an Associate Professor at the School of Environment and Natural Resources at The Ohio State University. His research focuses on understanding how complex governance systems operate, with an emphasis on examining how collaboration can be engendered by policy stakeholders who interact in policy networks. Other research interests include hydraulic fracturing and climate change policy and water access and affordability.

Michael Cox is an Associate Professor in the Environmental Studies Program at Dartmouth College, in Hanover, NH. His research and teaching are rooted in the concepts of the commons, collective action, and local governance. He has conducted fieldwork focused on community-based natural resource management in New Mexico, Colorado, Peru, Kenya, South Africa, and most recently in the Dominican Republic.

Matthew Hamilton is an Assistant Professor in the School of Environment and Natural Resources at The Ohio State University. His research examines how patterns of interactions among individual people, organizations, and decision-making processes influence the performance of environmental institutions. He primarily applies the ecology of games framework to study environmental decision-making in the contexts of climate change adaptation and wildfire risk governance.

Adam Douglas Henry is Distinguished Scholar and Professor of Environmental Politics at the University of Arizona. His interests occupy the intersection of public policy, sustainability science, and computational social science. His work applies network analysis to the study of the policy process in domains such as energy policy, urban water management, and invasive species management, among others.

Nicole Herweg is a Postdoctoral Researcher in the Department of Political Science at the University of Heidelberg, Germany. Her research interests are in comparative public policy. Her research focuses in particular on policy processes, energy policy, and European Union politics. She is the author of *European Union Policy-Making. The Regulatory Shift in Natural Gas Market Policy* (2017).

Karin Ingold is a Professor at the Institute of Political Science as well as the Oeschger Center of Climate Change Research at the University of Bern. She heads the Policy Analysis and Environmental Governance research group affiliated with the Environmental Social Science Department at Eawag, the Water Research Institute of the ETH Domain in Zurich. She is interested in studying complex environmental problems and the production of policy solutions.

Bryan D. Jones is J.J. "Jake" Pickle Regent's Chair in Congressional Studies at the University of Texas at Austin, and Director of the US Policy Agendas Project. His research centers on American public policy processes, including agenda-setting and decision-making. He is author or co-author of 15 books and has authored or co-authored articles in *The American Political Science Review*, *The American Journal of Political Science*, *The Journal of Politics*, *The Policy Studies Journal*, *The Journal of Public Administration Research and Theory*, and others.

Michael D. Jones is an Associate Professor of Political Science and Director of the Master of Public Policy and Administration at the Howard H. Baker Jr. Center for Public Policy at the University of Tennessee. He is currently editor-in-chief of the *Policy Studies Journal* and has published in a wide range of disparate disciplinary venues such as *Political Psychology*, *Social Science Quarterly*, *Policy and Politics*, and *Critical Policy Studies*, among others.

Andrew Karch is Arleen C. Carlson Professor of Political Science at the University of Minnesota. His research centers on the political determinants of public policy choices in the contemporary United States, with a special focus on federalism and state politics. His most recent book, co-authored with Shanna Rose of Claremont McKenna College, is *Responsive States: Federalism and American Public Policy* (2019).

Mark Lubell is a Professor in the Department of Environmental Science and Policy at University of California, Davis. He studies cooperation problems in the context of environmental policy and governance. He is one of the developers of the ecology of games framework, with empirical applications in the context of water governance and climate change adaptation.

Mark K. McBeth is a Professor in the Department of Political Science at Idaho State University. His research focuses on using the Narrative Policy Framework to better understand the role of narrative in public policy. Substantively, his research focuses on a wide range of environmental issues including wildlife policy, energy and the environment, and climate change.

Jack Mewhirter is an Associate Professor in the Department of Political Science at the University of Cincinnati. His research examines the origins, evolution, and resolution of collective action problems within complex governance systems. Much of his work is done within the context of regional water governance and policing.

Jamila Michener is an Associate Professor of Government at Cornell University. She studies poverty, racism, and public policy, with a particular focus on health and housing. She is author of the award-winning book, *Fragmented Democracy: Medicaid, Federalism, and Unequal Politics*.

Daniel Nohrstedt is Professor in political science in the Department of Government at Uppsala University, Sweden, where he is also research coordinator in the Centre of Natural Hazards and Disaster Science (CNDS). He specializes in the policy process, particularly the advocacy coalition framework, spanning empirical applications in environmental policy and governance issues.

Edella Schlager is a Professor in the School of Government and Public Policy at the University of Arizona. Her research focuses on comparative institutional analysis using the IAD Framework and its associated tools. Substantively, her research program centers on multi-level and multi-scale governance of transboundary watersheds in the Western United States and New York State.

Elizabeth A. Shanahan is a Professor in the Department of Political Science and Associate Vice President for Research at Montana State University. Her research centers on theoretical advancement paired with practical applications of the Narrative Policy Framework in the context of risk.

Saba Siddiki is an Associate Professor in the Department of Public Administration and International Affairs, Syracuse University. Her research expertise is in policy design and collaborative policymaking. She has studied these topics in the contexts of agriculture, aquaculture, transportation, and energy policy. She uses institutional analysis techniques in her assessment of policy design and has been particularly involved in the development of institutional analysis techniques affiliated with the institutional analysis and development framework.

Aaron Smith-Walter is an Assistant Professor in the Department of Political Science at the University of Massachusetts Lowell. His research focuses on the narrative policy framework, cultural theory, and images of public administration in apocalyptic fiction. He also serves as Vice-President of the Massachusetts Chapter of the American Society for Public Administration, the Associate and Managing Editor of the *Policy Studies Journal*, and the administrator of the Narrative Policy Framework Online Community.

Geoboo Song is an Associate Professor of Political Science, Public Policy, and Statistics and Analytics at the University of Arkansas. He also serves as

Associate Editor for the *Policy Studies Journal* (PSJ). He has been involved in scholarly research that holds theoretical, practical, and methodological implications for advancing public policy studies.

Mallory SoRelle is an Assistant Professor of Public Policy at the Sanford School of Public Policy at Duke University. Her research investigates how public policies influence socioeconomic and political inequality in the United States. She is the author of *Democracy Declined: The Failed Politics of Consumer Financial Protection* and has published articles on the politics of public policy in journals including *Perspectives on Politics*, *Publius*, and *Interest Groups & Advocacy*.

Francesca Vantaggiato is a Lecturer (Assistant Professor) in Public Policy at King's College London. Her research expertise is in environmental governance, regulatory governance, and bureaucratic politics, particularly in multi-level governance settings. She has studied these topics in various contexts.

Christopher M. Weible is Professor in the School of Public Affairs at the University of Colorado Denver. His research focuses on conflict and collaboration in policy processes. He is the Editor of the *Theories of the Policy Process*, Co-Editor of *Policy & Politics*, and Co-Director of the Center for Policy and Democracy.

Samuel Workman is Professor of Political Science and Director of the Institute for Policy Research and Public Affairs in the Rockefeller School of Policy and Politics at West Virginia University. His work is focused on public policy, agenda setting, the bureaucracy, regulatory policy, and research methodology. He is Principal Investigator on the Education Regulations Project (EdRegs), funded by the National Science Foundation (NSF), and the author of *The Dynamics of Bureaucracy in the U.S. Government*. His work has appeared in the *Policy Studies Journal*, the *Journal of Public Administration Research and Theory*, *Policy and Politics*, *Interest Groups and Advocacy*, and *Towards Data Science*, among others. He is a Founding Editor of *3Streams* with William Resh and Heath Brown. His popular writing focuses on data science and rural public policy, especially politics and public policy in Appalachia.

Nikolaos Zahariadis is Mertie Buckman Chair and Professor of International Studies at Rhodes College, USA. He has published widely on issues of comparative public policy and European political economy. When not researching, he wishes the corona virus would just die so he can resume taking long inspirational walks on the beach.

Reimut Zohlnhöfer is Professor of Political Science at the University of Heidelberg, Germany. His research interests include political economy, policy process theory, and West European politics. He is author and co-editor of several books, including *Decision-Making under Ambiguity and Time Constraints. Assessing the Multiple-Streams Framework*, co-edited with Friedbert W. Rüb (2016).

1 The Design of Policy Process Research

Samuel Workman and Christopher M. Weible

Introduction

Growth and advancement in policy process studies depend not merely on developing better theories but also on developing better methods. In this volume, we use *"methods" as an umbrella term to denote systematic techniques for applying the various approaches, frameworks, and theories of the policy process and systematically measuring their key concepts.*[1] Methods are ways researchers put theories into action, test them, and improve them. Methods involve strategies for actual application, including the research design (e.g., sampling approaches), conceptualization and measurement, data collection (e.g., experiments, field research, surveys), modeling (e.g., models of government innovation), and data analyses (e.g., quantitative or qualitative approaches). Methods also include precise ways in which researchers connect variables in explaining a phenomenon in context.

A focus on methods in the study of policy processes is notably absent. While some strands of research have best practices for applying a theory, these methods are rarely understood or communicated outside of a given research program. In other instances, some theories have existed for decades but have yet to develop standard and relatable methods for application. Whether the methods exist but are not communicated or yet developed, the result is the same – limited growth and advancement in the study of policy processes.

This volume supplements the well-established and widely used *Theories of the Policy Process* (Weible and Sabatier 2018). First published in 1999, *Theories of the Policy Process* has served as the primary compilation of the most established theoretical approaches in studying policy processes. It emerged from the need to develop better theories and to communicate these theories in a single volume for new and experienced researchers (Sabatier 1991).

Today, however, the intellectual landscape is very different. The increasingly global study of policy processes faces severe challenges with scholars applying theories in radically different national and cultural contexts (Tosun and Workman 2018). Questions increasingly arise about conducting policy process research comparatively among a global community of scholars. For example, scholars often interpret concepts differently and use different measures in applying the same theory. We should undoubtedly encourage

diversity, experimentation, and creativity to advance the field. However, without some standardization, best practices, common templates, and general strategies for applications, the generation of our shared knowledge will be stunted, and many of the lessons learned will be overlooked or lost. We do not all have to do things the same way, but what we do and the logic underlying our approaches should be clear. Indeed, in advancing the study of policy processes, there has always been a necessary tension between recognizing contextual particularities and drawing appropriate generalizations.

This volume remedies this situation by bringing attention to methods of the policy process. Our intent is not to establish orthodoxy or impose a rigid way of researching on the policy process community. Instead, we offer the leading researchers of policy process theories the opportunity to share their best and established practices for theoretical applications. The goal is to communicate the diverse ways that we conduct theoretically based research toward better methods of the policy process.

Why This Book and Why Now?

For academia, the contributions to the field of policy processes will be nontrivial. Never has an edited volume sought to tackle issues of methods for theories of the policy process, and the need for such an emphasis has never been greater. In this respect, this volume has no peers. Many of these methods develop along parallel tracks within each of the theoretical traditions. This volume will allow academics and students, for the first time, the ability to see outside these tracks and be aware of the commonalities across the different approaches, as well as identify the key differences.

This volume serves as a reference for applying the theories for students at the undergraduate and graduate levels and experienced scholars. This compilation of theories captures the most established approaches with international research communities, ongoing empirical applications, efforts to push comparative insights, emphasis on being transparent and public, and continual growth in contributing knowledge about policy processes (Weible 2018).[2] However, there are also many quality books and essential ideas in public policy that we teach in our courses and reference in our research that do not meet these criteria or fit in this volume, one of which is Peter John's (2018) book *How Far to Nudge?* Public policy is associated with the fields of public administration, regulation, broader political science, and others that supplement theory-based policy process research and form valuable comparisons to approaches in public policy. Many of these related fields might also inform how to advance theory-based policy process research methods and methodologies. Where this is the case, we will discuss these areas of opportunity.

In addition to illuminating methods to spur innovation in public policy scholarship, broader trends in interdisciplinary science place a premium on making social science research methods accessible to students, researchers, and those in practice. Social and natural systems have become intertwined

as human beings increasingly influence the systems that govern the natural world (Stromberg 2013; B. A. Jones et al. 2016). Climate change, food systems, and environmental justice are just a few of the myriad issues demanding interdisciplinary science for workable solutions. These problems beg for interdisciplinary science that not only sees social science as translational but as intrinsic to understanding and solving societal problems. Make no mistake about it; these problems are *policy* problems lying at the intersection of the natural world and governance systems.

The barrier to interdisciplinary science is language, or rather, jargon. Yet, in our experience, methods and methodologies offer a common language for interdisciplinary science. By shedding light on the methods of policy processes, this volume speaks to the range of possibilities and opportunities for collaborative research between natural scientists and policy scholars. These benefits push far beyond translating science and research to impactful social science on the front-end design of problem-oriented research and toward the betterment of societies.

The Approaches and Data Science

The approaches to research design and analysis presented in this volume represent unique data sciences and traditions. Thus, it is helpful to consider the emergent field of data science when thinking through the chapters to follow. As a subset to methods, we define data science as the set of processes for collecting, cleaning, organizing, storing, and analyzing data.

For the experienced scholar, the chapters lay out myriad possibilities for the types of data one might collect, innovative measures, and the many analytical techniques for testing and refining theories (e.g., statistical modeling or qualitative case studies). For the beginning scholar, these are second-order concerns. Often, the most significant barrier to applying theories is what data to collect and how to organize them for analysis. Thus, to the extent possible, the chapters attempt to give the inexperienced scholar a glimpse of what type of data is typically used, how it is collected, and how it is organized for analysis. For example, how should data be organized in a spreadsheet to be useful to the analyst (Broman and Woo 2018)? These questions seem basic, but in fields stretching from bio-medical data to business to public administration, data are often not collected and organized in a way that allows for analysis across platforms, theoretical traditions, or disciplines (see Workman 2020).

The chapters to follow offer an overview of how data are collected, organized, and analyzed in each of the theoretical traditions we cover. These are not compendiums on data science within the theoretical traditions, but the reader will leave with a good idea for getting started. Moreover, the theories covered in the chapters encapsulate not only standard operating procedures for data collection, organization, and analysis but also the norms and traditions within each theoretical approach. Thus, data science is a set of objective guidelines and *perspectives* on what works best, given data and context within theoretical traditions.

How Does This Book Relate to the *Theories* Book?

Distinctions between deductive and inductive research are common in courses on research design for budding social scientists and data analysts of all types. In general, deductive research proceeds with theory as the main driver and develops conjectures about the world for testing with empirical research designed for confirmation or falsification. On the other hand, inductive research begins with data and observation, developing descriptions and explanations of the world from the ground up.

While the distinction between deductive and inductive research is a useful pedagogical tool for understanding the relationship between data and theory building and testing, science is an iterative exercise in exploration, discovery, and explanation. The inductive scientist develops a theory from careful observation of the world and then uses that understanding to gather more data or undertake more observations to refine that theory. Likewise, the deductive scientist undertakes observation, data collection, and analysis that supports their working theory of the phenomena they study in whole or in part. The theory is then discarded or refined in light of the data or observations. All science is an iterative process of deduction and induction, differing only from how one starts. Discussions of research approaches often abbreviate the iterative process of data and theory, but the arc is much longer, involving observation, theory building, theory testing, evaluation of the evidence, and finally, feedback and theoretical refinement.

In this light, there is a clear relationship between this volume and *Theories of the Policy Process*. This volume closes the iterative loop in policy process research. The methods covered are the engines for the refinement of the theories in that volume. Whereas the *Theories* volume provides an overview of the state of the theoretical approaches to public policy and traces the development of the core ideas, this volume relates to conducting research within the theoretical approaches.

Elucidating the different methods and methodological approaches in policy process research also holds the promise of cross-fertilization. Some of the concepts across theoretical traditions are common or similar. Take the example of the concept of feedback – how does policy itself fuel resultant politics and decision-making in policy systems and relate to other concepts like learning or attention? Awareness of how different theories conceptualize and operationalize common concepts offers opportunities for innovations in and across multiple research programs and for bridging the silos that result from their decades of advancement.

Methods in the Context of Theory

This volume aims to provide a resource for those looking to research and contribute to these theoretical traditions. To that end, the authors have endeavored to be systematic and transparent without imposing any rigid set of rules for the beginning analyst. Some of the authors dig deeply into

measurement and particular statistical techniques, while others offer a broad view of the techniques generally employed in the area, giving the reader a range of opportunities for engagement.

The theories embrace various methods for analyzing data and refining conjectures about the world. That last bit is important. The nature of scientific publishing means that we spend lots of time specifying, explaining, and justifying a method for a given piece of work. This is necessary, of course, to give the reader confidence in the findings. As a result, though, we tend to spend less time in any given project thinking about how that process should inform and refine our theory (usually leaving that to the literature review and theory in the next project). This volume offers that opportunity for those steeped in techniques and modes of analysis to evaluate those approaches, how they might be adjusted, and what needs reform. We return to the theme of building or creating knowledge from our methods and theories in the concluding chapter.

For the reader of this volume, we think two distinctions in the ways researchers use the approaches below are useful for understanding theory building, testing, and feedback in the traditions. The first distinction is whether the mode of analysis is qualitative or quantitative. In this volume, the modes of analysis span from advanced statistical analysis to qualitative case studies and comparison, process tracing, and more. There is no substantive advantage to quantitative or qualitative research – both can be systematic and transparent enough to follow their logic.

In the background, the distinction between deterministic and probabilistic models of the world lurks. Deterministic models explain the world and can predict the world accurately and precisely with minimal variance, given certain conditions. However, strictly speaking, nearly all social science is probabilistic – it explains the range of likely behaviors or outcomes under certain conditions. Certain behaviors and outcomes are more likely within this range, and our models attempt to attach probabilities to these, even when working qualitatively. The reasons for the probabilistic nature of social science are obvious – there is randomness in human behavior, and, thus, in social, political, and economic systems that distinguish them from natural systems. Additionally, policy and governance systems are complex and generally not linear and additive (e.g., more akin to the human body than a car). Finally, there is no "right" methodological argument or method, merely what is theoretically and logistically useful for building knowledge in the field.

We offer a roadmap for understanding and reasoning about the methods and methodologies to follow. The roadmap is a set of considerations addressing the research questions we ask, the research design, how the approaches treat data and measure concepts, and the tools for analyzing data. These are points of departure for the beginning analyst, not hard-and-fast strictures. The reader should realize that these are more than mechanical elements of research on the policy process – they are cultures and traditions built around researchers with a shared perspective and intellectual interest.

Research Programs and Culture

We discuss methods in depth in our work, at conferences, and in our collaborations as academics. Unfortunately, we discuss and contemplate the culture or cultures of research much less often. As anyone who has attempted to replicate a scientific study will tell you, it is difficult without awareness of all the little "common practices" and ways of doing and interacting with colleagues and data.

The approaches in this volume represent teams of diverse scholars from myriad backgrounds and identities. Many of them span the globe and are accessible. There is no substitute for talking with folks and getting their perspectives. The reader will find the authors in this volume open and accepting. Where there is doubt about how to start or at key decision points in data collection, measurement, or analysis, scholars working within these traditions will be an inexhaustible source of advice and experience.

We urge the reader to consult them in thinking about, designing, and planning research. These cultures can help identify and answer the types of questions amenable to a particular theory, common research pitfalls, ways of collecting and analyzing data, and generally organizing and designing research. Innovation, creativity, and development of the field are the goals, but innovation and creativity spring from knowledge and perspective.

> Where you stand depends on where you sit.
>
> (Rufus Miles)

Roadmap for Reasoning

In the sections below, we present a strategy for thinking about the theories and methods to follow. We forego a discussion of substantive theory and conceptualization, choosing instead to elaborate on themes common across these and all social scientific approaches. Consult *Theories of the Policy Process* for an analytical assessment of each theoretical tradition's core ideas and concepts. In the concluding chapter, we return to these themes and discuss the chapters, attentive to the layout presented here.

Research Questions and Scope

Scope pertains to the contextual range in which the theory is operational – this is both a state and an aspiration. Scope arises from the systematic application of the theories and more organically from the intellectual interests of the associated research programs and the people therein. For example, one might study environmental policy because they think the environment is of interest. Another might study environmental policy indirectly through their interests in the operations and decisions of subnational and national governments. The accumulation of these decisions and the resulting theoretical refinements give a theory its scope. However, sometimes, we

learn the most about a theory when applying it outside its intended scope. Alternatively, a theory developed to study natural resource policy may not apply well to civil rights policy.

Public policy contains three settings that go a long way in delineating a theory's scope. The first is the substantive policy or problem (e.g., the environment, education, civil rights, and liberties). The second is the governance level within or beyond a political system (e.g., supranational, national, subnational). The third is the loci of political activity or the decision-making (or policy-making) venues of the collective action of interest, such as organizations and legislative bodies like parliaments, policy subsystems, or networks. For example, a theory and methods might focus on a specific governance level (e.g., subnational) and the adoption of an idea across city governments as found in scholarship associated with Innovation and Diffusion. Sometimes a specific substantive policy problem anchors an entire theory, such as the Institutional Analysis and Development (IAD) Framework's Common Pool Resource Theory. It can be useful to think about these three settings at the micro, meso, and macro levels. In general, micro refers to the applicability of an approach at the individual level. Meso level refers to groups, organizations, networks, and policy subsystems, where collections and coalitions of actors make policy. Macro-level approaches refer to the applicability to the level of governing institutions and countries.

The reader should note that scope is both systematic and intentional. The methodologies and methods in scope are a collection of arguments and decisions about theory building, testing, and development.[3] However, the scope is also logistical – a product of data availability and modes of analysis. In other words, scope is tethered to what data we can collect and how we can analyze them. Punctuated Equilibrium, for instance, is a theory of agenda-setting and resultant policy change. Its scope is mainly national or macro analyses as one cannot study agenda-setting and issue prioritization by analyzing only one issue on the agenda. Thus, it does not restrict itself to a specific policy or problem area. At the national or macro level, its locus is decision-making venues, such as legislatures. The theory also demands the explicit comparison across issues, decision-making venues, and even countries. Logistically, this implies studying the things governments consider, discuss, and act on – laws, budgets, speeches, hearings.

With each chapter, consider the scope of the theory and methods. What issues, levels of governance, and loci of political activities or decision-making venues appear in the typical analysis? Scope offers the first clue as to what types of data we need to apply any theory. It also delineates the kinds of appropriate questions for each theory.

Research Design as Comparison

All research design is comparison. Research design structures data collection and analysis to allow implicit or explicit comparisons across time, spatial units, or some theoretical baseline. For example, experimental designs allow us to

compare the effect of a treatment to a control group without treatment or intervention. Comparative research designs that, for example, analyze governance at the macro or national levels across more than one country allow comparisons of policy-making given variation in settings. Comparison is even operative for the scholars conducting field research on a single case study engaging in sense-making indirectly premised on relations with different contexts. In these examples and more, this research is explicitly or implicitly compared or triangulated with the analyst's knowledge of other cases and to learn something and understand the world a different way. The notion that learning requires comparison is operative at every quasi-experimental approach between these two endpoints.

Comparison across time implies the collection of data over periods, preferably substantial periods. The unit of time could be anything from seconds to years. Spatial units might refer to individuals, organizations, governments, countries, or even policy issues, among many others. Policy Feedback Theory (PFT), for example, asks how policies influence outcomes for individuals that then shape mass political behavior. This implies a time component – citizens react to policy over time. It may also compare how policies shape the behavior of different groups within the citizenry – gleaning information from comparisons across societal groups. Likewise, the Narrative Policy Framework (NPF) may examine the evolution of a given narrative over time or compare narratives curated by different actors.

The point of comparison need not be empirical. Some approaches advance with research designs that compare empirical data to some theoretical baseline. For example, Punctuated Equilibrium advanced its depiction of policy change by gauging the degree to which budgetary and other data diverge from the normal distribution. These simple, descriptive comparisons were the launching pad for better explanations of incremental and large-scale policy change. In addition, methods utilizing network analysis and relational data often compare empirical networks to random theoretical networks to study everything from power relations to likely information flows.

Baselines are important for gauging what we learn from comparison. For the beginning researcher, it is useful to consider how each chapter and research design sets up a *counterfactual*. Counterfactual is a staple tool of the qualitative researcher, especially those working in case study or process-tracing methods. Counterfactuals are exercises in establishing baselines. The researcher establishes a world or context in which a key causal factor is present or not and proceeds to examine whether the outcome would have been altered. These counterfactuals can be made through abstract or reasoned arguments, and other times they are not. Regardless, counterfactuals are a vital technique for making causal arguments (Brady 2011). In the example of Punctuated Equilibrium Theory (PET) or the Ecology of Games, the counterfactual might be a theoretical baseline such as the normal distribution or a random network graph. However, a counterfactual need not be theoretical.

Modern computing and statistical simulation create the possibility of exploring counterfactuals from statistical models in quasi-experimental

designs. When we include variables in models that tap into citizens' iden-
tities, effects for governmental types, or any other phenomenon of interest,
we implicitly say that there is a difference between these identities or types
compared to some others – a baseline or counterfactual. Thus, all methods
making causal arguments embody some form of counterfactual, whether
theoretical or empirical. Counterfactuals or baselines for comparison need
not embody the average or modal category or effect (e.g., the average citizen)
and are more often useful if they tap a comparison that exists in the sample
or population we study (e.g., transportation compared to environmental
policy, flat networks compared to hierarchies, or middle-aged black women
compared to white, middle-aged men). When counterfactuals and baselines
have a presence in the real world we study, our findings are given life and
meaning in that world and not just in the social science realm. The reader
would do well to ask how each chapter treats the design of counterfactuals
and how each designs research for exploring counterfactuals or baselines.
The first step in doing so is thinking about the measurement of key concepts.

Social science methodology and methods are an exercise in simplifying
the real world. We can understand it, explain it, make sense of it, and arrive at
probabilistic predictions about human behavior, the likely impacts of public
policy, and the workings of social systems. Parsimony is valuable because it
makes generalization easier, illuminating the critical components of social
systems to build knowledge about them. This process of simplification comes
with costs, however. The major stages of methodologies and methods involve
abstractions from the real, social world. So, the reader can think of meth-
odologies and methods as introducing slippage between reality and our
parsimonious models of behavior and systems. This slippage is valuable for
understanding, but there is always tension between adequately describing a
social system and maintaining enough simplicity such that knowledge can
accumulate. The two major points of abstraction are measurement and ana-
lysis, and we tackle them in turn.

Measurement

Often overlooked, measurement is the first stage in conducting any empirical
inquiry to apply a model to observations. Measurement has consequences
for the types of analyses that we can deploy on data and for the conjectures
we will draw about the world we observe, not to mention their generaliz-
ability. A fantastic mentor in measurement and statistical analysis relayed an
example of this to one of us.[4] Imagine that you are trying to measure the
concept of "bug health." Unless you are an entomologist, it is likely that you
only observe that a bug is either "alive" or "dead." This measure introduces
a simplification that leads to slippage in measuring the key concept. Why?
Because ill bugs will likely be measured as "alive." Furthermore, the cat-
egories "alive" or "dead" cannot distinguish between bugs that are very sick
or perfectly healthy. In other words, "alive" or "dead" are noisy measures of a
theoretical concept – "bug health" – useful but not perfect.

This example raises important considerations for measurement and points to measurement as the initial stage of deploying a model on data, keeping in mind that models, or simplifications, are necessary to build understanding and generalize about the world. On the leading edge of measurement is the distinction between the *unit of observation* and the *unit of analysis*. Imagine we want to study how students perform across countries. We then gather student testing scores. For our purposes, the unit of observation pertains to the unit of data on which we construct measures. In this example, the units of observation are students or perhaps cohorts or classes of students because we gather observations on students and aggregate them to the country level using various measures like a mean or median for each country. The units of analysis are the units to which we wish to extend inferences or conjectures – our generalizations. In this case, we want to distinguish student performance by country – so countries are the units of analysis.

Contrast this with our example of "bug health." There, the units of observation and analysis are the same – bugs. There is more slippage when using students to conjecture about countries than using bugs to conjecture about bugs. In each of the chapters below, readers should consider typical units of observation and units of analysis within each tradition. These considerations are paramount for understanding how we collectively might improve our measures to theorize about and understand the policy process. These distinctions hold regardless of the qualitative or quantitative nature of the research process. There is tremendous guidance in broader political science for thinking innovatively about maximizing observation to bring more weight to the conjectures we make about the unit of analysis (King, Keohane, and Verba 1994).

Choosing Indicators and Measures

The process of measurement involves two layers of choices. With each, the researcher makes consequential decisions about the amount of simplification that will result, and by extension, the ability to generalize or make accurate conjectures. The process of measurement begins with a theoretical concept. For reference on the development of key concepts for each theoretical tradition, *Theories of the Policy Process* is a great launching pad. In practice, we often go from a key concept (e.g., policy change) to a measure (e.g., count of new laws). However, doing so conceals an intervening choice that is immensely consequential for the fit of our measurement to the concept – the choice of an *indicator*.

Figure 1.1 displays the relationship between concept, indicator, and measure (see Goertz 2006). On the left, the figure shows a conceptual chart, and the right shows the chart applied to the concept of policy change. Concepts are the components of frameworks and the endpoints for causal arrows in our theories. We start with a theoretical concept that we must measure, and the first choice we confront is what indicator of the concept
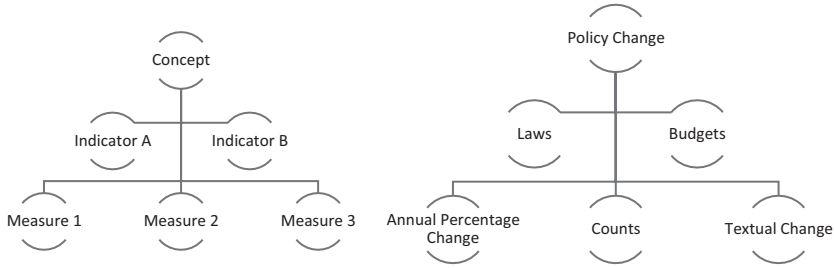
Figure 1.1 Concepts, indicators, and measures

to use. Each concept will have multiple potential indicators (i.e., sources of data). For example, in policy change, we might choose laws or budgets, but there are many more indicators (e.g., regulations, personnel, networks, agendas, coalitions). With the choice of indicator, we have made our first abstraction from the concept – laws will be a better choice in some contexts and budgets in others, but neither perfectly reflects the concept of policy change. Yet, each allows us to simplify the concept and make conjectures about policy change given certain conditions. For example, major substantive laws might change while budgets remain similar. Likewise, there could be drastic changes in budgets but very little change in substantive law.

Once we have chosen an indicator, we choose measurements constructed from observations in the indicator or source of data. Just as each concept has many potential indicators, the researcher may construct many potential measures from that indicator. In the example of laws, the analyst may choose to count the number of laws produced by a legislative body or examine the textual changes in new laws, comparing them to previous ones.

The reader will note that each choice of indicator and measures will allow generalizations that are better or worse in certain contexts. It should also be clear that each choice introduces noise in our efforts to represent the concept accurately. For this reason, multiple indicators and multiple measures are often the optimal strategies. In the digital age, data are more available than ever before but often require time and resources or advanced programming skills to collect them. We encourage the reader to consider how the chapters treat indicators and measures and whether research designs within the theoretical traditions incorporate multiple indicators and measures.

When the authors choose indicators, what is their logic or justification for doing so? Being open-minded and considerate about measurement is a major avenue for understanding and improving the approaches presented here. Attention to new indicators or innovative measures promises major advancements and contributions to the theories. This process of getting from a concept to a measure is often labeled operationalization. The fidelity of the measure to the concept is key to understanding the value we can place on our generalizations.

Validity and Reliability

The culmination of thinking about the choice of indicators, or sources of data, and measures is in evaluating the tradeoffs in generalizing about the policy process. The most pertinent dimensions are *validity* and *reliability*. We briefly cover these here as one of our chapters goes into great depth on the distinction as applied to data and measures – the reader would do well to consult Workman, Baumgartner, and Jones (2021) in Chapter 3.

In the basest sense, validity refers to how accurately a measure reflects some characteristic or quality of the data. Simply, is the measure correct? Face validity is the most straightforward of the variations – does the measure make sense in terms of its aims. Thus, a measure's validity is an evaluation of whether the measure is fitted to its purpose, whether it adequately differentiates between objects that differ, and likewise characterizes objects with similar traits in the same way. There are many forms of validity, but all relate to whether the measure accurately reflects characteristics of the objects to which it is applied and allows researchers to distinguish one object from another.

Reliability refers to whether observations, cases, subjects, or similar items receive similar values on the measurement scale or similar classifications. Reliability is not just an important methodological consideration; it is an important logistical consideration in measurement design (see Workman 2020). Arguments about validity tend towards the theoretical and conceptual. Researchers debate which measure is the "right" one given the theoretical concept. Where there is great debate about validity, reliability may become a pernicious problem. It is perhaps easiest to see this in classification and categorization systems.

Classification and categorization are fundamental to all scientific endeavors. Imagine a scenario where we attempt to code an object into a classification or categorization scheme. We assign a code and, over time, learn that, while we assign the "correct" code on average, there is a wide variance in the assignment of codes to an object of a given type or class. While we are right on average, we are very often wrong. In other words, reliability is low. In this case, the researcher is better off assigning the code that logistically is the most reliable. Without reliability in the assignment of the code, attaining reliability means re-coding *all* observations *again*. With reliability, we can have an informed debate about the correct code after coding all observations. If we agree, we can reassign those objects to the agreed-upon code en masse. Meanwhile, classification or categorization that is unreliable means coding all data anew.

In this way, debates about bias and reliability in open measurement systems differ from those in statistical modeling, where bias is often unknown or, at a minimum, is not transparent. The reader of the following chapters should consider both validity and reliability from a theoretical standpoint and the context of the traditions and norms of data collection and organization within each tradition. Both present theoretical and conceptual challenges

as well as logistical ones. Solving one does not necessarily mean assuaging the other. These considerations are even more important for the beginning researcher who is not privy to the standard procedures undertaken to collect and organize data within each theoretical tradition.

The reader will note that validity and reliability considerations are not confined to qualitative or quantitative research but are ubiquitous in all types of social inquiry. Whether the methodological approaches to follow rely on quantitative or qualitative designs is an important distinction as each has different ways of dealing with reliability and validity.

Above, we have argued that all research design is comparison. Addressing validity and reliability is always a key concern of measurement whether the research design aims at spatial or temporal comparisons. For the approaches to follow that examine temporal dynamics, reliability is an even greater concern, since measurement requires backward compatibility to construct trends and must be attentive to measurement drift – this is all made more difficult amid changing language and definitions of measurements in society writ large (B. D. Jones 2016). Yet, again, the inevitable tradeoff returns – if the only reliable longitudinal measure for a concept inadequately represents that concept, we might be enticed by the discovered longitudinal patterns, but those patterns might not offer valid conceptual or theoretical meaning.

In sum, consider how each theory treats time or temporal dynamics and the tradeoffs between reliability and validity in their measurements. Some approaches explicitly deal with temporal research designs, which are central to theorizing and research design and emphasize reliability and validity differently (e.g., Punctuated Equilibrium and PFT). Other approaches might have developed highly reliable measures for analyses across space and time, but at what loss of validity? The reader should note that the choice between spatial and temporal designs and balancing reliability and validity in measurements is just that – a choice. Even for approaches centered on temporal designs, one gets only so far without comparing the dynamics of one system, country, institution, network, narrative, or coalition to another. We conclude this section by noting that measurement represents a major area for developing and refining theoretical traditions. In general, we find that policy process research would benefit by modeling less and measuring more.

Tools of Analysis

When discussing the tools of analysis, it is easy to immediately jump to thinking about statistical models, qualitative comparative analysis, process tracing, network analysis, distributional analysis, and the like. However, more pertinent distinctions allow the reader to reason across chapters and approaches and not just within theoretical traditions. The advancement of knowledge about the policy process depends on what these tools of analysis can tell us in light of the comparisons at the heart of analyses.

We ask that the reader keep in mind these broad distinctions in analysis and think about how the different modes work within the methodological

approach presented in each chapter. Each will vary with the degree to which these motivate the work. They are essential in guiding analysis within projects and matter for understanding learning and refinement within each theoretical tradition.

Descriptive Versus Causal Inference

One might think that we would start with quantitative versus qualitative research. We think this entirely unhelpful in the context of drawing conjectures about the policy process. Our concern is with methods that are systematic and generate descriptions, explanations of, or predictions about the world that are verifiable in empirical data. Both quantitative and qualitative research fit the bill. Throughout the chapters to follow, the reader will note points where quantitative researchers call for more qualitative work and vice versa. This is healthy and presages a bright future for the study of public policy with innovative new insights and refinements of our current understanding.

The goal of science of all types is to understand the world, how it works, and why it works the way it does. This broad aim can be broken into three component concerns – to accurately *describe* the world, reduce the complexity in our descriptions of the world to *explain* it, and use these explanations to forecast or *predict* phenomena. Each of these is important. One can predict phenomena without understanding them very well, for instance, but without description and explanation, generalizations about the world (the policy process for us) will be limited and leave the analyst vulnerable to errors in other contexts.

Science is borne on the sibling endeavors of descriptive and causal inference. We engage in *descriptive inference* when describing something unobserved from what we observe (King et al. 1994). Descriptive inference is inevitable in all social science research. We never observe everything, which then raises questions about the extent that we can infer something from the observed to the unobserved and the confidence in making such inferences. However, descriptive inference does not end with an accurate description of an observation, class of observations, or phenomenon. It also encompasses efforts to delineate these observations, classes, or phenomena from others. Throughout this volume, the reader will note the tremendous strides in understanding the policy process made through descriptive inference.

PET, for instance, made the leap from case studies to large-scale examinations of public policy in part by descriptive inference – taking the distribution of policy outputs seriously. How do the output distributions of policy systems compare to one another, and what we might expect theoretically? They began with accurate description – what many considered to be policy output distributions with the shape of the theoretical "normal" distribution were not normal at all on close inspection. This theoretical and methodological leap was made possible by an earlier qualitative description of the destabilization of policy subsystems combined with visual displays of

the time trends around subsystem politics. These early case studies laid bare potential causal mechanisms for the distributions modern scholars examine.

To read academic policy journals is to understand that the step of descriptive inference is often assumed away or buried beneath the burden of issuing a model and making a causal argument. In our view, the discipline does not take descriptive inference seriously enough. It is worth noting how the chapters to follow treat descriptive inference. Does it appear central to the development of methods and theory, or is it a precursor to causal inference?

Descriptive inference also relates to the discussion of measurement. In any methodological approach, researchers may use some measures that are created for the project at hand alongside others gleaned from secondary sources. Naturally, measures custom-made for a given research design have greater internal validity for said design; however, there are costs regarding external validity and the ability to generalize across theories. Descriptive inference opens the door to synthesizing across theories.

For instance, the IAD Framework, the Ecology of Games Framework (EGF), and the Advocacy Coalition Framework (ACF) are all premised on identifying a polycentric system, subsystem, or action arena where actors engage the policy process. Tremendous work has gone into describing these conceptual elements that have empirical structure and implications. Each engages in descriptive inference in unique and innovative ways to identify a policy system, governance system, or institution. For example, when researchers specify polycentric governance systems and apply the EGF, great effort is necessary to accurately identify, describe, and bound the system empirically and geographically. These scholars do not measure all aspects of polycentric governance but rather observe what they can observe and infer the rest. The ACF offers various measures of advocacy coalitions using a variety of different designs to infer reasonably comprehensive descriptions of them.

For some approaches, descriptive inference is a prelude to causal inference. *Causal inference* seeks to make causal claims based on generalizations from empirical data. It is essential to remember our earlier discussion of the nature of social scientific models. There, we noted that most social scientific theories are probabilistic rather than deterministic. The key takeaway from that discussion is that human individual and social behavior exhibits an element of randomness. Randomness means that we can attach probabilities to certain behaviors and outcomes but not specify them with certainty. This makes some uneasy but is necessary for accurate description and explanation of human social systems. The source of uneasiness relates mainly to the ability of a theory and methodological approach to predict behavior. As we note above, prediction has little value where the system is not descriptive enough to foster explanation and understanding.

The logistical and ethical constraints on experimental design and the randomness in human social behavior leave a portion of policy process research in the realm of *quasi-experimental* causal research, especially research using quantitative techniques. Quasi-experimental designs substitute statistical

control for experimental control. The quality of causal inference depends on the data and the nature of the quasi-experimental research design. In terms of data, the reader will note that some of the chapters rely on samples of the population of interest (e.g., PFT). Others base analyses on entire populations of data (e.g., PET). When evaluating any research, the reader should pay close attention to the nature of samples used in drawing conjectures. Are samples randomly drawn from the population with weighting and balance as in studies of mass behavior (e.g., PFT), or are they stratified or purposive as is necessary for identifying governance systems (e.g., Advocacy Coalitions Framework)? There is also the larger question of what statistical analysis means on entire populations rather than samples – a question that the statistical discipline itself has yet to work out fully.

Causal inference also implies attention to the logic of causality, including identification, though identification is not synonymous with causality (Kahn and Whited 2018). Identification implies establishing a covariation between a cause and effect, a temporal relationship between them, and control for other causes and conditions. We will assume that, for most studies, there is some correlation or covariation between a hypothesized cause and effect. In addition, all observed and unobserved links between a cause and effect (also called "causal mechanisms") provide theoretical justification for the relationship. Thus, the conditions for identification of causality are merely our current best practices for making such claims. To be clear, theories in this volume do not assume direct observations of causality. Making causal claims essentially means setting up a research design and offering an argument for what we think might be a causal relationship. When it comes to causality, we never know for sure, and uncertainty always persists.

How each chapter treats time is extremely important for understanding causal arguments and claims. Measures that purport to reflect a cause cannot come to empirical realization in the stream of time before those tapping the effect. For example, an annual survey of public opinion that typically occurs in December cannot be argued to be a cause of behavior in January of the same year. Likewise, it is often the case that research designs attentive to temporal dynamics lead to modeling strategies that assume constant effects of some cause across time. However, policy dynamics teach us that this is unlikely. In fact, several of the chapters to follow explore the complex process of feedback and dependence, contagion, and tipping points that lead to volatility in causal effects.

At the heart of this discussion of temporal design is the distinction between linear additive systems and complex, dependent systems (Simon 1996). A linear additive system is the sum of its parts – understanding causal effects means isolating a part of the system and measuring its impacts. Much of physics and econometrics relies on the notion that systems have equilibria and can be parsed into linear additive components. The analog is an automobile. A skilled mechanic can isolate a problematic part, focus attention and effort on that part, and be assured of the effects of their endeavor. But the dependencies and hierarchies that structure policy dynamics are more akin

to a biological system like a body. For the doctor, causes are difficult to isolate due to dependent, overlapping systems. Furthermore, treating a specific system does not allow the doctor to map treatments to resulting outcomes straightforwardly. In this way, the policy process is more like biology than physics, at least at the stage where we theorize, model, and assess causality and outcomes.

Many of the chapters to follow explicitly conceptualize, measure, and analyze dependent processes. Feedback, an explicit form of dependence and endogeneity, characterizes both Punctuated Equilibrium and PFT. The networks lying at the center of the Ecology of Games are premised on reciprocal relationships (even if imbued with power) through time in determining governance arrangements and policy outcomes. In the ACF, policy change results from a special kind of dependence – competition within subsystems among advocacy coalitions such that actors learn from strategies and scientific and technical information. The Multiple Streams Framework (MSF) envisions separate systems for problems and politics such that outcomes are explicitly dependent on the coupling of independent systems. Innovation and Diffusion models suggest that subnational policy results from monitoring, mimicking, and learning from other actors or states under various conditions.

Dependence and Endogeneity

Policy processes are dependent on at least two layers. First, dependence comes from government hierarchy and the interdependence and overlap of jurisdictions. Features of government systems like federalism and overlapping jurisdictions induce dependency across levels. Even when jurisdictions do not overlap, policy diffusion and contagion foster dependency in policies and outcomes (see the chapter on Diffusion and Innovation and arguments about polycentricity in EGF and the IAD Framework).

The substantive nature of problems overlaps with governance features to add a layer of dependence. Many emergent policy problems are ill-fit to existing governance structures, subsystems, or networks (Lewallen 2021; Workman 2015, pp. 116–123; May, Workman, and Jones 2008). These problems engage multiple venues of policy-making and multiple levels of government. Climate change is an excellent example – it engages governance structures and actors at multiple levels across agriculture, environmental policy, energy, and transportation, to name a few. If scholarship takes the Anthropocene seriously, then dependence is increasing not just among social systems but also among human social and natural systems.

The statistical terminology for much of this is *endogeneity*. Endogeneity refers to dual causality or a statistical inability to distinguish causal ordering. Endogeneity is worrisome for the statistical analyst and especially the causal modeler. Though worry we will, dependence and endogeneity are a feature of social systems rather than an aberration. Statistical modeling techniques, especially in econometrics, deal with endogeneity by incorporating temporal designs to suss causal ordering and feedback. Network analysis, most

prominent in the chapter on the EGF, takes endogeneity at face value, dealing with it by examining the structure and organization of policy networks. Like the one on PFT, other chapters take as a given that policies, themselves a product of mass political behavior, feedback into and mold that same behavior.

Qualitative researchers deal with endogeneity by paying careful attention to describing and specifying causal mechanisms. Addressing causal mechanisms is long thought to be a strength of qualitative research where researchers spend time and resources comparing causal events, actors, and process-tracing policy outcomes. In each chapter to follow, the reader will note points in their development where qualitative work would prove particularly useful in advancing theory.

Visualizing Data

Finally, the reader will note that visualization plays a more prominent role in methodological approaches and methods than ever before. Some of the chapters make extensive use of figures. For example, the EGF comes to life in its depictions of the networks that form the fabric of polycentric governance systems. PET relies on visualizing trends in attention to various policy topics. These approaches to visualizing data are not throwaway elements of analysis but form the core methods for supporting theories. Each is developed for dealing visually with the types of concerns we discuss above.

The reader should pay careful attention to each figure and table. Hidden beneath lessons about approaches to methods and methodology in these chapters is a practical set of lessons about presenting and communicating research results. Note that figures appear as ways to simplify and communicate theory and conceptualization and tools for communicating results.

Upon reading this volume, the reader will follow the trail of citations and compose and organize a manuscript that is consistent with common practice in each of the theoretical traditions covered here. We encourage all readers, especially beginning researchers, to be attentive to these practical lessons as they familiarize themselves with the presented approaches.

Up Ahead

Our roadmap concludes with a preview of what is up ahead. Our chapter ordering follows that found in the *Theories of the Policy Process*. For each chapter, consider the points we raise above. The concluding chapter will return to many of these themes and situate them within the broader endeavor of science. The reader should also begin to imagine how they might improve on the approaches presented here – that is, after all, part of the point of such a volume – to spur refinement and innovation from a place of understanding.

In Chapter 2, Zohlnhöfer, Herweg, and Zahariadis, (2021) detail the rich theoretical application of the MSF. This chapter gives the reader an overview of a qualitatively lively and rich set of studies that has expanded considerably

in the comparative context. Their central concerns for the future development of the MSF relate to standards of measurement and common, shared conceptualization. Even in qualitative research, their assessment calls for aggregating the rich case studies of policy change to drive generalizations and the future of the approach.

In Chapter 3, Workman, Baumgartner, and Jones (2021) discuss the methodological approach and development of PET. They tease out the tradeoffs inherent in measurement related to validity and reliability, especially in systems of classification and categorization. They also detail how to extend the Comparative Agendas Project (CAP) to other contexts and finer levels of detail. They cover the general distributional approach to quantitative analysis and relate it to the key elements of the theory. The reader will find a wealth of citations for further exploration of key areas. They conclude their chapter with a call for collegiality in adapting content coding systems and for more qualitative work that addresses key issues and causal mechanisms.

Chapter 4 addresses PFT. SoRelle and Michener (2021) argue for renewal in Lowi's (1964) premise that policy causes politics. They sketch the methodological approach as applied to mass publics interacting with government policies and programs. The approach addresses how policies influence mass political behavior generally and has particular advantages for understanding social justice, inequality, and minority politics. They illuminate the key challenges faced by measuring mass behavior, government programs (for which interaction is ongoing), and the difficulties of attaining data. This chapter addresses these and lays out a strategy for overcoming them in future research. It summarizes the major data sources and techniques of analysis.

In Chapter 5, Henry, Ingold, Nohrstedt, and Weible (2021) explore the methods and methodological approach in the ACF. This chapter goes into depth on typical units of analysis and how to bound a study. They develop a lexicon to understand key concepts and ideas and how they attach to measurement strategies with step-by-step recommendations. They conclude their chapter with calls for transparency and efforts to standardize at least some elements of approaches. This endeavor will be key for future advancement of the theoretical tradition for a framework so widely applied.

In Chapter 6, Jones, McBeth, Shanahan, Smith-Walter, and Song (2021) develop the methodological approach and associated methods of the NPF. They specify their assumptions that support NPF research and the latest methods for data collection and analysis. Given the research questions or objectives, this chapter provides an overview of varied research designs for moving forward at micro, meso, or macro levels of analysis. As with many chapters in this volume, measurement is critical, and this chapter devotes space to operationalizing NPF concepts. Finally, the chapter concludes with ideas for future research.

Chapter 7 traces the methodological development of Innovation and Diffusion. In this chapter, Karch (2021) examines how the study of policy diffusion has evolved from early considerations of likeness or nearness geographically. The approach has made strides with measurement and evolved

alongside the interconnectedness of policy and governance structures. In particular, this chapter aims for the indeterminate nature of supposed causal mechanisms as they attach to observed diffusion. Karch also tackles thorny methodological problems related to sample bias and understanding unobserved policies that do not diffuse. A key concern for the literature moving forward is how it addresses similar processes ongoing comparatively or cross-nationally.

In Chapter 8, Schlager, Siddiki, and Cox (2021) explore IAD as it has evolved methodologically. The IAD was developed to study problems related to collective action and self-governance, and the theoretical tradition excels in this regard. With the ACF in Chapter 5, these approaches offer a wealth of lessons for bounding a study and rooting it in the action of the engaged public and policy actors on the ground with authority to alter policy dynamics. They note that the case study is the main engine of analysis and remains the most versatile, but they explore the promising experimental research designs emerging in recent studies. The major innovation in recent work is the development and deployment of the institutional grammar tool – its value increasing with modern approaches to computing and machine learning.

Finally, Chapter 9 examines methodological development within the EGF. Lubell, Hamilton, Mewhirter, Vantaggiato, and Berardo (2021) give us what is likely the best illumination of complexity and dependence in the inner workings of policy and governance systems. Network analysis is a key engine for the EGF. These studies come at complexity and dependence head-on. They note that there are key questions relating to evolving complex systems. The temporal dynamics of such systems present tremendous problems for data collection – most significant time and resources. They note that the tradition would also benefit from expansion into other substantive realms beyond environmental governance.

We conclude in Chapter 10 with themes across the methodological traditions. We use what we have learned to discuss four principles for contributing scientific knowledge to our understanding of the policy process. In particular, we discuss how we can collectively be better in generating knowledge that is sound and derived from clear, rich methodological traditions. We finish with the notion that science is a creative endeavor, and as such, we should understand how we bring ourselves into our research.

To close this introduction, we want to remind readers of the reasons for doing any of this. As a social science, our goal is to develop knowledge about policy processes. While we conduct our science through the interplay of theories and methods, our continued progress hinges upon the people involved and the course and cover of our collective dialogues and engagements. With certainty, this volume will not be the final word on the methods of the policy process. But, more provisionally, it might prompt more words about methods in our discussions and lead to improvements in our science and contributions to societies.

Notes

1 We use "theories" generically to represent any research approach to focus the scope of inquiry, help specify assumptions, and define and relate concepts. This includes the possibility of establishing various relational forms, such as hypotheses or propositions. See Weible (2018, p. 1) for elaborations.
2 In this volume, we refer to policy process research or policy process studies as the field which consists of theories and methods of the policy process. Of course, policy process studies are broader than these theories and methods.
3 Whereas methods refer to the tools and techniques for conducting research, methodologies refer to the justification or rationale for choosing methods, a discussion point to which we return in the conclusion.
4 Chris Adolph at the University of Washington was famous for this example in statistical courses on measurement. It is illustrative of the slippage that happens during observation and measurement and helped one of us understand these things immensely.

References

Brady, Henry E. 2011. *Causation and Explanation in Social Science*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199604456.013.0049.

Broman, Karl W., and Kara H. Woo. 2018. "Data Organization in Spreadsheets." *The American Statistician* 72 (1): 2–10. https://doi.org/10.1080/00031305.2017.1375989.

Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.

Henry, Adam Douglas, Karin Ingold, Christopher M. Weible, and Daniel Nohrstedt. 2021. "Advocacy Coalition Framework: Advice on Applications and Methods." In *Methods of the Policy Process*. Milton Park: Routledge.

Herweg, Nicole, Nicholaos Zahariadis, and Reimut Zohlnhöfer. 2021. "How to Conduct a Multiple Streams Study." In *Methods of the Policy Process*. Milton Park: Routledge.

John, Peter. 2018. *How Far to Nudge? Assessing Behavioural Public Policy*. Cheltenham: Edward Elgar.

Jones, Benjamin A., Robert P. Berrens, Hank C. Jenkins-Smith, Carol L. Silva, Deven E. Carlson, Joseph T. Ripberger, Kuhika Gupta, and Nina Carlson. 2016. "Valuation in the Anthropocene: Exploring Options for Alternative Operations of the Glen Canyon Dam." *Water Resources and Economics* 14 (April): 13–30. https://doi.org/10/gkf9hz.

Jones, Bryan D. 2016. "The Comparative Policy Agendas Projects as Measurement Systems: Response to Dowding, Hindmoor and Martin." *Journal of Public Policy* 36 (1): 31–46. https://doi.org/10/gkr4hk.

Jones, Michael D., Mark K. McBeth, Elizabeth A. Shanahan, Aaron Smith-Walter, and Geoboo Song. 2021). "Conducting Narrative Policy Framework Research: From Theory to Methods." In *Methods of the Policy Process*. Milton Park, Routledge.

Kahn, R., and Toni M. Whited. 2018. "Identification Is Not Causality, and Vice Versa." *The Review of Corporate Finance Studies* 7 (1): 1–21. https://doi.org/10/gkx2wf.

Karch, Andrew. 2021. "Innovation and Diffusion: Connecting Theory and Method." In *Methods of the Policy Process*. Milton Park: Routledge.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

Lewallen, Jonathan. 2021. "Subsystems and Ill-fitting Problems: Clarifying a Concept." *Policy Studies Journal*, May: psj.12432. https://doi.org/10/gj9s5z.

Lowi, Theodore J. 1964. "American Business, Public Policy, Case Studies, and Political Theory." *World Politics* 16 (July): 677–715.

Lubell, Mark, Matthew Hamilton, Jack Mewhirter, Francesca Pia Vantaggiato, and Ramiro Berardo. 2021. "Methodological Approaches to Studying Polycentric Governance with the Ecology of Games Framework." In *Methods of the Policy Process*. Milton Park: Routledge.

May, Peter J., Samuel Workman, and Bryan D. Jones. 2008. "Organizing Attention: Responses of the Bureaucracy to Agenda Disruption." *Journal of Public Administration Research and Theory* 18 (4): 517–541. https://doi.org/10.1093/jopart/mun015.

Sabatier, Paul A. 1991. "Toward Better Theories of the Policy Process." *PS: Political Science and Politics* 24 (2): 147. https://doi.org/10/fcb8j6.

Sabatier, Paul A. 1999. "The Need for Better Theories in Theories of the Policy Process." In P.A. Sabatier (ed.) *Theories of the Policy Process* (pp. 3–17). Boulder, CO: Westview Press.

Schlager, Edella, Saba Siddiki, and Michael Cox. 2021. "Methods for Analyzing Social Dilemmas and Institutional Arrangements within the Institutional Analysis and Development Framework." In *Methods of the Policy Process*. Milton Park: Routledge.

Simon, Herbert. 1996. *The Sciences of the Artificial*, Third Edition. Cambridge, MA: The MIT Press.

SoRelle, Mallory, and Jamila Michener. 2021. "Methods for Applying Policy Feedback Theory." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First Edition. Milton Park, UK: Routledge.

Stromberg, Joseph. 2013. "What Is the Anthropocene and Are We in It?" *Smithsonian Magazine*. www.smithsonianmag.com/science-nature/what-is-the-anthropocene-and-are-we-in-it-164801414/.

Tosun, Jale, and Samuel Workman. 2018. "Struggle and Triumph in Fusing Policy Process and Comparative Research." In Christopher M. Weible (ed.) *Theories of the Policy Process* (pp. 329–362). Milton Park: Routledge.

Weible, Christopher M. 2018. "Introduction: The Scope and Focus of Policy Process Research and Theory." In Christopher M. Weible and Paul A. Sabatier (eds) *Theories of the Policy Process* (pp. 1–13). Milton Park: Routledge.

Weible, Christopher M., and Paul A. Sabatier (eds). 2018. *Theories of the Policy Process*. Milton Park: Routledge.

Workman, Samuel. 2015. *The Dynamics of Bureaucracy in the U.S. Government: How Congress and Federal Agencies Process Information and Solve Problems*. Cambridge: Cambridge University Press.

Workman, Samuel. 2020. "Four Principles of Data Collection." *Towards Data Science*, May 12, 2020. https://towardsdatascience.com/four-principles-of-data-collection-82ad07938ac1.

Workman, Samuel, Frank R. Baumgartner, and Bryan D. Jones. 2021. "The Code and Craft of Punctuated Equilibrium." In *Methods of the Policy Process*, First Edition. Milton Park: Routledge.

2 How to Conduct a Multiple Streams Study

Reimut Zohlnhöfer, Nicole Herweg, and
Nikolaos Zahariadis

Introduction

The Multiple Streams Framework (MSF) is one of the most frequently applied approaches to study the policy process. Recent reviews (Jones et al. 2016; Rawat and Morris 2016) have found literally hundreds of empirical applications. One reason for the success of the MSF is probably its easy accessibility and the high *prima facie* plausibility of its concepts and basic ideas. The metaphors in Kingdon's (1984) book (streams, primeval soup, etc.) "are very powerful and have great purchase" (Béland and Howlett 2016, 224) because they help to get an idea of how policy processes work.

Nonetheless, the framework's strength is also its main weakness because these metaphors have not been translated into a commonly shared understanding, definition, and operationalization in empirical MSF studies. On the contrary, critics stress that a great part of the MSF literature is characterized by isolated case studies that barely talk to each other and, consequently, hinder knowledge accumulation (Cairney and Jones 2016). Worse still, while scholars "use the same vocabulary they do not all share the same definition of concepts" (Jones et al. 2016, 30). The bottom line is that too many empirical MSF studies lack a conceptual and methodological foundation that would allow assessing the framework's explanatory potential.

This chapter aims at providing readers with the means to conduct a reliable MSF study. The question of how MSF hypotheses should be tested and how individual concepts should be operationalized has received hardly any scholarly attention (as an exception, see Engler and Herweg 2019 for quantitative MSF studies). We are not aware of any papers that deal with the methods of MSF case studies and provide suggestions on how to conduct them. Hence, in this chapter, we report what we consider best practices from the literature while in some instances we also suggest new paths. Therefore, it is to some extent explorative and we hope to spark a debate about how best to conduct MSF research. Moreover, due to space restrictions we focus on how to analyze agenda-setting from an MSF perspective only. We are confident that scholars will be able to transfer our methodological suggestions to the analysis of decision-coupling.

We proceed by discussing what to consider when developing a research question (Section "Develop Your Research Question(s)"), how to translate a research question into MSF hypotheses (Section "Specify MSF Hypotheses"), how to operationalize these hypotheses (Section "Operationalization Issues"), and, finally, how to test them empirically (Section "Choose the Method that Best Matches Your Research Interest"). We conclude with suggestions for future MSF research designs (Section "The Way Forward to Empirically Sound MSF Studies").

From Theory to Application: Steps to a Solid MSF Analysis

Develop Your Research Question(s)

Kingdon (1984, 3) starts from the following research questions: "[W]hy [do] some subjects become prominent on the policy agenda and others do not, and why [are] some alternatives for choice […] seriously considered while others are neglected?" At the heart of these questions are two different dependent variables: the selection of alternatives and the actual agenda change.

However, since its initial publication, the MSF has been extended to further policy stages, such as decision-making or implementation. Resulting from these extensions, the MSF also covers the following research questions: Why does a policy get adopted? Does the policy adopted differ from the original proposal, and if so, why? Which factors make it more probable that policy implementation is successful? Consequently, MSF analyses can also choose policy change or the implementation output as dependent variables.

Finally, researchers might focus on specific elements of the framework, for instance, on policy entrepreneurs' success metrics or how the activities of problem brokers (Knaggård 2015) in the problem stream influence problem recognition. Another example concerns the causes or consequences of open policy windows. These foci go along with different dependent variables, namely policy entrepreneurs, problem recognition, and the occurrence of policy windows.

While it is possible to conduct MSF-guided analyses explaining any of the above-listed dependent variables, it may become necessary to adjust the form or structure of causal mechanisms that link the dependent variable with one or more independent variables. The causal mechanisms that apply for agenda change and the selection of alternatives cannot be transferred one to one to policy adoption, for example. There are a number of helpful papers discussing these issues (for an overview, see Herweg, Zahariadis, and Zohlnhöfer 2018), so scholars do not need to start from scratch, but it may not always suffice to start out from Kingdon's landmark book alone.

Next to deciding which dependent variable to analyze, researchers must specify their unit and level of analysis. Remember that Kingdon (1984) derived his ideas inductively from observations of agenda change at the federal level of the United States, so his generalized propositions are tailored

to explain agenda-setting in this specific institutional setting. Consequently, applying it outside the framework's initial scope requires addressing how this transfer affects the MSF's setup. One important aspect is the difference between democracies and non-democracies. Researchers analyzing non-democracies must address how the framework's key concepts need to be adapted before conducting an MSF-led analysis. For example, the national mood or balance of interest groups may not be as important in China (but see Liu and Xu 2021). Instead, some propose "political attention" (Mu 2018, 4–5), understood as the policy goals of central government and the institutional context they create in terms of goal conflict and hierarchical authority.

However, even when applying the MSF to democracies other than the US, researchers should critically assess whether different institutional contexts require modifying the framework. Herweg, Huß, and Zohlnhöfer (2015), for instance, provide recommendations on how to factor differences between presidential and parliamentary systems in the MSF and highlight the prominent role of political parties in the policy and the political stream of parliamentary democracies. Similarly, Sanjurjo (2020b) discusses in detail which adjustment needs arise if researchers apply the MSF to Latin American politics.

By the same token, applying the MSF to the non-national level of analysis (e.g., subnational or inter/supranational level) may require adaptations. For instance, if the research question deals with agenda change in the European Union, the political stream may need adjusting. Again, these issues have been discussed in the literature already (e.g., Herweg 2016b, Zahariadis 2008).

In a nutshell, developing an MSF research question includes (i) discussing explicitly whether and where the framework requires adaptations, (ii) identifying functional equivalents of key concepts, and (iii) spelling out how these concepts are connected causally (see Box 2.1). In the next section, we focus on the latter challenge: Formulating MSF hypotheses.

Box 2.1 Developing your MSF research question(s)

(1) Specify your dependent variables

- Selection of alternatives
- Agenda change
- Policy change
- Implementation output
- MSF key elements (e.g., activities of policy entrepreneurs, occurrence of policy windows)

(2) Specify your unit(s) of analysis

- Democracies
- Non-democracies
- Else (e.g., international organizations, policy communities, …)

(3) Specify your level of analysis. Levels of analysis the MSF has already been applied to:

- National
- Subnational
- Supranational
- Else (e.g., parties)

(4) If you apply the MSF outside its original scope, explicitly deal with the following issues:

- Discussion of adaptation requirements
- If required, identification of functional equivalents of the framework's key concepts
- If required, statement how these key concepts are connected causally.

Specify MSF Hypotheses

Specifying testable hypotheses is fundamental for every good research program (King, Keohane, and Verba 1994). Hypotheses help empirical research in at least two respects. First, hypotheses structure empirical analyses. They define what the relevant actors and causal mechanisms are and how they interact to produce a specific outcome (e.g., a stream becoming ready for coupling or agenda-change occurring). Thus, they help researchers avoid simple storytelling by checking whether the theoretically derived expectations can be detected in the empirical material. Second, hypotheses – together with appropriate operationalizations – allow researchers to define in advance which evidence would lead them to conclude that the theory is corroborated – and, more importantly even (King, Keohane, and Verba 1994), which observations would lead us to reject an expectation of a theoretical framework. Thus, formulating hypotheses is an important instrument to avoid confirmation bias by presenting the material in a way that allows corroborating the theory.

Much research in the MSF tradition indeed has failed to produce explicit hypotheses. This may have to do with the fact that the vast majority of relevant studies were studies of individual cases (Jones et al. 2016; Rawat and Morris 2016). It is at least open to debate whether or not it makes sense to formulate explicit hypotheses in these kinds of studies because a single case cannot be used to reject a (nondeterministic) hypothesis. Neither would the success of a hypothesis in a single case increase our trust in that hypothesis substantially. Nonetheless, for the reasons discussed above, even single case studies can benefit from formulating explicit hypotheses (or, to be more humble: theoretical expectations). Being explicit about theoretical expectations does not only promise benefits for the individual case study, though. It may also allow aggregating the findings of the individual cases:

While we might not be very impressed if a single case study corroborates a specific hypothesis, it may be much more convincing if the same hypothesis is confirmed empirically more frequently.

While most MSF-related studies have abstained from using explicit hypotheses (Jones et al. 2016), there is nothing in the MSF that would keep us from constructing hypotheses from the framework. To the contrary, while some earlier studies have pioneered the deduction of MSF hypotheses (e.g., Blankenau 2001; Boscarino 2009), in the recent literature, a number of generic hypotheses, i.e., hypotheses that are not specific to individual cases, have been suggested (Herweg, Huß, and Zohlnhöfer 2015; Herweg, Zahariadis, and Zohlnhöfer 2018). Among these, the core hypothesis, which sums up the central idea of the MSF, is the following: "Agenda change becomes more likely if (a) a policy window opens, (b) the streams are ready for coupling and (c) a policy entrepreneur promotes the agenda change" (Herweg, Zahariadis, and Zohlnhöfer 2018, 30). Other hypotheses deal with developments in the individual streams, about the policy window or the policy entrepreneur. Box 2.2 summarizes the generic hypotheses on agenda-setting. Other hypotheses discuss decision-coupling, i.e., the processes taking place after the decision agenda has changed (Herweg, Zahariadis, and Zohlnhöfer 2018). Still others are more detailed, discussing individual elements of the framework like the relation between the structure of the policy community and the pace and scope of new policy proposals in the policy stream (Zahariadis 2003).

In sum, scholars should definitely deduce hypotheses from the MSF in their empirical work – even if they are only conducting single case studies. Hypotheses will help prevent storytelling and a confirmation bias in the study. Moreover, we encourage researchers to employ somewhat established hypotheses that have been suggested in the literature already. We do not deem established hypotheses necessarily to be better suited or empirically more appropriate than hypotheses each individual researcher could come up with. Rather, to develop MSF further and refine the framework and the current hypotheses, it makes sense to start from established expectations, which researchers can amend or modify in the light of new empirical data. The point is to replicate hypotheses in similar environments or to expand/adapt (or not) their applicability in unexpected settings. In either case, the hypotheses acquire greater analytical weight and promote knowledge accumulation. This will allow the scholarly community to define scope conditions for some hypotheses and refine others.

Box 2.2 MSF hypotheses on agenda-setting

Hypothesis for the framework as a whole

Agenda change becomes more likely if (a) a policy window opens, (b) the streams are ready for coupling, and (c) a policy entrepreneur promotes the agenda change.

Hypotheses for the framework's key elements
Problem stream
A problem broker is likely to be more successful in framing a
  condition as a problem, the more an indicator changes to the
  negative, the more harmful a focusing event is and the more
  definite a government program does not work as expected.

Political stream
Policy proposals that fit the general ideology of a government or
  majority in legislature have a better chance of gaining agenda status.

Policy stream
If a policy proposal does not fulfill the selection criteria, the
  likelihood of gaining agenda status, and thus being coupled,
  decreases significantly.
As the integration of policy communities decreases, it becomes more
  likely that entirely new ideas can become viable policy alternatives.

Policy window
The policy window opens in the problem stream due to at least one
  of the following changes: change of indicators, focusing events, or
  feedback.
The more a condition puts a policy-maker's reelection at risk, the
  more likely it is to open a policy window in the problem stream.
The policy window opens in the political stream due to at least one
  of the following changes: changes in legislature, election of a new
  government, a change in the national mood.

Policy entrepreneur
The policy entrepreneur is more likely to successfully couple the
  streams during an open policy window, (a) the more access to core
  policy-makers she has and (b) the more persistent she is.

Source: Herweg, Zahariadis, and Zohlnhöfer (2018, 30).

Operationalization Issues

Because many key MSF concepts are not clearly defined in Kingdon's (1984) landmark book, they lack analytical precision. Hence, scholars who wish to apply the MSF empirically need to pay particular attention to the definition of the key concepts and to specifying and justifying their respective operationalizations. We discuss the key MSF concepts in turn, starting with the three streams and then looking at the policy window, policy entrepreneurs, and coupling. Note, however, that it is neither possible nor necessary to collect all of the data mentioned below for every research project. Time and money are scarce resources for all scholars, and the word limits

of journal articles often do not allow for presenting these data. Therefore, in many cases, the secondary literature may provide usable data for some MSF elements. Other cases may be so unambiguous (e.g., 9/11 as a focusing event for law-and-order projects) that it will not be necessary to put a lot of effort into further data collection. Nonetheless, depending on their research question, scholars should consider the variables we discuss in the following.

The Problem Stream

According to the MSF, the problem stream consists of all conditions that policy-makers or people around them perceive as problematic. To capture policy-makers' attention, conditions must become known to them. According to MSF thinking, this happens via one or a combination of the following attention generating mechanisms: indicators, focusing events, and feedback. Moreover, and particularly important for our purposes, there are no objective problems but only socially constructed ones. Consequently, there need not be a strong correlation between "objective" problem indicators and developments in the problem stream.

To make it on the agenda, conditions must be framed as public problems. This frame has to be popularized and policy-makers must be convinced that certain conditions deserve attention and possible action. In this regard, the problem broker is pivotal. The problem broker is defined as an actor who "frame[s] conditions as public problems and work[s] to make policy makers accept these frames" (Knaggård 2015, 450). It is important to keep in mind that problem brokers (and conditions looking for political attention) are in competition for policy-makers' attention. Hence, scholars should consider the conditions under which a problem makes it on the agenda.

To operationalize the problem stream empirically, researchers must check whether indicators, focusing events or feedback draw policy-makers' attention to a specific condition. Which kind of attention generating mechanism is relevant, varies with the research question and policy field. Indicators can direct policy-makers' attention to an issue if the level of an indicator changes markedly (e.g., a rise in the housing affordability index draws attention to the state of the housing market; see Tiernan and Burke 2002). Second, focusing events include disasters (e.g., floods, see O'Donovan 2017), crises (e.g., the 2003–2005 Stability and Growth Pact crisis, see Saurugger and Terpan 2016), personal experiences (e.g., importance of obesity-related personal health problems of leading politicians for evaluating childhood obesity as problematic, see Craig et al. 2010), and symbols (e.g., Knut, the polar bear born in the Berlin Zoo, was used as a symbol for global warming; see Brunner 2008). The third path to generate attention, feedback, can occur via formal channels such as systematic monitoring and evaluation studies or informal channels such as feedback from the bureaucracy.

Nonetheless, as suggested above, problems usually contain a "perceptual, interpretive element" (Kingdon 1984, 115). Therefore, only pointing to some "objective" problems will not always be enough; rather, researchers will need to demonstrate that actors were aware of a condition and that they

indeed thought it necessary to deal with this problem politically. Potential sources for these kinds of perceptions are interviews with key actors, their speeches or memoirs, but also the minutes of specific commissions (for the latter, see Sanjurjo 2020a).

In some cases, problems are so large that they "simply bowl over everything standing in the way of prominence on the agenda" (Kingdon 1984, 101); this is usually reflected in the salience of the issue among the electorate (at least in a democracy) at a given point in time. Likewise, if a government runs the risk of failing to attain a major goal (e.g., a core election promise), the situation will likely be viewed as problematic. In instances like these, even in the absence of an explicit quote from the relevant policy-makers, researchers can argue that doing nothing was not an option and that policy-makers were highly likely to have perceived the respective condition as a problem. In these cases, researchers should at least explicate which circumstances make it likely that policy-makers considered a condition a policy problem.

In some, above all in quantitative MSF applications, it will not always be possible to use data that reflect policy-makers' views of a condition as a problem. So, existing indicators like unemployment rates, poverty rates, or greenhouse gas emissions in principle can be used. The choice of an "objective" indicator becomes less convincing, however, the less politically salient and the less prominent the respective indicator is. Unemployment, for example, is usually a highly salient indicator. Hence, it is likely that governments may want to respond to an increase in joblessness. Biodiversity, however, may be a much less obvious problem indicator in many contexts. Therefore, authors should justify their choice of problem indicator by arguing why it is plausible that policy-makers should have perceived the specific indicator as indicating a policy problem to which they should attend. DeLeo and Duarte's (2021) article on the US opioid crisis provides an excellent example: the authors show that the indicators they choose are the only ones the relevant government department reports on its website and corroborate their choice using expert interviews.

Alternatively, researchers could follow Engler and Herweg's (2019) suggestion to look at media coverage about specific issues. The more the media speak about an issue, the more likely it should become a public problem. The same could be said about public opinion in the form of survey questions about the most important problem in a country. These data are publicly available at least for some countries, for example, in the Comparative Agendas Project. Media data can also be self-coded with the help of computer software (Soroka and Wlezien 2019).

If researchers examine how different issues compete for policy-makers' attention, they must explain why one specific problem received attention and not another. Explanations could refer, for instance, to the electoral relevance or the magnitude of the problem. Again, data on media reporting or opinion polls on the most important problems could be used to specify which other issues competed with the condition under study at a specific point in time and how much space was left for the issue of interest to the researcher.

For example, Sanjurjo (2020a), in his study on gun control policies in Latin America, used surveys on the most important problems to show at which points in time the issue of internal security was high on the public agenda and thus required policy-makers' attention. If scholars are investigating very prominent conditions like the economic situation, unemployment, etc. or very important focusing events (9/11, Covid-19 pandemic), it may be less important to point out that there was enough agenda space for the issue. If, however, one is looking into forest or agricultural policy, that is, issue areas that are rarely in the public's eye, one should make plausible that there was agenda space left for the issue. Hence, to analyze the problem stream researchers must document which indicators, focusing events, and/or feedback attracted policy-makers' attention to conditions which at least one problem broker managed to frame as problematic. If the analysis reveals that there is such a condition, the problem stream is ready for coupling. When we discuss policy windows, we address how a problematic condition leads to the opening of a problem window.

The Policy Stream

When analyzing the policy stream, researchers must identify the relevant policy community, its members and structure, the ideas it generates, and the softening-up process these ideas must survive.

A policy community "is mainly a loose connection of civil servants, interest groups, academics, researchers and consultants (the so-called hidden participants), who engage in working out alternatives to the policy problems of a specific policy field" (Herweg 2016a, 132). Herweg (2016a, 132) suggests that an "actor's engagement in working out policy alternatives" should serve "as a distinction criterion for being considered a policy community's member." If a person contributes ideas that other members are aware of, (s)he can be regarded a member of the policy community. Given the loose connection between the members, most scholars define the policy community only vaguely. For instance, in his analysis of urban road pricing, Dudley (2013) refers to those actors that propose solutions to road congestions as members of the policy community. In other cases, scholars were able to delineate policy communities more precisely. For example, the European gas policy community analyzed by Herweg (2016a, 2017) was to a substantial degree created by the European Commission. Consequently, a specific forum, the Madrid Forum, existed in which the policy community started exchanging views. Münter (2005), in his book-length account of British devolution policies in the 1970s and 1990s, looks at very different sources. For the (failed) attempt to devolve competencies in the late 1970s, he compiled the policy community's members and their proposals by analyzing various green books, white books, reports of party commissions, and bills. In contrast, in the 1990s, all proponents of devolution for Scotland assembled in the Scottish Constitutional Convention. Therefore, Münter uses its proceedings and final report to identify the respective policy community and to reconstruct the softening-up process. If no such focal points for the deliberations

can be identified, scholars could start by collecting relevant policy proposals (e.g., in parliamentary hearings) and follow the interactions between the various actors. Ideally, researchers can apply a discourse network analysis (Leifeld 2020), for example, based on newspaper articles or statements in social media, in order to analyze systematically the relations between the members of the policy community.

Alternatively, researchers could rely on expert surveys (see Novotný, Satoh, and Nagel 2021). Identifying key actors in the policy stream should pose no problem and researchers could then ask these actors with whom they interacted, whom they sent information, and from whom they received relevant information. This procedure does not only allow identifying all relevant actors (and thus a policy community's size) but also enables scholars to assess the centrality of actors in the policy stream. Those people who are named frequently are likely to be more important than those who are mentioned only a few times. Moreover, scholars can employ these data in a policy network analysis to explore the policy community's structure, i.e., the degree of a policy community's integration: Is everybody talking to everybody else or are there two or more camps? Nonetheless, this procedure is highly resource intensive and its success critically depends upon the willingness of the policy community's members to answer the survey. Therefore, the procedure is unlikely to be applicable in single-authored studies, in studies that seek to apply the whole framework (rather than just the policy stream) and in comparative studies. In these cases, a discourse network analysis based on publicly available written sources is probably preferable.

Once the policy community's members and its structure are identified, scholars must specify which ideas the policy community has developed and analyze the softening-up process, i.e., the process by which policy proposals are discussed, rejected, or amended. Relevant sources, which may help reconstruct the softening-up process, are policy papers, minutes of committees or working groups, publications in specialized periodicals, committee hearings, newspaper coverage in salient cases, and expert interviews with individual members of the policy community.

Note that it may not be necessary to trace specific ideas back to their very origin. Sometimes it is important to follow the development of policy alternatives in a detailed fashion (e.g., to understand why a specific proposal popped up), while at other times it may suffice to contend oneself with establishing that certain ideas fulfill the criteria of survival. MSF expects that when ideas meet these criteria, they are more likely to become viable policy alternatives. The criteria include technical feasibility, value acceptability, tolerable costs, and receptivity among decision-makers.

Researchers should assess empirically to what extent policy proposals satisfy the criteria of survival from the perspective of the policy community's members. This can be done by analyzing the discourse among policy experts. If, for example, policy experts in a parliamentary hearing do not question that a policy is technically feasible, we can conclude that it fulfills that specific criterion. Moreover, sometimes scholars can assume the fulfillment of

specific criteria present without going into the details of the softening-up process (unless there is evidence to the contrary). For example, if the policy under study is regulatory in nature, we can assume that costs are tolerable. Similarly, if there is clear evidence that certain policy-makers have signaled their interest in a proposal, decision-makers' receptivity is present without being explicitly stated in the policy community's exchanges.

Value acceptability is difficult to ascertain, especially if some in the policy community are skeptical while others are enthusiastic. There is no consensus on how large the support for a proposal has to be among the policy community, largely because the answer varies by community. In more integrated communities, actors are included in the process and may even act as veto players (Zahariadis 2003; Zohlnhöfer, Herweg, and Huß 2016). Side payments or institutional manipulation to garner political support are reasonable indicators of value acceptability (Zahariadis 2021).

It is clear that a proposal does not need unanimous (or even majority) backing in the policy community to be considered a viable policy alternative. In principle, the policy community can be divided between two or even three proposals. But when can these proposals be regarded as viable policy alternatives? When does the softening-up process come to a (temporary) end? We suggest that this is the case when most of the policy experts have made up their minds and back one proposal. If many policy experts are undecided and do not explicitly back one policy proposal, no viable policy alternative exists and coupling should be difficult. If, however, many members of the policy community have endorsed one (although not necessarily the same) proposal, we can assume that the softening-up process has produced at least one worked-out alternative that is ready for coupling. If more than one proposal is supported by a sizable share of the policy community, two (or more) viable options exist.

The policy stream can be considered as ready for coupling if at least one alternative is available that meets the criteria of survival.

The Political Stream

The political stream consists of three elements in democracies: the government and parliament, interest groups, and the national mood. The guiding question is whether these elements support the rise of an issue on the governmental agenda. We will discuss how to operationalize the elements of the political stream in turn.

Although policy-makers are assumed to have unclear policy preferences, the ideological affiliation of politicians may provide researchers with a rough idea of a proposal's political appeal. Relatedly, in many cases, parties or policy-makers "own" certain issues, i.e., voters tend to view specific actors as particularly competent in dealing with specific issues and the issue owners are likely to benefit electorally if "their" issues become prominent on the agenda (Petrocik 1996; Seeberg 2017). Hence, elected policy-makers can be expected to be more receptive to deal with issues that they or their parties

own or that are in line with their party's policy program. Travis and Zahariadis (2002), for instance, assume that in the US, Democrats are more inclined to spend money in foreign aid programs than Republicans. Therefore, they operationalize this element by looking at the president's party and the party that holds the majority in the Senate.

Nonetheless, precisely because policy-makers' policy preferences are assumed to be unclear, their ideological orientation is often only a rough predictor of appeal. The same proposal can be framed very differently, making it compatible with very different ideological orientations. For example, we can think of cuts to unemployment benefits as being closer to the ideology of conservative rather than social-democratic governments. If, however, a policy entrepreneur can convince a social-democratic government that the cut is necessary in order to deal with another problem (e.g., unemployment – an issue these parties "own"), then that government might be willing to adopt the cut despite its ideological position. Thus, researchers have to look closely at the positions of parliamentarians and ministers and the framing of issues when policy-makers support policies that do not align with their ideological orientation at first glance.

Whereas data on the composition of government and parliament are readily available, it is less clear how to operationalize interest group receptivity to agenda change. In most instances, it is quite evident which interest groups show an interest in a specific issue and researchers can gather their positions by analyzing parliamentary hearings or press coverage. Unless interest group activities all point in the same direction, however, policy-makers' perceptions of the balance of support are relevant (Kingdon 1984). In MSF thinking, this balance is not an objective fact. Rather, actors in and around government will consider whether there is enough support for their proposal among interest groups to have a reasonable chance of getting it on the agenda or whether interest group opposition is so strong that the project is doomed to fail.

But how can we measure the balance of support? Policy-makers' perceptions of this balance among interest groups can be read off their memoirs, interviews in the press, or expert interviews (Staff 2020). If, however, the issue is not taken up in memoirs or interviews, scholars can also try to assess the balance of support and assume that policy-makers will have calculated similarly. Resources and intensity are particularly relevant (Kingdon 1984). Regarding resources, interest groups that have many members, spend much money in election campaigns or rally many supporters for demonstrations (Sanjurjo 2020a) can potentially mobilize many voters and hence count as important. The same is true for representatives of economically vital sectors.

Regarding intensity, policy-makers are more likely to listen to interest groups that constantly and extensively talk about an issue than to groups that only occasionally come forward with their demands. One could operationalize this element of the political stream with data on the number of press releases or social media tweets or posts on a specific issue.

In corporatist systems, interest groups that are represented in the committees where policy projects are negotiated will be regarded as more

important than outsiders. On the informal level, some interest groups have close relations to officials or former representatives of an interest group are now MPs or civil servants. If such links exist, these interest groups should also be considered influential.

The third element of the political stream is the national mood, which is defined as "a rather large number of people out in the country (…) thinking along certain common lines" (Kingdon 1984, 153). It is a very challenging concept to measure. Similar to interest group activities, what matters is how policy-makers perceive the national mood. Consequently, Kingdon (1984) warned that the national mood is not necessarily congruent with opinion polls. Thus, preferably, researchers should rely on information of policy-makers' perceptions of the national mood. For example, in his study on German labor market reforms, Zohlnhöfer (2016) discusses the relevant actors' perceptions of the national mood with the help of several sources, mostly memoirs, and is able to show that key actors misjudged the national mood. Dolan (2021) performs a quantitative media analysis to assess how the Australian Millennium Draught affected the national mood. Other possible sources for how actors perceived the national mood are interviews with the relevant actors – either press interviews (preferably during the policy process, i.e., without the benefit of hindsight) or interviews that are conducted by the researcher herself. When such data collection is too expensive or time-consuming to obtain, we propose using public opinion data. Given the professionalization of policy-makers in recent decades, we argue that Kingdon's original warning against the use of survey data is too strict because policy-makers often commission (and make use of) opinion polls themselves (Herweg, Huß, and Zohlnhöfer 2015). Therefore, we can presume that these survey data will shape policy-makers' perceptions of the national mood. Indeed, some recent MSF contributions have used opinion poll data to operationalize the national mood (Cook and Rinfret 2013; Dolan 2021, 13; Sanjurjo 2020a; Tiernan and Burke 2002).

Trickier than operationalizing the political stream's elements is specifying when the political stream is ready for coupling. The political forces comprising the political stream might point in opposite directions: For example, the national mood might be favorable while the majority of powerful interest groups is opposed to a proposal. According to recent research (Herweg, Huß, Zohlnhöfer 2015; Herweg, Zahariadis, Zohlnhöfer 2018), government and parliament should be considered the most relevant actors in the political stream – who, however, might be influenced by the national mood and interest group campaigns. Government and parliament are assigned a more prominent role simply because these are the actors that are responsible for adoption of a proposal. Nonetheless, the relative importance of the elements of the political stream could also depend on the salience of an issue. If a topic is highly salient, the national mood is likely to be influential (see Zahariadis 2015). If, in contrast, the public and media more or less ignore an issue, the role of interest groups becomes more important (Kingdon 1984).

When is the political stream ready for coupling? At the agenda-setting stage, it is not yet necessary that parliamentary majorities are forthcoming. What is necessary, however, is that a key policy-maker like the responsible minister or an influential member of parliament, i.e., a political entrepreneur (Roberts and King 1991, 152), actively supports the proposal and is willing to bring together a majority for it during decision-making (cf. Herweg, Zahariadis, and Zohlnhöfer 2018; Zohlnhöfer 2016). If such a political entrepreneur exists, the political stream is ready for coupling.

Policy Window

A policy window constitutes "an opportunity for advocates of proposals to push their pet solutions, or to push attention to their special problems" (Kingdon 1984, 173). It can open in the problem stream or in the political stream.

For a policy window to open in the problem stream (problem window), two conditions must be met: First, the problem stream is ready for coupling. Second, policy-makers must deem the issue important and relevant. Herweg, Huß, and Zohlnhöfer (2015) argue that conditions that put a policy-maker's reelection at risk are more likely to gain her scarce attention, i.e., an issue's relevance can be read off its chances to make a difference at the next election. To capture this second condition empirically, researchers must make plausible that policy-makers believed that their reelection was threatened. For instance, Zohlnhöfer (2016), deriving evidence from people involved in the policy process, makes it plausible that the German government under Gerhard Schröder believed that the stubbornly high level of unemployment would endanger its reelection chances and that labor market reforms would alleviate the problem. Similarly, Dolan (2021, 177) shows that the serious threat of being ousted triggered Prime Minister Howard's response to the Australian Millennium Draught.

A policy window opens in the political stream (political window) if the composition of government or parliament changes or the national mood shifts. The third element of the political stream, interest groups, is not relevant for the opening of a policy window because interest groups focus on blocking or adapting proposals that are already on the governmental agenda and do not tend to set agendas on their own (Kingdon 1984). Thus, the opening of a political window can be measured via electoral changes (e.g., election of a new government, change of single ministers, new MPs) or changes in the national mood.

Note, however, that it is not always evident if a specific development in the problem or the political stream constitutes a policy window for a specific policy project. While some policy windows like a change of government are easy to identify and can be employed for many different policy projects, other policy windows are less obvious and need justification to confirm skillful exploitation by policy entrepreneurs. Moreover, some potential policy windows go unnoticed (or at least unused), while in other cases policy

entrepreneurs believe that a policy window had opened but fail to bring about agenda change. Researchers should take advantage of these negative cases to specify the conditions that make a policy window promising for entrepreneurial coupling attempts.

Policy Entrepreneurs and Coupling

Policy entrepreneurs couple the streams during open policy windows. According to Kingdon (1984, 189), "one can nearly always pinpoint a particular person, or at most a few persons, who were central in moving a subject up on the agenda and into position for enactment." How do researchers actually pinpoint these persons? Since it is irrelevant where policy entrepreneurs are located (e.g., inside or outside the government or administration), researchers identify policy entrepreneurs via their activities. More precisely, they look for actors who invest time, energy, reputation, or money to move an issue higher on the agenda.

For instance, in their analysis of greenhouse gas emission regulation in the US, Cook and Rinfret (2013) identify Lisa Jackson, the Environmental Protection Agency (EPA) administrator, as a policy entrepreneur. They did this by comparing the number of public speeches and addresses on climate change she had given with that of her predecessors (hers being tremendously higher). Furthermore, the authors documented that Jackson dealt with climate change since the beginning of her term in office and that she urged the EPA to push the issue forward. Similarly, Staff (2020), in his UK case study on security privatization, identifies the Labour MP Bruce George as the policy entrepreneur due to his long-term engagement in the issue area but also thanks to evidence from expert interviews.

While identifying one or several policy entrepreneurs is important, operationalizing their characteristics and strategies is essential. Favorable characteristics are persistence, political well-connectedness, access to policy-makers, and negotiating skills.

Following Kingdon (1984, 190), persistent policy entrepreneurs

> spend a great deal of time giving talks, writing position papers, sending letters to important people, drafting bills, testifying before congressional committees and executive branch commissions, and having lunch, all with the aim of pushing their ideas in whatever way and forum might further the cause.

Bernd Buchheit, the policy entrepreneur for German labor market reforms, is a prime example of an actor who went out of his way to get his pet proposal on the agenda (Zohlnhöfer 2016).

Political connectedness refers to policy entrepreneurs' relational profile in the policy community. If policy entrepreneurs hold a central position, they are more likely to be able to influence others. One advanced way to measure whether a policy entrepreneur is politically well-connected is to conduct a

network analysis (cf. Novotný, Satoh, and Nagel 2021). Alternatively, his/her formal role may help. Returning to the above examples, the EPA administrator is a key policy-maker with formal decision-making competences. Similarly, MPs have easy access to policy-makers. Resulting from their daily business, they are personally acquainted with other key decision-makers, interest groups representatives, or journalists. Consequently, they are politically well-connected. It is far more demanding to measure how well a policy entrepreneur is connected if (s)he belongs to the hidden participants. In that case, interviews with relevant actors might be informative. Actors mentioned by many interviewees are likely better connected than actors only few interviewees mention.

Access to policy-makers refers to having a claim for a hearing. This claim results from at least one of the following sources: "expertise; an ability to speak for others, as in the case of the leader of a powerful interest group; or an authoritative decision-making position, such as the presidency or a congressional committee chairmanship" (Kingdon 1984, 189).

Negotiating skills capture how competently policy entrepreneurs push for their favorite proposal in negotiations with other key actors. Measuring this skill directly is rather challenging. An indirect way to measure it is to analyze how well policy entrepreneurs make use of strategies to couple the streams. Strategies they have at their disposal are, for instance, framing of a problem, affect priming, "salami tactics," and the use of symbols (Zahariadis 2003, 14).

One of the core activities of policy entrepreneurs is coupling the three streams when a policy window opens. But what do we observe when the streams are coupled? Most MSF case studies deal with the coupling process either as a black box or only implicitly, i.e., the authors observe that the streams are ready for coupling and a policy window opens on the input side and conclude that coupling has taken place if they observe agenda change on the output side (cf. Dolan 2021). In contrast, recent studies provide some first ideas of empirically observing coupling. One example is Sanjurjo's (2020a) study of gun control policy in Latin America. In his case study on Brazil, for instance, the author shows that once the streams were ready for coupling, the policy entrepreneurs started arguing that their policy project – disarmament (policy stream) – would be a good solution for the country's notorious problems of violence and homicides (problem stream). Next, they sought for political support via media campaigns and used their access to key policy-makers at the state and federal levels (political stream) to get the policy on the agenda. Similarly, Staff (2020), in his study of private security regulation in the UK, shows how a policy entrepreneur had been advocating his pet project, the regulation of the private security industry, for decades (policy stream). Persistently championing his project in communications with potential political entrepreneurs, he eventually was able to convince a political entrepreneur that the regulation of the industry would help solve the problem of criminality (problem stream), which was a high priority for the incoming Labour government (political stream). Hence, the issue came

on the agenda although the incoming Home Secretary was not particularly interested in the project.

To describe and explain the coupling process, data from various sources need to be gathered. In the two examples above, the authors used interviews with people who were actively involved in the policy processes, policy papers from the policy entrepreneurs, but also the secondary literature.

Choose the Method that Best Matches Your Research Interest

Having been derived inductively, it is unsurprising that most MSF studies are qualitative case studies (Jones et al. 2016). However, depending on the research interest, a quantitative MSF analysis might be the better choice as medium- to large-N analyses would add weight to the MSF's explanatory power. This section draws attention to methodological issues.

Qualitative MSF Research

Having decided on the research question that is best answered by way of qualitative research and having selected cases and time period (Gerring 2017; Seawright and Gerring 2008), we focus on issues of case study protocol. How does one conduct an MSF qualitative application?

MSF case study designs have to contain variance in the dependent variable. Hence, scholars should not content themselves with investigations of successful couplings. Instead, analysis of failed couplings is important to test the MSF's core hypothesis. For example, Venters, Hauptli, and Cohen-Vogel (2012) analyze how the state of the political stream prevented agenda change regarding the introduction of a national sales tax for education in the US during the Nixon administration. Similarly, Zahariadis (1996) investigates the attempts at privatization of British Rail by focusing on four political windows and two problem windows between 1974 and 1992. The comparative analysis of failed and successful couplings reveals how important the positions and strategies of policy entrepreneurs are. Other studies that look at failed couplings include Münter (2005) and Sanjurjo (2020a).

In light of the well-known problem of "too many variables, too few observations," single cases may increase observations by lengthening the time component (e.g., Zahariadis 2005) or compare cases across units and over time within and across cases (e.g., Zahariadis 1995). Finally, it is important to remember that the number of observations that can be used to assess the MSF is usually higher than the number of cases, as scholars might be able to deduce many observable implications of the MSF for every agenda-setting process they investigate (cf. Hall 2008; King, Keohane, and Verba 1994).

MSF lends itself to qualitative hypothesis testing via process tracing. The latter has been used extensively in social science applications as a diagnostic and analytical tool to identify *causal* mechanisms that link independent and dependent variables (Collier 2011). It contains careful description and analytical tests of causal chains of events. The tests use hypotheses and available

evidence to construct and compare alternative causal chains of events and make plausible inferences about their scope and validity (see Van Evera 1997). They vary in terms of standards and may affirm (straw in the wind and hoop), confirm (smoking gun and doubly decisive), and/or eliminate hypotheses. Rigorous MSF hypothesis tests may draw inspiration from Staff (2020) who used them to evaluate and discuss the explanatory power of potential causal mechanisms in private security policy decisions. Of course, not all four tests need to be performed. Rather the point is to adjudicate between different hypotheses to gain more analytical leverage between narratives and sources of evidence.

MSF studies, like all other case studies, must collect, analyze, and interpret data in ways that may be replicated by other researchers (Yin 2018). The quality of such replication depends on construct validity, internal validity, external validity, and reliability (e.g., King, Keohane, and Verba 1994). While considerable work has been done in dealing with validity threats in qualitative research, data reliability has been less explored but is equally important. We are concerned with demonstrating that the same results can be obtained by repeating the data collection procedure and analysis. In other words, other investigators should in principle be able to follow the same procedures and replicate the findings. Two strategies for ensuring reliability of case studies include creation of the case study protocol and development of a database (Yin 2018). Relevant documents to be included are an overview of the project, field procedures, sources and their appropriateness, guiding questions, and a report outline. MSF researchers also need to specify a timeline of events perhaps by consulting media sources to keep track of what happened and when. It helps guide data collection, especially interview questions directed at specific interviewees. A template is provided in Annex A.

When researchers use interviews, they need to decide whom to interview and ask the "right" questions (Tracy 2020). The sample of interviewees and the questions asked depend on the research question (for an example, see Tunstall et al. 2016). Generally, however, an ideal sample of interviewees should contain a mixture of relevant public and private actors. For example, interviewees should be members of parliament, agency heads and other civil servants, social actors from relevant interest groups, academics, journalists, and others depending on the issue at hand. The idea is to get relevant information by involved actors which then needs to be triangulated with other sources of evidence. Triangulation is important not only because it cuts down on the cost of interviews but also because it provides an important check on the veracity of evidence. People may lie to increase their sense of importance. If other sources, media interviews or archival research, suggest otherwise, it may be wise to collect more information. They may also point to potential reliability problems or serve as counterfactuals. For example, Zahariadis (2015) argues that Greek policy-makers felt trapped by a gigantic

tidal wave of public opinion which demanded no compromise in the name dispute between Greece and (now) North Macedonia. In all his interviews, policy-makers complained of the constraining effects of the national mood (in the form of public opinion). Had some policy-makers or other sources in the media reported that public opinion did not matter, that would be cause for more data collection. Triangulation identifies these discrepancies and increases data reliability.

The sample questions in the Online Annex are designed as a template in concurrent designs to help researchers collect data and tease out MSF implications. Of course, they are indicative, subject to adaptation depending on the project, and executed in at least two panel waves, meaning researchers need to conduct interviews at several temporal intervals to capture any dynamic changes. MSF studies that examine policy choices retrospectively (e.g., Ackrill and Kay 2011; Dolan 2021; Zohlnhöfer 2016) require a twofold adaptation of interview techniques. First, the sample interview questions need to address a single issue's agenda position or adoption (see Kagan 2019). Consequently, the questions need to be very specific to the issue at hand and have to inquire about how it was viewed relative to other issues at different points in time. Annex B contains a hypothetical, adaptable list of questions. Second, the questions need to identify temporal changes in the past, which can only be done by inquiring how one's thinking changed or reading about it in archival research. This means interviews will likely be supplemented by archival research, which has to include primary sources, e.g., memoirs, press releases, interviews, in addition to media reports and agency documents. Given the frequent semi-structured nature of interviews, coding is very difficult, but the questions may have to include some form of variance. It is best to specify these terms *a priori* in the protocol so that there is a common definition of what "more" or "less" means in what context.

Quantitative MSF Research

So far, the few MSF analyses applying quantitative methods predominantly either used a variant of regression analysis (DeLeo and Duarte 2021; Goyal 2021; Travis and Zahariadis 2002) or Qualitative Comparative Analysis (QCA) (Sager and Thomann 2016; Shephard et al. 2021). Therefore, we focus on these two methods in this section and start by discussing how well the methods' logics correspond to MSF (for a summary, see Box 2.3).

Regression analysis fits well the MSF's probabilistic logic. Testing the MSF core hypothesis would call for logistic regression analysis since it allows for assessing the size of an independent variable's effect (e.g., policy window) on the dependent binary variable (agenda status: change or stability). However, MSF attributes great importance to the right timing of events, which logistic regression analysis does not capture adequately. To model the time dimension more precisely, researchers may opt for event history analysis

(Box-Steffensmeier and Jones 1997). This regression for longitudinal data incorporates the concept of "right timing" by investigating the effect of the independent variables on the time until an event occurs (see Goyal 2021 for an application).

Notwithstanding its advantages, researchers need to consider two challenges when testing MSF with regression analysis. First, variance in the dependent variable is imperative. Hence, scholars must code agenda (or policy) change *and* stability. Second, testing the framework's entire core hypothesis requires assessing the combined effect of the five MSF key concepts (streams, policy window, policy entrepreneur) on agenda change. To model this effect statistically, researchers use interaction terms. Following Brambor, Clark, and Golder's (2006) suggestion to include all constitutive terms in a regression model, the resulting specification would consist of 31 independent variables. Furthermore, this specification leaves researchers with one fivefold and many fourfold, threefold, and twofold interaction terms, which is impossible to interpret. Engler and Herweg (2019) suggest overcoming this drawback by either condensing the streams' readiness for coupling into one variable or by testing partial couplings only. For example, DeLeo and Duarte (2021) use regression analysis to explore the dynamics of the problem stream and conduct a qualitative case study of the other streams and coupling activities.

In contrast to regression analysis, QCA is based on a deterministic logic and tests implicational hypotheses on the relationship between one or more conditions and an outcome. Consequently, QCA analyses can answer which (combination of) factors are necessary/sufficient for agenda change. Resulting from different assumptions about how the dependent and independent variable(s) are causally connected, it is quite evident that QCA does not capture MSF hypotheses accurately. To address this problem, we recommend paying particular attention to QCA's consistency value. Consistency informs about the degree to which the empirical evidence is in accordance with the hypothesis that a set of conditions (streams, policy window, policy entrepreneur) is necessary for agenda change (cf. Schneider & Wagemann 2012). Although measurement of how the degree of alignment between empirical evidence and theoretical expectations still differs from analyzing probable agenda change, examining the consistency value brings QCA closer to MSF's probabilistic logic.

Another problem is that timing matters in MSF reasoning while QCA tends to neglect temporality. However, recent QCA research introduces possible ways to include a time dimension in QCA applications (Fischer and Maggetti 2017; Verweij and Vis 2021).

Despite the challenges MSF-guided QCA applications face, they come with a significant advantage. They allow for examining individual factors' interplay and their combined effect on the outcome. This feature is a vital asset in terms of capturing the causal mechanisms the MSF expects to find.

Box 2.3 Promising quantitative methods for MSF analyses

	Regression analysis	*QCA*
Algebra	Linear algebra: covariational hypotheses on the relationship between one or more independent variables and a dependent variable.	Boolean algebra: implicational hypotheses on the relationship between one or more conditions and an outcome.
Mathematical logic	Probabilistic logic	Deterministic logic
Translation of the mathematical logic to MSF reasoning	Exemplified by the following MSF hypothesis: If a policy window opens, agenda change becomes more likely.	
	Logistic regression: If a policy window is open, agenda change becomes more likely.	If a policy window opens, the agenda changes.
		Hypothesis on necessary conditions: Only if a policy window opens, the agenda changes.
Advantage	Assessment of the size of an independent variable's effect	Examination of the interplay of individual factors and their combined effect on the outcome.
Challenge	Interpretation of interaction terms	Research questions and hypotheses covered by QCA differ from those derived by the MSF.

Source: Own compilation based on Engler and Herweg (2019).

Of course, it is possible to think of other applicable quantitative methods. Researchers who are interested in explaining the generation of alternatives, for example, could conduct network analyses to capture the impact of a policy community's structure on the survival of policy alternatives (Novotný, Satoh, and Nagel 2021). Another recent quantitative MSF example is Fowler's (2020) contribution who applies seemingly unrelated regression analysis to test the conditional nature of the three streams in affecting policy adaptation and implementation and their interdependent relationship. Hence, depending on the research question, a huge variety of possible quantitative methods could guide MSF analysis.

The Way Forward to Empirically Sound MSF Studies

Judging by the high number of empirical applications, MSF theorizing has been a tremendous academic success. Nonetheless, many – but not all – MSF (case) studies to date lack falsifiable hypotheses, a shared understanding of the main concepts, rigorous operationalization, and systematic empirical assessment of theoretical expectations.

The lack of explicit hypotheses and the adaptation to various contexts and different dependent variables are problems that have been dealt with in the recent literature, and researchers should consult the relevant papers when planning their study. Similarly, there are attempts to generate a common understanding of the main concepts of the framework, as hopefully has become clear in this chapter. The questions of operationalization and the systematic analysis of empirical evidence are slightly more cumbersome. With regard to operationalization, we have tried to discuss all relevant elements of the approach and provide ideas of how to measure them. As suggested in the introduction, this is an explorative endeavor to some degree, and there are likely other or maybe even better ways of operationalizing certain concepts. Hence, we do not consider our suggestions a definitive list. If you find better ways to measure some concepts, let us know!

Regarding the way we analyze data, quantitative or qualitative, there does not exist one single right method for MSF analyses. On the contrary, it depends on the research question and the available data. Nonetheless, qualitative and quantitative MSF studies differ in terms of the challenges they pose and the findings they might deliver. What they do not differ in, however, is the need to explicitly define and operationalize dependent and independent variables and their causal linkages (King, Keohane, and Verba 1994). Put differently, we recommend stating explicit hypotheses in addition to definition and operationalization of the MSF's key concepts in qualitative analyses, too.

In contrast, quantitative analyses must deal with the challenge of data availability. For instance, data on the readiness of the policy stream for coupling are not included in the databases widely used in political science research. Consequently, researchers may choose not to include these variables in the analysis, specify variables that can be measured but do not fully correspond with MSF concepts, or collect own data suited for validly measuring the variables. Whatever the choice, proxy variables and limitations of quantitative analyses are common remedies, and we would encourage researchers to test the MSF quantitatively. Quantitative analyses add weight to MSF's explanatory power that case studies cannot provide.

Translating the intriguing metaphors of the MSF into validly measured empirical concepts and systematically and rigorously testing its basic ideas are intricate tasks. But to exploit the MSF's explanatory potential in full, this is what MSF scholars need to continue doing. In this chapter, we provided some ways forward for empirically sound MSF studies, hoping to spark debates on how to best test the MSF. We believe this is a challenging but fertile endeavor for policy process research.

References

Ackrill, Robert, and Adrian Kay. 2011. "Multiple Streams in EU Policy-Making: The Case of the 2005 Sugar Reform." *Journal of European Public Policy* 18(1): 72–89.

Béland, Daniel, and Michael Howlett. 2016. "The Role and Impact of the Multiple-Streams Approach in Comparative Policy Analysis." *Journal of Comparative Policy Analysis* 18(3): 221–227.

Blankenau, Joe. 2001. "The Fate of National Health Insurance in Canada and the United States: A Multiple Streams Explanation." *Policy Studies Journal* 29: 38–55.

Boscarino, Jessica E. 2009. "Surfing for Problems: Advocacy Group Strategy in US Forestry Policy." *Policy Studies Journal* 37(3):415–434.

Box-Steffensmeier, Janet, and Bradford S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *American Journal of Political Science* 41(4): 1414–1461.

Brambor, Thomas, William R. Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1): 63–82.

Brunner, Steffen. 2008. "Understanding Policy Change: Multiple Streams and Emission Trading in Germany." *Global Environmental Change* 18(3): 501–507.

Cairney, Paul, and Michael D. Jones. 2016. "Kingdon's Multiple Streams Approach: What Is the Empirical Impact of this Universal Theory?" *Policy Studies Journal* 44(1): 37–58.

Cook, Jeffrey J., and Sara R. Rinfret. 2013. "The Environmental Protection Agency Regulates Greenhouse Gas Emissions: Is Anyone Paying Attention?" *Review of Policy Research* 30(3), 263–280.

Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science and Politics* 44(4): 823–830.

Craig, Rebekah L., Holly C. Felix, Jada F. Walker, and Martha M. Philipps 2010. "Public Health Professionals as Policy Entrepreneurs: Arkansas's Childhood Obesity Policy Experience." *American Journal of Public Health* 100(11): 2047–2052.

DeLeo, Rob A., and Alex Duarte 2021. "Does Data Drive Policymaking? A Multiple Streams Perspective on the Relationship between Indicators and Agenda Setting." *Policy Studies Journal* online first (doi: 10.1111/psj.12419).

Dolan, Dana A., 2021: "Multiple Partial Couplings in the Multiple Streams Framework: The Case of Extreme Weather and Climate Change Adaptation." *Policy Studies Journal* 49(1): 164–189.

Dudley, Geoff. 2013. "Why Do Ideas Succeed and Fail over Time? The Role of Narratives in Policy Windows and the Case of the London Congestion Charge." *Journal of European Public Policy* 20(8): 1139–1156.

Engler, Fabian and Nicole Herweg. 2019. "Of Barriers to Entry for Medium- and Large-N Multiple Streams Applications: Methodologic and Conceptual Considerations." *Policy Studies Journal* 47(4): 905–926.

Fischer, Manuel, and Martino Maggetti. 2017. "Qualitative Comparative Analysis and the Study of Policy Processes." *Journal of Comparative Policy Analysis* 19(4): 345–361.

Fowler, Luke. 2020. "Using the Multiple Streams Framework to Connect Policy Adoption to Implementation." *Policy Studies Journal* online first (doi: 10.1111/psj.12381).

Gerring, Jason. 2017. *Case Study Research*, 2nd ed. Cambridge: Cambridge UP.

Goyal, Nihit. 2021. "Policy Diffusion through Multiple Streams: The (Non-) Adoption of Energy Conservation Building Code in India." *Policy Studies Journal* online first (doi: 10.1111/psj.12415).

Hall, Peter A. 2008. "Systematic Process Analysis: When and How to Use It." *European Political Science* 7(3): 304–317.

Herweg, Nicole. 2016a. "Clarifying the Concept of Policy Communities in the Multiple Streams Framework." In Zohlnhöfer, Reimut and Friedbert Rüb (eds.), *Decision-Making under Ambiguity and Time Constraints: Assessing the Multiple Streams Framework*. Colchester: ECPR Press, 125–145.

Herweg, Nicole. 2016b. "Explaining European Agenda-Setting Using the Multiple Streams Framework: The Case of European Natural Gas Regulation." *Policy Sciences* 49(1): 13–33.

Herweg, Nicole. 2017. *European Union Policy-Making. The Regulatory Shift in Natural Gas Market Policy*. Basingstoke: Palgrave Macmillan.

Herweg, Nicole, Christian Huß, and Reimut Zohlnhöfer. 2015. "Straightening the Three Streams: Theorizing Extensions of the Multiple Streams Framework." *European Journal of Political Research* 54(3): 435–449.

Herweg, Nicole, Nikolaos Zahariadis, and Reimut Zohlnhöfer. 2018. "The Multiple Streams Framework: Foundations, Refinements and Empirical Applications" In Weible, Christopher M. and Paul A. Sabatier (eds.), *Theories of the Policy Process*. London: Routledge, 17–54.

Jones, Michael D., Holly L. Peterson, Jonathan J. Pierce, Nicole Herweg, Amiel Bernal, Holly Lamberta, and Nikolaos Zahariadis. 2016. "A River Runs through It: A Multiple Streams Meta-Review." *Policy Studies Journal* 44(1): 13–36.

Kagan, Jennifer A. 2019. "Multiple Streams in Hawaii: How the Aloha State Adopted a 100% Renewable Portfolio Standard." *Review of Policy Research* 36(2): 217–241.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry. Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton UP.

Kingdon, John W. 1984. *Agendas, Alternatives, and Public Policies*. Boston, MA: Little, Brown.

Knaggård, Åsa. 2015. "The Multiple Streams Framework and the Problem Broker." *European Journal of Political Research* 54(3): 450–465.

Leifeld, Philip. 2020. "Policy Debates and Discourse Network Analysis: A Research Agenda." *Politics and Governance* 8(2): 180–183.

Liu, Dawei, and Hang Xu. 2021. "A Rational Policy Decision or Political Deal? A Multiple Streams' Examination of the Russia-China Natural Gas Pipeline." *Energy Policy* 148 (doi: /10.1016/j.enpol.2020.111973).

Mu, Rui. 2018. "Coupling of Problems, Political Attention, Policies and Institutional Conditions: Explaining the Performance of Environmental Targets in the National Five-Year Plans in China" *Sustainability* 10 (doi: /10.3390/su10051477).

Münter, Michael. 2005. *Verfassungsreform im Einheitsstaat. Die Politik der Dezentralisierung in Großbritannien*. Wiesbaden: VS.

Novotný, Vilém, Keiichi Satoh, and Melanie Nagel. 2021. "Refining the Multiple Streams Framework's Integration Concept: Renewable Energy Policy and Ecological Modernization in Germany and Japan in Comparative Perspective." *Journal of Comparative Policy Analysis* 23(3): 291–309.

O'Donovan, Kristin. 2017. "An Assessment of Aggregate Focusing Events, Disaster Experience, and Policy Change." *Risk, Hazards & Crisis in Public Policy* 8(3): 201–219.

Petrocik, John R., 1996. "Issue Ownership in Presidential Elections, with a 1980 Case Study." *American Journal of Political Science* 40(3): 825–850.

Rawat, Pragati, and John Charles Morris. 2016. "Kingdon's 'Streams' Model at Thirty: Still Relevant in the 21st Century?" *Politics & Policy* 44(4): 608–638.

Roberts, Nancy C. and Paula J. King. 1991. "Policy Entrepreneurs: Their Activity Structure and Function in the Policy Process." *Journal of Public Administration Research and Theory* 1(2): 147–175.

Sager, Fritz, and Eva Thomann. 2016. "Multiple Streams in Member State Implementation: Politics, Problem Construction and Policy Paths in Swiss Asylum Policy." *Journal of Public Policy* 37(3): 287–314.

Sanjurjo, Diego. 2020a. *Gun Control Policies in Latin America*. Cham: Palgrave.

Sanjurjo, Diego. 2020b. "Taking the Multiple Streams Framework for a Walk in Latin America." *Policy Sciences* 53(1): 205–221.

Saurugger, Sabine and Fabien Terpan. 2016. "Do Crises Lead to Policy Change? The Multiple Streams Framework and the European Union's Economic Governance Instruments." *Policy Sciences* 49(1): 35–53.

Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge UP.

Seawright, Jason, and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61(2): 294–308.

Seeberg, Henrik Bech. 2017. "How Stable is Political Parties' Issue Ownership? A Cross-Time, Cross-National Analysis." *Political Studies* 65(2): 475–492.

Shephard, Daniel D., Anne Ellersiek, Johannes Meuer, Christian Rupietta, Ruth Mayne, and Paul Cairney. 2021. "Kingdon's Multiple Streams Approach in New Political Contexts: Consolidation, Configuration, and New Findings." *Governance* 34(2): 523–543.

Soroka, Stuart and Christopher Wlezien. 2019. "Tracking the Coverage of Public Policy in Mass Media." *Policy Studies Journal* 47(2): 471–491.

Staff, Helge. 2020. *The Political Economy of Private Security. How European States Privatize, Regulate and Produce Domestic Security*. Wien/Zürich: Lit.

Tiernan, Anne and Terry Burke. 2002. "A Load of Old Garbage: Applying Garbage-Can Theory to Contemporary Housing Policy" *Australian Journal of Public Administration* 61(3): 86–97.

Tracy, Sarah J., 2020. *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact,* 2nd ed. Hoboken, NJ: Wiley.

Travis, Rick and Nikolaos Zahariadis. 2002. "A Multiple Streams Model of U.S. Foreign Aid Policy." *Policy Studies Journal* 30(4): 495–514.

Tunstall, Ashley M., Christopher M. Weible, Elizabeth A. Tomsich and Angela R. Gover. 2016. "Understanding Policy Reform in Colorado's Domestic Violence Offender Treatment Standards." *Social Policy & Administration* 50(5): 580–598.

Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell UP.

Venters, Monoka, Meghan V. Hauptli and Lora Cohen-Vogel. 2012. "Federal Solutions to School Fiscal Crises: Lessons from Nixon's Failed National Sales Tax for Education." *Educational Policy* 26(1): 35–57.

Verweij, Stefan and Barbara Vis. 2021. "Three Strategies to Track Configurations over Time with Qualitative Comparative Analysis." *European Political Science Review* 13(1): 95–111.

Yin, Robert K. 2018. *Case Study Research: Design and Applications*, 6th ed. Thousand Oaks, CA: Sage.

Zahariadis, Nikolaos. 1995. *Markets, States and Public Policy: Privatization in Britain and France.* Ann Arbor, MI: University of Michigan Press.

Zahariadis, Nikolaos. 1996. "Selling British Rail: An Idea Whose Time Has Come?" *Comparative Political Studies* 29(4): 400–422.

Zahariadis, Nikolaos. 2003. *Ambiguity and Choice in Public Policy: Political Manipulation in Democratic Societies.* Washington, DC: Georgetown UP.

Zahariadis, Nikolaos. 2005. *Essence of Political Manipulation: Emotion, Institutions and Greek Foreign Policy.* New York, NY: Peter Lang.

Zahariadis, Nikolaos. 2008. "Ambiguity and Choice in European Public Policy." *Journal of European Public Policy* 15(4): 514–530.

Zahariadis, Nikolaos. 2015. "The Shield of Herakles: Multiple Streams and the Emotional Endowment Effect." *European Journal of Political Research* 54(3): 466–481.

Zahariadis, Nikolaos. 2021. "Coalitions and Values in the Flow of Policy Solutions." In Zittoun, Philippe, Frank Fischer, and Nikolaos Zahariadis (eds.), *The Politics of Formulating Policy Solutions: Arguments, Arenas, and Coalitions.* Bristol: Policy Press, 93–114.

Zohlnhöfer, Reimut. 2016. "Putting Together the Pieces of the Puzzle: Explaining German Labor Market Reforms with a Modified Multiple-Streams Approach." *Policy Studies Journal* 44(1): 83–107.

Zohlnhöfer, Reimut, Nicole Herweg, and Christian Huß. 2016. "Bringing Formal Political Institutions into the Multiple Streams Framework: An Analytical Proposal for Comparative Policy Analysis." *Journal of Comparative Policy Analysis* 18(3): 243–256.

Annex A: MSF Case Study Protocol Template

1 **Statement of Purpose**
   a   Define the main research question being addressed by this study
   b   Identify any additional research questions that will be addressed

2 **Research Design**
   a   Identify whether single-case or multiple-case designs will be used
   b   Describe the object of study (e.g., agenda change)
   c   Identify any propositions or sub-questions derived from each research question
   d   Specify the mechanisms that connect dependent and independent variables
   e   Elaborate on the case selection rationale

3 **Case Study Procedures and Roles**
   a   Procedures governing field research, whether interviews or document analysis
   b   Roles of case study research team members, if any

4 **Data Collection**
   a   Identify the data to be collected
   b   Define the data collection plan
   c   Discuss how the data will be stored

d   Specify what collection techniques will be used (e.g., interviews, archival research, document analysis, etc.)

e   For each data source (e.g., interviewee or document), state purpose for inclusion. Be specific about the kind/use of information expected

f   If applicable, include informed consent for data sources

g   For all data sources, summarize key information in written form

h   For interviews, add the following:
- State how the interview will be conducted; e.g., take notes? Audiotape? Online?
- Develop an interview guide/survey that lists the questions or issues to be explored
- Be sensitive to the fact that questions may differ slightly according to stakeholders being interviewed

5   **Analysis**
a   Identify the criteria for interpreting case study findings

b   Propose which (combination of) data are used to address which research question/sub-question/proposition

c   Cluster preferences by actor, time, and value (e.g., for/against a proposal) to visualize any emergent patterns

6   **Plan Validity and Data Reliability**
a   Construct validity – show that appropriate indicators are used for the concepts being studied. Tactics include triangulation which may entail using multiple sources of evidence and establishing chains of evidence

b   Reliability – would other researchers be able to replicate the findings?

c   Internal validity – show how one accounts for alternative explanations

d   External validity – identify the population domain to which findings may be generalized. Specify the outer limits of inference

7   **Reporting**
a   Identify target audience and relationship to the broader literature

8   **Schedule**
a   State time estimates for all of the major steps including deliverables: Planning, Data collection, Data analysis, and Reporting

Annex B: Sample Interview Questions for Designs over Time and Single Issue/Sector – Past

1   What role did you/your agency play in the policy reform (or analytical topic)?

 2 Were there problems with this policy that you felt needed to be addressed? Which ones, why, and in what way?
 3 When did you recognize it was a problem? Did you believe it received adequate attention from policy-makers up to that point? Why or why not?
 4 Was problem recognition sudden or incremental? Why?
 5 Was there a particular event or confluent factor that facilitated reform as an urgent issue? Why?
 6 Did you support or oppose reform? What kind of reform did you propose (if any) and why?
 7 Was your proposal/opposition supported and why?
 8 Did you change your preferences/position over time? Why?
 9 Were other actors/agencies involved in recognizing it as a problem?
10 Why do you believe the reform effort received serious attention/found receptive audiences among policy-makers? Other political actors? The general public? When and why?
11 Do you think interest groups played a role? The media? Specific groups, supranational institutions, or national governments?
12 Can you think of any specific factors, not actors, e.g., institutional, environmental, global issue, within or outside the policy community that may have contributed toward creating support or opposition to reform?
13 Can you list them in terms of importance to the reform effort?
14 Were there specific individuals/actors within the policy community who played a big role in pushing for the reform?
15 Can you tell me how these actors went about gaining support for or generating opposition to the reform ideas?
16 Is there anything else that you think I should ask/may be interested in?

3 The Code and Craft of Punctuated Equilibrium

Samuel Workman, Frank R. Baumgartner, and Bryan D. Jones

Introduction

At the center of punctuated equilibrium is the study of agenda-setting (Schattschneider 1960, Kingdon 1984). Which issues does government prioritize for action? How do individuals, organizations, groups, and policy systems make these decisions? Punctuated equilibrium reconciles a rift in the policy literature. Subsystems scholars described (mostly) incremental change, while agenda-setting scholars routinely noted large-scale policy reform. In particular, the theory attempts to explain large-scale agenda shifts even in the face of stable institutions and fixed preferences (Jones 1994). Punctuated equilibrium characterizes agenda-setting and policy change in long periods of incremental reform and short bursts of dramatic change (Baumgartner and Jones 2009). Both sides are correct. Indeed, blockage, institutionally induced stability, and incrementalism characterize policymaking. At the same time, dramatic change is possible. Moreover, the factors that make one possible make the other inevitable.

The approach emphasizes how individuals, organizations, and governing systems allocate attention to problems and prioritize issues that present interdependency and tradeoffs (Jones and Baumgartner 2005). The theory began by explaining how attention and positive feedback link seemingly stable policy subsystems to broader policy change through increased attention and positive feedback. Since that time, the theory has evolved to understand how policy systems and actors in them search for and process information about problems (Workman, Jones, and Jochim 2009, Baumgartner and Jones 2015, Workman, Shafran, and Bark 2017, Fagan and McGee 2020). Issue attention is the key to understanding long and broad policy change patterns and teaches us something about institutional adaptation to new problems and information (Workman, Jones, and Jochim 2009).

Though its roots lie in the US political system, the theory benefits greatly from expansion to other countries via comparative research, venues in the policy process, and extension to other social phenomena (Epp 2015, 2018). The conceptual core of all these extensions is issue attention. For a more robust discussion of the evolution and development of punctuated

equilibrium, refer to Jones and Baumgartner (2012) and Jones, Baumgartner, and Mortensen (2017) in *Theories of the Policy Process, 4th Edition*. For a discussion of comparative research, see Tosun and Workman (2017), who discuss the conceptual elements in detail and offer insight into the theory's substantive direction, or Baumgartner, Breunig, and Grossman (2019), who provide a variety of examples of comparative analysis using the approach. Jones, Epp, and Baumgartner (2019) argue that extreme punctuations are indicators of maladaptive policymaking systems due to a lack of attention to critical problems. They indicate how democracies are generally more adept at adapting to complex, changing environments, pointing to the approach's usefulness in addressing broader comparative research questions.

Chapter Roadmap

In what follows, we first address four persistent methodological themes that thread through the development and current state of research in the field. These themes are (1) using *text-as-data* in analyses of the discourse that underlies the connections between problems and solutions; (2) making sure policy topics are *comparable across time and space* so that one may assess reliable time series of changes in the policy discourse; (3) *collaboration with machines* in documenting the policy discourse; and (4) attending to *full distributions of change* rather than average changes across time. Keep these in mind as we delve into the in-depth discussion of the measurement system and the various methodological approaches characteristic in punctuated equilibrium analyses.

After discussing these themes, we dive into the measurement system, giving special attention to its underlying logic and current construction. We elaborate on the logic by considering the tradeoff between validity and reliability in constructing a flexible system to researchers' interests. We also offer advice on extending the system to other topics at finer levels of detail and the guidelines for doing so. We review the three paths to further development identified in the recent literature and conclude with a brief tour of the techniques researchers have used to classify text while maintaining the coding and measurement system.

We then discuss the various methodological approaches. The first category of techniques falls into the realm of descriptive inference. It focuses on assessing the shape of policy change distributions by comparing empirical distributions to others or theoretical ones. These comparisons illuminate aspects of the policy process across different policy process stages, institutions, and governing systems. The second category of approaches involves causal inference and regression techniques that differ significantly from the standard approach.

Finally, we offer our thoughts on the future of the approach, areas for improvement, and avenues for empirical and methodological development. We include our sense of where the process might illuminate elements of politics and policymaking where it has yet to be employed.

Four Persistent Themes

For our purposes, we want to identify the key elements of the empirical approach to agenda-setting. At the outset, four empirical themes undergird the development of the theory. Those working in the theoretical tradition pioneered many of these developments in public policy and political science more generally. Concern for these themes leads straightforwardly to much of the methodological innovation in the tradition.

Text-as-Data

The agenda-setting component of punctuated equilibrium requires measuring the discourse about policy topics across time. There are two key questions. First, how does the dialogue about a single topic change? For example, how much attention does a political system pay to environmental issues compared to defense issues? Second, how does the entire policymaking agenda change – the entire set of policy topics confronting policymakers, groups, institutions, and governments? The theory's empirical base is categorizing these substantive topics – the "what" of government. In general, this comes in the form of policy texts (e.g., congressional hearings, parliamentary questions, regulations, party platforms, or manifestos). The research design broadly demands categorizing texts by topics to map a substantive agenda. Before we can analyze an agenda, we must know what it is. These descriptive elements allow comparing changes in the agenda across time and space (more on that below).

Those working in the field have pioneered many techniques for classifying text, first by hand, then in collaboration with machines. In our experience, onlookers often associate the theory with the examination of budgetary change distributions. Many of the statistical innovations used to develop and test the theory indeed come from studying budgets. Still, the core of the measurement system and theoretical development depend on classifying text and analyzing the topics gleaned from them. These comprise most of the studies and efforts within the tradition.

Comparability Across Time and Space

Comparing the policy dialogue across time (so a policy topic means the same thing in 1950 as in 1990) and space (so a policy topic in 1990 means the same thing in the US as Denmark) is achieved by categorizing that dialogue into a topical classification system. The specific classification system used by policy scholars in the Comparative Agendas Project (CAP; www.comparativeagendas.net/) is one example. The CAP system classifies "chunks" of the policy dialogue (such as the title of a congressional hearing or a parliamentary question) into a policy topic system that consists of 22 major topics and 220 subtopics arranged hierarchically. Each subtopic can be traced backward in time, making a time series of similar congressional hearings addressing the subtopic.

There are many ways to classify text, and the choices present tradeoffs for the analyst. Most systems privilege the validity of classification, while others also elevate reliability. For purposes here, validity is whether an object (text in this case) is classified correctly. For example, is a congressional hearing about fish coded as "fish." In our experience, most systems for classifying text are designed to retrieve information, not compare topics across space and time (Jones 2015). Such systems err on the side of comprehensiveness, assigning multiple topics to each observation. The goal is to return *all* relevant information, even if that means providing much that is irrelevant.

The study of policy agendas requires measures of topics that are comparable across time and space. Comparison and trends are difficult under systems built for information retrieval. Thus, topical categorization must be backward compatible and consistent across groups, institutions, governing systems, and different stages of the policy process. In other words, the measurement system attempts to elevate reliability while keeping a keen eye on validity. Elevating reliability will seem a dire sacrifice for some, but the system's development allows for debate about validity and for reforming topic codes, so long as we achieve reliability. Given the aims of comparison, a lack of reliability renders arguments about validity logistically moot. We deepen the discussion of this logic when elaborating on the measurement system that has evolved from the research program.

Without adhering to the reliability principle – that reliability through time is fundamental to understanding trends – time-series analyses of agenda-setting processes are impossible. Suppose one is tracing attention to a policy topic across time but uses multiple categories to capture validity. In that case, one cannot create reliable time series tracing the policy topic back in time. That is the backward compatibility criterion. The idea is simple: Do not create new classification bins if an existing one will do. If a new classification is necessary, go back and recategorize any observations that might fit into it. In sum, make sure that the classifications used are consistent over time. Jones and Baumgartner developed the Policy Agendas Project to address this reliability problem in the mid-1990s. Policy classification systems existed then but were based on information retrieval rather than the goal of agenda-setting research – to trace changes in government priorities across time. Only the Budget Authorization category system developed by the US Office of Management and Budget holds fast to the reliability principle. That system is too limited to study the policy dialogue across time and institutions. A comparison of changes in Budget Authority topics with Policy Agendas categories addressed across time show similar patterns, but with a budgetary lag, as one would expect (Jones, Theriault, and Whyman 2019, pp. 67–68).

Collaboration with Machines

At the time of writing, the CAP catalogs the topics governments face across 25 countries, many more institutions, and a range of elements of the policy process (see www.comparativeagendas.net). The systematic collection and

coding of the data is a massive undertaking, requiring significant allocations of effort from project teams around the world. As the project has developed, it is impossible to hand-code every observation as when Baumgartner and Jones began so long ago. However, the investment in large data sets hand-coded reliably has paid dividends for those working on modern research directions.

The initial, hand-coded data sets provide a wealth of training data for implementing machine learning algorithms for classification (Workman 2015; Fagan and McGee 2020). In essence, the hand-coded data teach the machine to identify substantive content and assign it a topical code. New data sets form in one of two ways. If previously hand-coded data reasonably match the new content (e.g., same institution or part of the policy process), existing data are used to train the machine. Alternatively, suppose the prospective data are in some way new or different from existing data. In that case, the researcher may hand-code a proportion of the new documents, using them to train the machine to code the remaining documents (Loftis and Mortensen 2018; Sebok and Boda 2021).

Note that these are supervised machine learning techniques in all cases – there is a real collaboration between researcher and machine. This approach differs from blunt force classification strategies found in topic model implementation, usually using a Latent Dirichlet Model. These models produce classifications entirely derived from the data on hand. Owing to drift in language that machines rarely identify, these models produce consistent, reliable topic coding over time only as an artifact of the data. Topic modeling is not a bad thing; we like letting the data speak. However, these models are dependent on the complete set of data, not data plagued by drift or updated frequently. Researchers of government activity deal with an ever-expanding stream of data through time. Our concern with reliability across time and space means we cannot dispense with humans if our goal is to compare institutions and governments or construct the trends over time vital to policy research (Sabatier 1987).

Attention to Distributions of Change

Having a sound strategy for measuring text is one thing, but analyzing the data and making meaningful comparisons is another. The fourth empirical theme running through the development of the theory is attention to features of empirical distributions besides averages or other measures of the center mass. In particular, the theory seeks to explain the shape of agenda and policy changes. A Gaussian or normal distribution of policy changes indicates that policies are changing relatively mildly, with no large-scale jumps or lurches. However, policy punctuations are exactly those lurches. Because these lurches and punctuations generally characterize policymaking, policy changes do not typically form a normal distribution. The classic pattern or shape of change is a peaked center, accounting for many small, incremental changes clustered around the center, along with "fat tails" – more large changes than one might expect.

The governing system and the institutions that comprise it spend long periods under-responding to policy problems and then short periods of extreme change and dislocation as rusty gears lurch into motion and the system over-responds. While police mistreatment of Black citizens may have been occurring for years, the political system has ignored this, only suddenly to focus attention on the problem and leap into action. An empirical distribution of policy changes measures attention to policing and attention to all other policy topics addressed by the government – the full public agenda.

These dynamics demand measures beyond the typical preoccupation with the center of empirical distributions, necessitating consideration of the variance and the shape using statistics like kurtosis, skew, and various ways of accounting for unequal variance across the distribution. Of the three attributes of an empirical distribution – the center, dispersion, and shape – punctuated equilibrium brings dispersion and shape into sharp relief in explanations of policy change. This methodological focus is perhaps what most distinguishes the approach from traditional political science or policy scholarship.

To accurately describe policy dynamics, the theory shifts focus away from the center mass of empirical distributions. The theory takes this exercise in descriptive inference seriously. Accurate description allows for comparison across time and spatial units. In more recent work, it also allows for researchers to implement regressions whose focus is on these other features of distributions (more on this when we discuss methodological approaches).

We mention these persistent themes because they are invariably present across nearly all the work on punctuated equilibrium and several outgrowths. The reader would do well to keep these in mind as we delve more deeply into the measurement system and set of methodological approaches. However, we remind you that getting started assessing policy dynamics in this tradition begins with a concern for text and classification. The possibilities for analysis are endless and range from the highest reaches of government to the most local of local levels across any number of country-specific contexts. The set of CAPs are more than data but a *system* of extensible measurement.

The Measurement System

The research program's beating heart is a systematic effort to catalog attention to policy issues over long periods. The data sets fueling the research span issues, countries, levels of government, and stages of the policy process (Jones, Baumgartner, and Breunig, et al. 2009, Baumgartner, Breunig, and Grossman 2019). The process starts by identifying a set of texts (e.g., congressional hearings, parliamentary questions, or Environmental Protection Agency [EPA] reports) or other data amenable to categorization by topic (e.g., budget functions). Remember, the approach is rooted in the study of agenda-setting, so the researcher needs to think topically about the potential data. Take the text below, summarized from the Congressional Information Service (CIS) descriptor of a 2018 US Senate committee hearing:

Examining wildfire projections for 2018.

Discussion of wildfires in the US revolves around national forests and other federal lands, usually in the west. We code this hearing into the major topic of Public Lands and the subtopic of Public Lands. The key to the coding is that the hearing's topical content is as essential for understanding the policy as the value statements attached to the hearing's purpose. At the base of the categorization is the assumption that a well-trained layperson should be able to code the topical substance from a gestalt summary of the text or the text itself.

Conceptual Logic

The CAP topical coding scheme contains a set of major topics and subtopics nested underneath these. The classification system is premised on two rules: adherence to which makes the measurement system infinitely extensible and adaptable.

The first of these rules is that we code observations into mutually exclusive topic categories. An observation receives one and only one topic code. We know it hurts. Categorization can be a noisy endeavor. However, this quality of the classification system allows for comparing policy issues across time, countries, processes, and institutions. But what about validity? Shouldn't we assign "correct" codes even if we must give more than one?

Figure 3.1 illustrates the tradeoff that exists between validity and reliability when reliability is high and low. The horizontal axis displays the location of the "correct" or valid code τ. For the distribution of codes, *A*, the
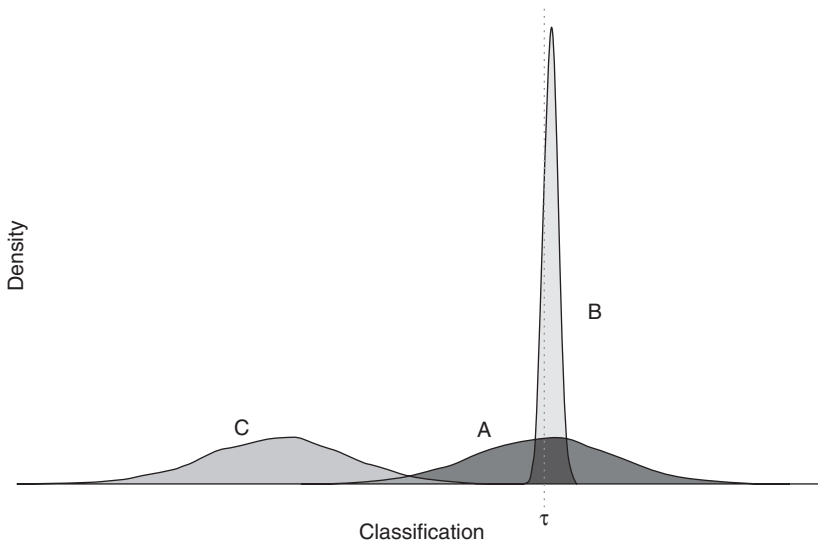

Figure 3.1 The tradeoff of validity and reliability

valid code is achieved on average. However, note that the reliability of the code is extremely low. While we are right, on average, we are wrong often and sometimes by a lot. Now, consider the distribution of codes, *B*. This distribution of codes is slightly biased on average by a small amount. However, distribution *B*'s reliability is very high – similar observations receive the same code, even if slightly off. While we would prefer classification that is correct (valid) and reliable, distribution *B* is far more valuable for comparison purposes. Finally, notice distribution *C* is neither valid nor reliable and thus has no analytical value.

When classification aims to construct trends over time or compare across spatial units, we would prefer the slightly imperfect but reliable distribution, *B*. This tradeoff is a unique feature of such publicly available, open-source classification systems like the CAP. We can have open debates about the validity of any mutually exclusive, reliable topical code. If adjustment is necessary, then reform easily comes from reclassifying the set of observations en masse. However, if reliability is lacking, no matter how valid on average, reclassification means coding each observation again (a logistical nightmare) if the goal is a comparison across time and space. Bias in statistical models is almost always bad because its magnitude is unknown. In open-source classification systems geared for comparison, however, reliability allows open debate and context-based reform of the classification scheme.

The second rule regards the classification system's structure and pertains to adapting the system to specific purposes. *The system is hierarchical, using major topics and subtopics within these.* Researchers can develop new, detailed coding schemes at a lower level so long as those are nested hierarchically – they do not cross the topics located above them in the hierarchy (May, Workman, and Jones 2008, Koski, Siddiki, et al. 2016, Workman 2021). New topic codes must be integrated into the topic system at each level of the hierarchy. Figure 3.2 illustrates how this allows for the extension of the coding scheme to more detailed policy issues.

Figure 3.2 illustrates how CAP subtopics could be aggregated to study topics that are more detailed than the existing scheme. A cursory search "fish" in the data on legislative hearings in the US turned up these five major topics, among others. Though this is not an exhaustive list for studying fish, it is illustrative of how one would do so. It also illustrates how policy issues
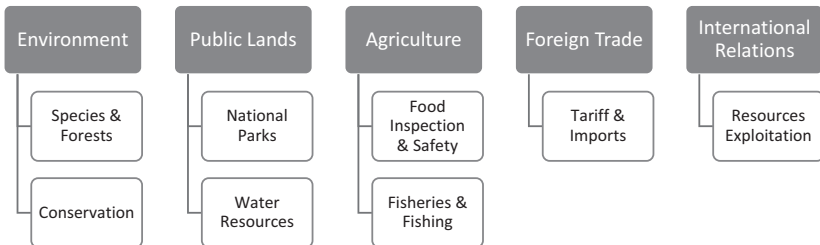

Figure 3.2 Topics related to the study of fish in the CAP

overlap and are interdependent. We offer this example because one of our most common queries is how to use the scheme to study topics of interest but in finer levels of aggregation. Not every one of these legislative hearings would be relevant, but the search is narrowed tremendously over a keyword search on existing databases geared for information retrieval.

How to Compile a Data Set

Now that we understand the logic of the coding system, what about logistics? How does one compile a data set? There are three approaches that, in our experience, yield quality, reliable data sets that allow for comparisons across time, spatial units, and different parts of the policy process. Bear in mind that each of these still depends on adherence to the two rules above. The best approach depends on the substantive context and the size of the task.

The first method is merely *hand-coding* the entire data set for topical content. Hand-coding may seem daunting, but there are instances where training the machine consumes similar effort to hand-coding. Hand-coding is especially important in two cases. The first pertains to extending the coding system to new contexts – countries, institutions, or government levels. While a researcher may well be able to use hand-coding to train a machine classifier later, some initial hand-coding will likely be necessary (instead of piggybacking off existing data sets). Hand-coding will also be required where the researcher aims to extend the topical coding scheme itself – into more specific and finely detailed issues (e.g., elementary and secondary education).

Supervised machine learning is the second way to compile a data set. This approach holds the promise of making enormous data collection efforts scalable but still requires initial hand-coding and human supervision. Analysts hand-code some representative sample (usually a random sample) of observations using this method. This hand-coded sample is then split evenly into representative subsamples. One subsample is the training data set for coding the other (hand-coded) subsample. The researcher can use the hand codes as a check for the machine's classification. The process iterates until the researcher is confident that the machine is ready to classify new data. The repeated process is dependent on representative samples and the choice of machine learning algorithms. With each interlude in the iterative process, human coders adjust the machine's classification in areas where it is unreliable. Early researchers used a mix of three algorithms and looked for the degree of agreement among them (Workman 2015, pp. 167–168). Later implementations use Bayesian priors, maximum entropy, and logistic regression (Loftis and Mortensen 2018). Recently, a team in Hungary has focused on repeated interactions between human coders and their machine helpers (Sebok and Boda 2021).

An alternate supervised machine learning protocol uses classified data from a different data set as the training set for entirely new data. In instances where the data constitute a "living" data set, updated periodically, this

represents a low marginal cost relative to the initial investment of effort. In instances where the data are not updated and represent a new frontier, exercise caution as text varies widely in different contexts and institutional settings. For example, bureaucratic regulations contain more detailed, technical, and scientific language than legislation. As a result, regulations are less prone to normal drift in the meanings attached to text in different contexts.

The third way of compiling a data set is to develop a crosswalk from an existing set of classifications to the CAP topic scheme. If a system codes data reliably (and with attention to topical content), these codes can be crosswalked to the CAP's policy content codes. For an example of how this works, see both the US and UK budgetary data sets. In the US, the crosswalk maps budget subfunctions to corresponding topical codes in the CAP scheme. For the UK, a similar process uses the *Blue Book on Government Expenditure*. The crosswalk between classification schemes has thus yielded crucial new data sets and ways to measure policy change. In fact, without the ability to crosswalk, these data sets may never have come together (as beyond rudimentary classification, one could not infer substance from money).

The construction and maintenance of the measurement system and associated data sets is a collective endeavor. Anyone seeking to employ the measurement system will find a wealth of help from the various project teams and researchers working in the area. Consult the CAP website to locate contact information for the various project teams.

With a reliably coded data set, there are two general concerns for analysis. The first is describing the distribution of policy change accurately. Much of the work in agenda-setting begins with the effort to get that stochastic process that generates the observations right – an exercise in descriptive inference.

Descriptive Inference

Research design aims to foster comparisons across time and spatial units like countries, institutions, organizations, or individuals. Comparison across time or space involves assessing some outcome relative to a baseline. Without a base of comparison, it is challenging to ascertain what we learn, theoretically and empirically, from a different outcome.

There are two ways to make comparisons and draw lessons from the data we observe. The first is to compare some theoretical outcomes. Game theorists expect a specific outcome in equilibrium – comparing empirical outcomes to game-theoretic equilibria teaches us something about social phenomena and strategy. The information processing theory that undergirds punctuated equilibrium relies on assessing a *distribution of outcomes.* The logic of comparison is no different here. We might compare student grades to a normal distribution; s-curves may indicate mimicking or contagion rather than learning. The second way to make comparisons is to assess one empirical distribution relative to another, serving as a baseline. How do the outcomes from presidential systems, for instance, compare with those of parliamentary systems?

Statistically, this difference is between *parametric* and *empirical* distributions. The classic bell curve, or normal distribution, is a theoretical, parametric distribution – meaning it has known properties and behaviors. Empirical distributions are built up from observational data and may not have known properties or behaviors. Sometimes empirical distributions map relatively well onto known parametric distributions; this allows us to build statistical models.

We should note that punctuated equilibrium does not imply any specific analytic approach. Many researchers eschew the distributional approach familiar in much of the research (Miller 2007, Baumgartner, Berry, et al. 2009, Workman 2015, Koski, Siddiki, et al. 2016, Bark and Bell 2018, Archuleta 2020, Russell 2020). Still, the distributional approach provides a useful gateway to the newcomer seeking to apply, adapt, and further develop the theory. The distributional approach syncs with the theory's early concern for explaining the simultaneous existence of large-scale and incremental policy change embodied in agenda-setting and subsystems literature, respectively.

Describing Empirical Distributions

The shape and contour of empirical distributions are central to the craft of comparison in punctuated equilibrium. Distributions have three general characteristics that allow us to compare them with others. Figure 3.3 illustrates the three features.
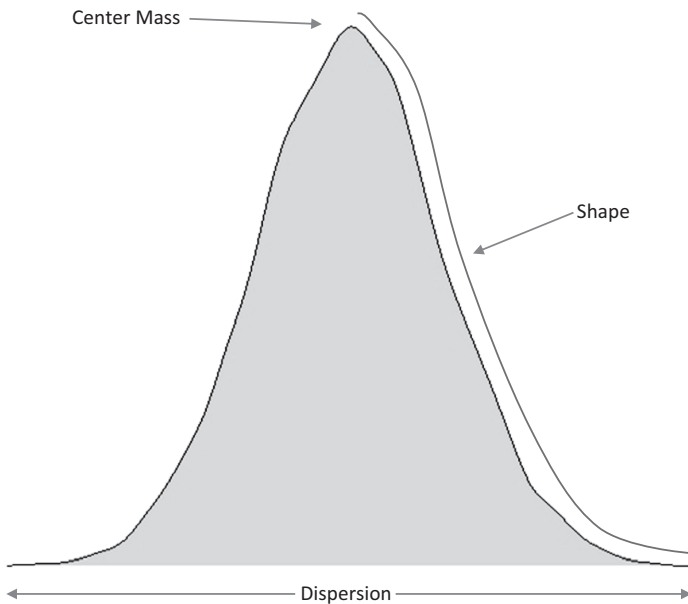

Figure 3.3 Three characteristics of distributions

The first and most often used characteristic of a distribution is its *center mass*. The center mass is familiar – it is the mean, median, or mode – the location of the most cases or the highest density of cases if using a density graph. The center mass is the basis for such simple comparisons as differences in means tests. The center mass is also the subject of most regression models – coefficients pertain to average effects.

The second feature of distributions is *dispersion*. The dispersion is less familiar but central to most statistical analyses – measures like the range, standard deviation, variance, interquartile range, and the minimum and maximum pertain to the distribution's dispersion. Dispersion often appears in descriptive tables and dictates inference in statistical models, whose standard errors relate the dispersion of estimates to the model coefficients.

The third feature of distributions, *shape*, is less familiar. The distribution's shape describes the distribution's contour; skew and kurtosis measures pertain to shape. Skew measures the symmetry, or asymmetry, in a distribution – does the distribution lean left or right? A negatively skewed distribution has a long tail to the left of its center mass. Alternatively, a positively skewed distribution has a long tail to the right of the distribution's center mass.

Kurtosis is another measure of a distribution's shape – specifically, it is a measure of the distribution's tails. In theory, kurtosis does not indicate anything about the distribution's center, though we note that "peaked" centers always attend empirical policy distributions. Kurtosis is a measure of the combined tails of a distribution relative to the rest. A distribution with "fat" tails is leptokurtic – having excess cases or density in the tails *compared to the normal distribution*. Platykurtic distributions have fewer values in the tails and more values farther from the center mass (the mean).

These features of a distribution offer avenues for comparison along the three characteristics, whether with theoretical or parametric distributions or other empirical distributions. Punctuated equilibrium focuses attention on large, rapid agenda and policy changes that result from positive feedback. If a system behaves in such a manner and we later observe its distribution, it will produce a highly leptokurtic distribution. During the periods of stability, cases will cluster near the mean, but occasionally there will be periods of dramatic adjustment, generating fat tails. Thus, we can assess the process that generates the distribution by looking at its shape. The key goal for analysis is accurately characterizing the distribution's shape – the stochastic process that generates the distribution of empirical outcomes. Then, the analysis proceeds to use the characterization of the distribution, often its shape, to compare the observed distribution to others or theoretical distributions.

Punctuated equilibrium implies that policy output measures follow a pattern of small changes occasionally interrupted by large-scale changes. Looking at one policy across time can highlight a policy punctuation at a particular moment in time, but it cannot indicate whether the entire system is subject to such punctuations. In a distribution of policy outputs, it is the change that interests us. So, we use the relative change as the fundamental unit for examination. Relative changes are constructed as follows:

$$\frac{\left(\text{Policy}_t - \text{Policy}_{t-1}\right)}{\text{Policy}_{t-1}} \tag{3.1}$$

Relative change is policy at time t (the present) minus policy at time $t-1$ (previous policy) divided by policy at time $t-1$ (previous policy). Multiplying the result by 100 would then give percentage change values. Then we prepare a frequency distribution based on the policy changes rather than the level of policy. Usually, this involves changes for each year for all policy topics.

By comparing kurtosis values in, say, stock market returns (a change distribution for all stocks in an index such as the Dow–Jones) to kurtosis values in changes in attention to policy topics in congressional or parliamentary lawmaking, we generally find that legislation has descriptively more cases or density in the tails than market returns. However, this also hints at something more. Generically, the process of lawmaking has more "friction" than the process of trading stocks. Legislation is a set of rusty gears difficult to turn, but they turn quite a bit when they do. By design, stock trading is more free-flowing (Jones, Sulkin, and Larsen 2003). Now that we know the focus of comparison is the shape, we can begin to lay a path to arrive at and reason through comparisons of policy distributions.

Recipe for Comparison

We view the steps below as a set of choices and opportunities rather than a set of rules. When engaging in work on punctuated equilibrium, the researcher need not always take the distributional approach. Much of the work we cite above uses other techniques. However, the distributional approach will be less familiar to those just getting started working in the field, perhaps because it is relatively unusual in political science and public policy outside of punctuated equilibrium approaches.

Taking the distributional approach implies that the researcher is interested in the prevalence of large-scale policy or agenda changes. This approach is a launching pad for thinking about how differences in policy and governing systems lead to peculiar distributions of outcomes. Different concerns will lead to other approaches.

How does one begin to make comparisons? Our belief here is that the researcher should always "look" at the data. The simplest way to compare distributions is the eyeball test – histograms and bar charts. Graph the distribution of agenda or policy outcomes across their range to get a sense of where the cases are and their prevalence at given values (or codes). Once the researcher understands the distribution, a comparison with a parametric (e.g., normal) or other empirical distribution is possible.

The researcher must scrutinize the data. Data often take the general shape of a well-known parametric distribution like the normal, and researchers proceed to model data accordingly. In our experience, a general "bell" shape

goads the researcher to complacency – failing to assess where the cases fall empirically. An excellent first step is to examine the distribution of outcomes as they relate to a parametric distribution (e.g., the normal or Poisson).

A second approach for longitudinal data is to examine trends. Policy and governing systems are sets of processes that generate a distribution of outcomes over time (Padgett 1980, Jones, Baumgartner, and True 1998). The dynamics of the system are knowable through those outcomes. We urge caution in the construction of trends. Reasoning about trends can devolve into event-centered analysis – understanding the entire distribution over time due to shocks caused by external events. An insight of punctuated equilibrium is that the external event or shock is a theoretical diversion – the significant bit is how a policy or governing system *processes* the event or adapts to the shock. In other words, events do not determine outcomes; how the system processes and adapts to events generates observable outcomes. After all, the distribution or trend we observe is the response filtered through these systems.

Beyond graphical analysis, the discussion of the three features of distributions gives us a wealth of options for comparison. We recommend calculating the kurtosis values of the distribution and comparing these. Any standard statistical program (including MS Excel) will calculate kurtosis. We recommend L-kurtosis in many cases since standard kurtosis scores can be unstable in the small sample sizes that characterize lots of policy research. See Breunig and Jones (2011) and Breunig and Koski (2006) for a complete explanation of L-kurtosis.

Since larger kurtosis values indicate more cases in the tails of a distribution, the values suggest what issues, institutions, and processes are more prone to large changes. The researcher can begin to reason through all the policy, political, and administrative factors that might explain varying kurtosis levels across distributions. Social scientists are used to thinking about variation *within* a distribution of policy outcomes. Historically, data in public policy have been expensive in terms of effort. In the modern, digitized world, data are cheaper, and collection and storage methods foster increasing returns to scale. Developing and testing theory across distributions is less familiar but offers tremendous insights into how governing systems prioritize and respond to problems.

An example of combining the intuition of kurtosis values and graphical analysis is the quantile–quantile, or q-q, plot. A q-q plot graphs the quantiles of an empirical distribution (e.g., count of hearings on pollution) against the quantiles of a theoretical, parametric distribution (e.g., normal or Poisson distribution). If the q-q plot aligns cases in a straight line at a 45-degree angle, then the theoretical distribution is a good approximation of the empirical distribution. The theoretical distribution is then useful for understanding the stochastic process generating the observed empirical distribution. Note that q-q plots facilitate a wide range of empirical and theoretical comparisons.

Examples of the Distributional Approach

The distributional approach originated in the field's study of budgets, a fertile ground for methodological innovation and development. (We suspect the theory is often mistakenly viewed as a budget theory.) Budgets illustrate the importance of correctly characterizing the stochastic process that generates observed change. These initial forays into the distributional approach also provided the basis for extending the theory conceptually and methodologically.

Incremental budgeting should be normally distributed according to budget theory. However, Jones, Baumgartner, and True (1998) characterize budgets as leptokurtic and statistically different from the normal distribution – budgets had excess probability (or cases) in the tails. Substantively, we can understand this as greater numbers of large changes and small changes than expected and fewer moderate changes. The initial research compared budget outcomes to the normal distribution to test theories of budget incrementalism. This research illustrates the value of comparing empirical distributions to theoretical distributions and illuminates an avenue of analysis and theory-building. We can begin to understand the ways empirical outcomes depart from theoretical and develop explanations to test by comparing output distributions from different institutions or processes (Jones, Baumgartner, and Breunig, et al. 2009).

Jones, Sulkin, and Larsen (2003) develop a theory of decision costs to explain variation in empirical distributions. Their analysis extends the distributional approach to outcomes other than budgets such as the private sector – assessing stock market returns. The question here was not so much whether these distributions diverged from the normal (they did), but in what ways they differed from one another. They develop a theory of institutional friction that is rooted in *decision costs* faced by an organization. The higher the cost of making a collective decision, the more delay in addressing the issue, and the more likely that a policy punctuation will occur as friction is suddenly overcome. Decision costs form the conceptual seed that leads to a full-blown theory of attention allocation rooted in the way institutions process information about problems (Jones and Baumgartner 2005).

Boushey's (2010) examination of state policy diffusion is an excellent example of comparing a theoretical expectation to an empirical distribution. If states engage in evaluating policies in terms of information – data and research – then the "shape" of policy diffusion should approximate the s-curve of the cumulative density of the normal distribution. Instead, Boushey finds that the shape of state policy adoptions approximates the cumulative density of the logistic distribution. Adoption is far more rapid than could be explained by an incremental learning model. Substantively, this means that the empirical distribution of state policy adoptions is more likely to result from copying or mimicking behavior rather than policy learning. Further, the degree of learning varies by policy topic.

A trio of studies extends the distributional approach to local and state governments, respectively. Robinson, Caver, et al. (2007) examine Texas state school districts and find that higher levels of bureaucratization lead to "smoother" budget changes. The study illustrates the influence of organization and professionalism in alleviating budget volatility. Meanwhile, Park and Sapotichne (2019) study city budgets and again find that bureaucratic professionalism mitigates against extreme policy punctuations. Breunig and Koski (2006) build an original data set of state budgets across topics. They show that state budget changes differ by topic and state, introducing the field to L-kurtosis and comparing those values across states.

The distributional approach is even useful for understanding behavior in the private sector. In a landmark study, Epp (2015, 2018) examines 1200 private firms' spending patterns and compares market versus deliberative resource allocation processes. Market processes aggregate the independent decisions of multiple actors, while deliberative processes are the product of interdependency brought on by the search for consensus. Epp finds that market processes lead to less punctuated resource allocation.

The distributional approach may yield leverage on extant hypotheses and be an end goal. The modern methods of analysis often use the distributional approach as the starting point for familiar analyses (e.g., regression). Fagan, Jones, and Wlezien (2017) use the distributional approach to compare punctuated budgetary outcomes in 24 countries. Using kurtosis measures as the dependent variable, they conduct a regression analysis that finds federal systems are more prone to punctuations than others. In many respects, this type of analysis is a bellwether for future research.

These studies are not an exhaustive list as there are too many well-done studies to cover. We present these to illustrate the development of the distributional approach and shed light on how they build toward the full Punctuated Equilibrium Theory (PET) methodological toolkit. These methodological strategies derive more from textual data than budgets, though the two are not mutually exclusive. They are also more likely to develop regression analyses that are more familiar to the reader.

Agenda Composition and Diversity

The discussion above uses a single frequency distribution comprised of policy changes in policy topics across time. Annual changes in how much congressional committees focus on environmental policy each year over the approximately 70-year period measured by the CAP yields only 69 data points, not enough to construct a stable frequency distribution. Nevertheless, annual changes in all major policy topics can be combined into one comprehensive frequency distribution, allowing the researcher to examine the kurtosis of several frequency distributions in different domains. How do the kurtosis topic changes addressed in congressional hearings across all topics compare to the kurtosis of laws passed? It turns out that the kurtosis for hearings is much closer to a normal distribution with modest tails than the

lawmaking distribution with its more extreme tails. Put simply, it is easier to get a topic scheduled for a hearing in Congress than to get Congress to pass and the president to sign a law. Friction or resistance is much greater for lawmaking.

Suppose we want to compare changes in the composition of items on the agenda as revealed in congressional hearings on a year-to-year (or Congress-to-Congress) basis? Is the legislative agenda dispersed across many topics or concentrated on a few? How does that change across time? Economists use the *Herfindahl Index* to assess market concentration – the square of the percentage of the market a firm holds summed across all firms. Political scientists use this measure to gauge agenda concentration, along with a second measure, *Shannon's entropy* (Shannon 1948, Baumgartner and Jones 2015, pp. 54–56). The calculation of Shannon's entropy *H* is as follows:

$$H(X) = -\sum_{i=1}^{n} P(x_i) \star \log_b P(x_i) \tag{3.2}$$

Entropy is the sum of the proportion of a variable, *P(X)*, multiplied by the log of that proportion across its categories. Many statistical programs have built-in functions for calculating entropy from a set of categories. To easily calculate entropy in a spreadsheet, organize the data such that periods appear in rows and the categories appear in columns. For each column, calculate the proportion of cases falling within that category and multiply this proportion by the same log. Summing across columns yields the entropy for the agenda in any given year.

Tracing the Herfindahl or entropy indexes across time shows that the composition is more open in some periods than others. In the US, this was a period from the late 1950s to the late 1970s. A second simpler "quick and dirty" method is to tabulate whether Congress addressed a topic in that year with at least one hearing (or held at least one vote on the topic, passed one law). This approach simplifies comparing agendas across domains and countries (Jones, Theriault, and Whyman 2019). For a thorough treatment of calculating entropy and its implications, see Boydstun, Bevan, and Thomas (2014).

Theory Development and Testing

PET addresses large-scale agenda and policy changes as they coexist in a sea of incremental reform. The theory directs attention to the dispersion and shape of empirical observations. Whereas statistical modeling exercises usually address the center mass or mean outcome of interest, PET focuses on the entire range of outcomes and their relative distribution across that range. This conceptual foundation has tremendous meaning for how statistical modeling proceeds.

What started as punctuated equilibrium within policy subsystems has evolved into a full-fledged theory of how policy and governing systems

generate, process, and search for information. A central concern of the theory is how information processing contributes to the shape of policy change and the adaptability of social systems (Workman, Jones, and Jochim 2009; Koski and Workman 2018).

The theory posits that governing and policy systems oscillate between over-responding and under-responding to information. Most of the time, policy systems under-respond to information – the response only partly addresses a problem. Under-response means that pressure for a government response mounts as the system accumulates error – the gap between response and the problem increases little by little. Once error accumulates and surpasses the response threshold, the governing system's rusty gears lurch into motion and produce a large response. Rather than matching the problem's demands, the governing system often over-responds to the problem once in motion.

In some cases, the over-reaction can feed on itself, resulting in a policy bubble (Jones, Thomas, and Wolfe 2014). So, policymaking is a process of continual approximation, always under- and over-shooting a problem such that the response is disproportionate or ill-fit to the problem. Systems may differ in their *efficiency*. More efficient systems will respond more proportionately. Less efficient systems, those with higher levels of *institutional friction*, will be less proportionate. These differences will be apparent in the distribution of outputs each system generates.

In the aftermath, a policy researcher observes a distribution of agenda or policy change with a tremendous amount of incremental change punctuated by more dramatic change. Institutional friction, feedback, and attention limits explain the characteristic pattern of policy dynamics. In addition to these, policies, particularly those that are locked-in, can contribute to punctuated policy dynamics – locking-in policy with no provision for revisiting or adjusting it all but guarantees that policy will become out of sync with the problem.

Feedback, Positive and Negative

The problem of policy lock-in is especially relevant to the relationship between elected and administrative institutions. The tenets of democratic governance would have administrative units make policy at the direction of elected officials who construct detailed legislation constraining bureaucratic departments. The irony of this is, of course, that locking-in policy causes errors to accumulate. Should the legislature have difficulty in prioritizing the problem for reform, then the error builds to the point of system failure.

Modern bureaucracies have broad discretion to make policies under existing legislation. Workman (2015) examines the relationship between legislative agenda-setting and bureaucratic policymaking in the US. Using a data set of a quarter-million regulatory proposals, he finds that the bureaucratic agenda is responsive to the legislative agenda and not necessarily lawmaking in real time. Legislative hearings on a policy topic are associated with regulatory proposals on the same topic. When the legislature prioritizes

issues, the bureaucracy responds with policy, providing an essential mode of policy adjustment and adaptability to policy problems.

The study uses *error correction models* to gauge the gap between attention in the legislature and bureaucratic policymaking. Error correction models assess how quickly two variables return to their long-run relationship after departures from their equilibria. Thus, the error correction model allows for evaluating the governing system's sensitivity to being out of kilter with the set of problems.

Mortensen and Green-Pedersen (2014) use error correction models to understand how agenda change connects to the creation and termination of government departments. Policy must adjust to changes in issues, but institutions make policy. Agency creation, termination, and reorganization facilitate the match between institutions and the set of problems on the agenda. In this case, the error correction pertains to adjusting the set of institutions available for given groups of issues (see also Baekgaard, Mortensen, and Seeberg 2018).

In the context of PET, we can assume that there is some informational change (e.g., shifting attention to an issue, a new topic, a redefined problem) and a governing system response in terms of policy or institutional reform. Whether a policy response or institutional reform, the error correction approach tells us how sensitive a policy area is to *under-responses* of the system. In essence, the error correction model measures the strength of *negative feedback*.

The original formulation of PET was concerned with policy change in the wake of the destabilization of institutions (e.g., policy subsystems). It is possible that destabilization instigates widely oscillating policy dynamics as the system over-responds and attempts to correct the over-response with an under-response.

Robinson, Flink, and King (2013) tackle the phenomena of destabilized, oscillating policy responses. Using budgeting in local school districts, they find support for the corrective hypothesis – large positive increases in budgets lead to more extensive cuts in succeeding periods. Likewise, large budget cuts spur large increases in the succeeding periods. With the creative categorization of large cuts and increases, they use a simple logit model to estimate the probabilities of corrective dynamics and policy oscillation. In another study, Breunig and Koski (2012) find that punctuated budget outcomes lead to less long-term growth, confirming policy feedback to have consequences for volatile policy change and substantive policy outcomes.

Institutional Friction

Friction in policymaking systems derives from four sets of costs: information costs, cognitive costs, decision costs, and transaction costs (Jones, Sulkin, and Larsen 2003, p. 154). Decision costs are key to punctuated equilibrium and are the most common and highest costs in the political sphere (cognitive costs are the realm of psychology, and transaction costs lie at the core of

economics). Studies based on the distributional approach find friction varies across countries, institutions, different policy process stages, and between the public and private sectors (Jones, Baumgartner, Breunig, et al. 2009, Breunig 2011, Epp 2015).

Several studies have begun to hone the mechanisms linking decision costs to punctuated policy and agenda change distributions. Breunig and Koski (2009) employ time-series specification and likelihood methods to examine the influence of states' institutional characteristics on incremental and punctuated budget changes. They find that institutionally strong governors induce punctuated budget changes over and above divided government or budget stringency.

Breunig and Koski (2018) find that governors can use their agenda-setting power to shift attention to specific substantive budget categories in more recent work. In doing so, governors amplify policy punctuations in the realm of increases and decreases. They use a quantile regression model that estimates how the effect of an independent variable varies across the quantiles of an empirical distribution. Substantively, we can expect that some independent variables foster incremental change (i.e., more influential in the center of the distribution), while others are more connected to policy punctuations (i.e., the tails of the distribution).

When we think of nonconstant variance (i.e., heterogeneity), we usually contemplate controlling for variance in the dependent variable. Quantile regression holds promise for understanding how influences on policy change vary across values of the dependent variable. Rather than a point estimate of the average effect (e.g., regression estimates), we take the variability of these theorized causes more seriously. The method holds tremendous promise for future research.

Since friction is a product of decision costs, it is hardly surprising that management, professional expertise, and organization influence punctuated policy dynamics. In a critical study, Robinson, Caver, et al. (2007) categorize budget changes by size and find that bureaucratization decreases the propensity for punctuated budgetary change. Flink (2015) uses a similar methodology to find that personnel turnover and low performance scores contribute to punctuated budgetary change. These studies are notable for bringing public administration into the realm of PET. Extending the analysis to American cities, Park and Sapotichne (2019) find that city managers and administrative officer discretion decreases punctuated budgetary outcomes. Previously, most studies tackled institutional characteristics. Researchers are beginning to examine how more organic forms of organization influence agenda and policy change with these studies.

Issue Complexity

A key insight of punctuated equilibrium, indeed, many policy process theories, is the complexity of policy issues. Issues can be complex in two ways. The advocacy coalition framework (chapter in volume) excels at understanding

how scientific and technical complexity comes into play in determining the course of policy change. Complexity in the context of PET pertains to the interdependence and tradeoffs inherent in policy problems. For instance, climate change is the confluence of agriculture, transportation, energy, and emergency management, among others. These issues are linked and present tradeoffs; improving on one issue may mean doing worse on others.

Methodologically, this presents a problem. At some limit, both organizational attention and resources are finite. The budget contains a fixed pool of resources for a given period. There are only so many committees in the legislature to attend to an issue. Increased attention on one issue means decreased attention to others. This interdependency can create a problem for statistical estimation under some circumstances.

Breunig and Busemeyer (2012) analyze budgets in the context of discretionary and mandatory spending priorities. The study treats budgets as compositional data, calculating the additive log ratios of the proportions before analysis. They use seemingly unrelated regressions to model the tradeoffs inherent in the budget categories of 21 countries in terms of issue-shares of the budget. Using the same approach on state budget data, Adolph, Breunig, and Koski (2018) find that partisans finance issues they prioritize by raiding the issues prioritized by their opponents.

Efforts to understand agenda complexity do not stop at modeling the interdependence in a causal framework. Measurement models can "map" the agenda space, illuminating the relationships among issues. Jones and Baumgartner (2004) use multidimensional scaling to assess the congruence between public priorities and the government agenda. Jochim and Jones (2012) scale 18 issues in the US Congress, finding that polarization is uneven across issues.

Workman (2015, pp. 116–123) shows that the bureaucratic policy agenda changes dramatically with a shift in partisan control of Congress. Using time-series factor analysis familiar in studying money in economics, Workman shows that shifts in partisan control of Congress alter the agenda map. He documents issue-bundling and issue-shuffling as partisan shifts link issues and break others apart. Workman, Carlson, et al. (2021) deploy a finite mixture factor model to map the agenda space for US education policy groups. The mixture model allows for an estimation of the issue space while controlling for the variance across groups. In the latter study, the measurement model establishes the validity of an extensible topic-coding scheme for education policy.

The Politics of Search

Issue complexity has consequences for how individuals and organizations search for information about problems and not just their responses. On a practical level, Koski and Workman (2018) argue that it affects governments' problem-solving capacity. Baumgartner and Jones (2015) conceptualize two forms of information search. Expert search is needed when the goal is to

understand one dimension of a problem, especially related to scientific, technical, or administrative expertise. Entropic search is necessary when there is uncertainty surrounding the relevant dimensions of a problem. For instance, climate change requires an entropic search for information as many of the dimensions are interdependent.

The research on the politics of search is the youngest of the various offshoots of PET. Workman, Shafran, and Bark (2017) broke the ice with a study of bureaucratic testimony at legislative hearings. They use likelihood methods for counts and simulation to assess the relationship between legislative uncertainty about a problem and the propensity to call bureaucratic experts to hearings. The study shows that topic uncertainty in the legislature leads to higher numbers of bureaucrats testifying. Among these, more of them are careerists rather than appointees, and the bureaucrats testify first, setting the course of the debate on the issue.

Fagan and McGee (2020) used reports at the Congressional Research Service coded by CAP codes to model the relationship between issue salience and the search for expertise by policymakers. Using time-series cross-sectional analysis, they find issue salience increases the request for reporting from the legislative bureaucracy. The context for exploring how policymakers and public organizations seek out information is ripe with opportunity from various methodological approaches.

Future Directions

In this chapter, we have given the reader a brief overview of PET, the classification and measurement system that underlies much of the analysis in the field, and outlined a range of methodological approaches that link to various conceptual and theoretical concerns in the area of research. We caution the reader that there are many more statistical approaches to testing PET. We chose studies that illustrate the development of the methods alongside the important conceptual and theoretical questions guiding research. Perusing this chapter and these studies will provide the reader with a strategy for analysis.

Here, we outline what we think are fruitful areas for extension of the work and some constraints – though many of these present opportunities for researchers hoping to get started. We can parse these opportunities and limitations in terms of statistical approaches and extensions of the data sets and the topic-coding scheme. Extensions of the data and classification schemes have low technical buy-in but require much effort. Pushing the boundaries of methodological approaches requires lower effort but demands more expertise.

Methodological Approaches

Budgetary data led the field's early methodological development. Some caution is needed when extending these approaches to other output

measures. Budgetary data are unique in bundling two measurement features in a single score – direction and magnitude of change.

Text-based data present a different problem. All else equal, increases in the number of laws, regulations, acts of parliament, or other policies, represent greater degrees of policy change and vice versa for decreases. But words are not numbers. A few words may inaugurate drastic policy change, and lengthy documents may do little to change the course of policy reform. So, textual data introduce more noise than budget measures in assessing the magnitude of change.

Directionality is also a challenge. Falling budgets mean "less," but it is not clear that a reduced number of laws means less policy, especially given our discussion of policy lock-in above. These are not substantive problems so much as measurement problems. Much of the current work relies on the normal distribution and relative policy changes (usually percent change). Alternative distributions will better characterize non-budgetary data. In addition, new conceptual development will be necessary to better link these distributions to changes in laws, hearings, parliamentary questions, speeches, and other non-budgetary data.

Laws, regulations, speeches, acts of parliament are all bound at zero – that is, there are no negative values, just less attention, down to none. We think a fruitful avenue for methodological development would come with greater reliance on likelihood methods on untransformed data, particularly methods for counts. The theoretical distributions for count data offer natural controls for phenomena like thresholds, dependency among outcome values, and dealing with long positive tails (e.g., overdispersion, zero-inflated, and hurdle models). These data also beg the question of the usefulness of skew as a secondary shape parameter.

Research on the policy process extends to ever more nuanced issues. Data collection is ongoing on many projects where the sample sizes are small by necessity. Given these, fitting the distribution that most accurately reflects the stochastic process generating the data is paramount. In these contexts, extreme value distributions and rare event frameworks may prove insightful.

In general, we think likelihood methods in which the researcher fits an appropriate stochastic distribution to the data and then parameterizes the dispersion or shape of the distribution offer tremendous opportunities. These parameters allow for theory testing and further understanding of causal mechanisms underlying punctuated policy dynamics more directly. See i-Marin, Hurka, et al. (2019) for an example of parameterizing a different characteristic of a distribution. Likewise, Workman, Robinson, and Bark (2021) parameterize a generalized Paretian distribution's shape parameter to model the organization's impact on bureaucratic regulations.

Our final recommendation is not statistical at all. Qualitative case studies dominated early research on PET. We believe that additional qualitative work might illuminate the causal mechanisms associated with decision-making under friction, the search for information in specific contexts, and the process of problem definition – how organizations construct meaning in

information, data, and research. These, in turn, provide fodder for large-N studies testing and refining these elements. For example, Shpaizman (2016) demonstrates using the CAP coding system to choose policies for intensive case studies. A team of Hungarian scholars led by Sebok Miklos and Boda Zsolt (2021) qualitatively isolate governing regimes in Hungary since 1867. They analyze the issue agendas of each regime quantitatively and qualitatively and provide qualitative analysis of salient issues in each regime. Archuleta (2020) examines the process of pension reform in the US military. The study illuminates how pension policy remained institutionalized on the old industrial model rather than updating in the wake of dramatic societal change. There are many more examples of such studies, and the potential of unifying qualitative and quantitative analyses from a policy agenda perspective is great. Generally, qualitative research and historical tracking of policy developments and key decision points have a large potential for theoretical advance, and we would like to see more of it.

Data and Classification

Data and classification offer some straightforward opportunities to contribute to the research area. We think of these contributions along three fronts – adding spatial units, extending the system to more detailed policy content, and new developments in collection or classification.

The data sets housed at the CAP contain data from 23 countries and 2 American states. Most of the countries are western-style democracies. Though they represent tremendous cultural and institutional variation, there is plenty of room to add and learn from countries and systems yet to be included, particularly those without democratic norms. Expansion to the nations and governments of the global south and east offers an opportunity to test many of the theory's central propositions in contexts outside western democracies and in less familiar cultural contexts. The decision-making model at the core of PET is one of bounded rationality. It should apply just as well to a monarchy or a dictatorship as it should to a system with multiple competing power sources. Indeed, Epp (2018) has applied the theory successfully to decisions made by private corporations. By adding to the range of tests of the theory, we can better assess it.

The project contains data from China and Russia, though authoritarian regimes are certainly underrepresented. Data collection in these systems can be difficult as measures can be unreliable, and researchers must undertake strategies to triangulate indicators to arrive at reliable data. Early research on authoritarian regimes suggests they exacerbate punctuated policy change (Baumgartner, Carammia, et al. 2017), though fiscal management seems to matter (Xiao, Wang, and Liu 2019).

There are only two states available to date. So, a focus on the American states is a fruitful line of research. So far, the main avenue for comparison is between states and the federal government. A more significant collection of states would allow state-level institutional differences to shine a light on the

agenda, much as they have budgets. Also, researchers could assess the effects of partisan regimes on the agenda since partisan control of state governments varies greatly. In general, spatial expansion across countries and down to lower government levels will advance theory development. For instance, Baekgaard, Larsen, and Mortensen (2019) have applied the system to city council agendas across Denmark. Automated text classification systems can allow the collection and classification of millions of documents, allowing further expansion of the theoretical approach to more systems.

Far fewer scholars make use of the policy subtopics than of the major topics. One can study more context-specific issues by selecting bundles of subtopics and aggregating them to understand policy dynamics at a more detailed level. The other approach is to extend the classification system to issues of much finer detail. By maintaining the hierarchical structure of the classification system and categorizing observations into new, mutually exclusive categories, the researcher can extend the topic classification system to an infinite level of detail. For example, see Workman, Carlson, et al. (2021) on education policy, May, Workman, and Jones (2008) on emergency management, and Koski, Siddiki, et al. (2016) on the agendas of food policy councils.

It is surprising to those of us who know the CAP well that it has not reached its potential in promoting, subsidizing, encouraging, and allowing higher and lower cost qualitative work on individual policy domains. By bringing together so many policy documents and classifying them by topic, we allow anyone to consolidate and identify all documents relating to hundreds of possible research questions with minimal effort. For example, every congressional hearing on the topic of immigration, or endangered species protection, or banking, or discrimination against women (to pick a few examples), is identifiable with a few mouse clicks. One may then combine these with laws, stories in the Congressional Quarterly, relevant data from other countries, or several other indicators. Such a compilation of information used to take months or years and could be the beginning of a historically rich qualitative policy history. The CAP is not only a repository of quantitative data requiring the most sophisticated statistical methods to generate new publications. It is a good starting point for qualitative research.

A final note. For policy researchers, the Comparative Policy Agendas data system has become the standard metric for categorizing policies. It is not perfect, nor is any other measuring system. However, as we have explained here, it can be expanded so long as hierarchy rules and attention to reliability and backward compatibility are honored. Occasionally, we still see independently derived policy categorization systems that cannot be integrated into the comparative system described here. In any case, we urge scholars to think about and plan how a new system might integrate with CAP during development. We welcome the opportunity to discuss category development and extension and are happy to offer help and consultation.

There is yet much to do in the development and testing of PET. We appreciate that the theory is alive and evolving both substantively and methodologically. The field continues to welcome ingenuity, innovation, and

problem–solving. Contributing to the research endeavor is straightforward, as is carving a niche in the field.

References

Adolph, Christopher, Christian Breunig, and Chris Koski. 2018. "The Political Economy of Budget Tradeoffs." *Journal of Public Policy* (Cambridge University Press (CUP)) 40: 25–50. doi:10.1017/S0143814X18000326.

Archuleta, Brandon J. 2020. *Twenty Years of Service: The Politics of Military Pension Policy and the Long Road to Reform*. Lawrence, KS: University Press of Kansas.

Baekgaard, Martin, Søren K. Larsen, and Peter B. Mortensen. 2019. "Negative Feedback, Political Attention, and Public Policy." *Public Administration* 97: 210–225.

Baekgaard, Martin, Peter B. Mortensen, and Henrik Bech Seeberg. 2018. "The Bureaucracy and the Policy Agenda." *Journal of Public Administration Research and Theory* (Oxford University Press (OUP)) 28: 239–253. doi:10.1093/jopart/mux045.

Bark, Tracey, and Elizabeth Bell. 2018. "Issue Prioritization by Bureaucratic Leaders: The Influence of Institutional Structure." *Administration & Society* (SAGE Publications) 51: 915–950. doi:10.1177/0095399718810194.

Baumgartner, Frank R., Jeffery M. Berry, Marie Hojnacki, David C. Kimball, and Beth L. Leech. 2009. *Lobbying and Policy Change: Who Wins, Who Loses, and Why*. Chicago, IL: University of Chicago Press.

Baumgartner, Frank R., Chrstian Breunig, and Emiliano Grossman. 2019. *Comparative Policy Agendas: Theory, Tools, Data*. Oxford University Press. www.ebook.de/de/product/35055846/comparative_policy_agendas_theory_tools_data.html.

Baumgartner, Frank R., Marcello Carammia, Derek A. Epp, Ben Noble, Beatriz Rey, and Tevfik Murat Yildirim. 2017. "Budgetary Change in Authoritarian and Democratic Regimes." *Journal of European Public Policy* (Informa UK Limited) 24: 792–808. doi:10.1080/13501763.2017.1296482.

Baumgartner, Frank R., and Bryan D. Jones. 2009. *Agendas and Instability in American Politics*. 2nd Edition. Chicago, IL: University of Chicago Press.

———. 2015. *The Politics of Information: Problem Definition and the Course of Public Policy in America*. Chicago, IL: University of Chicago Press.

Boushey, Graeme. 2010. *Policy Diffusion Dynamics in America*. Cambridge: Cambridge University Press.

Boystun, Amber E., Shaun Bevan, and Herschel F. Thomas, III. 2014. "The Importance of Attention Diversity and How to Measure It." *Policy Studies Journal* 42(2): 173–196.

Breunig, Christian. 2011. "Reduction, Stasis, and Expansion of Budgets in Advanced Democracies." *Comparative Political Studies* 44: 1060–1088.

Breunig, Christian, and Marius R. Busemeyer. 2012. "Fiscal Austerity and the Tradeoff between Public Investment and Social Spending." *Journal of European Public Policy* (Informa UK Limited) 19: 921–938. doi:10.1080/13501763.2011.614158.

Breunig, Christian, and Bryan D. Jones. 2011. "Stochastic Process Methods with an Application to Budgetary Data." *Political Analysis* (Cambridge University Press (CUP)) 19: 103–117. doi:10.1093/pan/mpq038.

Breunig, Christian, and Chris Koski. 2006. "Punctuated Equilibria and Budgets in the American States." *Policy Studies Journal* (Wiley) 34: 363–379. doi:https://doi.org/10.1111/j.1541-0072.2006.00177.x.

———. 2009. "Punctuated Budgets and Governors' Institutional Powers." *American Politics Research* (SAGE Publications) 37: 1116–1138. doi:10.1177/1532673X09333583.

———. 2012. "The Tortoise or the Hare? Incrementalism, Punctuations, and Their Consequences." *Policy Studies Journal*. 40(1): 45–68.

———. 2018. "Topping Off and Bottoming Out: Setting Budget Priorities through Executive Power." *Policy Studies Journal* (Wiley) 48: 342–366. doi:10.1111/psj.12247.

Epp, Derek A. 2015. "Punctuated Equilibria in the Private Sector and the Stability of Market Systems." *Policy Studies Journal* (Wiley) 43: 417–436. doi:10.1111/psj.12107.

———. 2018. *The Structure of Policy Change*. Chicago, IL: University of Chicago Press.

Fagan, Edward J., Bryan D. Jones, and Christopher Wlezien. 2017. "Representative Systems and Policy Punctuations." *Journal of European Public Policy* (Informa UK Limited) 24: 809–831. doi:10.1080/13501763.2017.1296483.

Fagan, Edward J., and Zachary A. McGee. 2020. "Problem Solving and the Demand for Expert Information in Congress." *Legislative Studies Quarterly* (Wiley). doi:10.1111/lsq.12323.

Flink, Carla M. 2015. "Rethinking Punctuated Equilibrium Theory: A Public Administration Approach to Budgetary Changes." *Policy Studies Journal* (Wiley) 45: 101–120. doi:10.1111/psj.12114.

Hanyu Xiao, Xiao Hu Wang, and Cheol Liu. 2019. "Budgetary Punctuations: A Fiscal Management Perspective." *Policy Studies Journal*. https://doi.org/10.1111/psj.12362

i-Marin, Xavier Fernandez, Steffen Hurka, Christoph Knill, and Yves Steinebach. 2019. "Systemic Dynamics of Policy Change: Overcoming Some Blind Spots of Punctuated Equilibrium Theory." *Policy Studies Journal* (Wiley). doi:10.1111/psj.12379.

Jochim, Ashley E., and Bryan D. Jones. 2012. "Issue Politics in a Polarized Congress." *Political Research Quarterly* (SAGE Publications) 66: 352–369. doi:10.1177/1065912912448930.

Jones, Bryan D. 1994. *Reconceiving Decision-Making in Democratic Politics: Attention, Choice, and Public Policy*. Chicago, IL: Chicago University Press.

———. 2015. "The Comparative Policy Agendas Projects as Measurement Systems: Response to Dowding, Hindmoor and Martin." *Journal of Public Policy* (Cambridge University Press (CUP)) 36: 31–46. doi:10.1017/S0143814X15000161.

Jones, Bryan D., and Frank R. Baumgartner. 2004. "Representation and Agenda Setting." *Policy Studies Journal* 32: 1–24.

———. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago, IL: Chicago University Press.

———. 2012. "From There to Here: Punctuated Equilibrium to the General Punctuation Thesis to a Theory of Government Information Processing." *Policy Studies Journal* (Wiley) 40: 1–20. doi:10.1111/j.1541-0072.2011.00431.x.Jones, Bryan D., Frank R. Baumgartner, Christian Breunig, Christopher Wlezien, Stuart Soroka, Martial Foucault, Abel Francois, et al. 2009. "A General Empirical Law of Public Budgets: A Comparative Analysis." *American Journal of Political Science* 53: 855–873.

Jones, Bryan D., Frank R. Baumgartner, and Peter B. Mortensen. 2017. "Punctuated Equilibrium Theory: Explaining Stability and Change in Public Policymaking."

In *Theories of the Policy Process*, edited by Paul A. Sabatier Christopher M. Weible. Milton Park: Taylor & Francis Inc. pp. 55–101, www.ebook.de/de/product/28096500/christopher_m_weible_paul_a_sabatier_theories_of_the_policy_process.html.

Jones, Bryan D., Frank R. Baumgartner, and James L. True. 1998. "Policy Punctuations: US Budget Authority, 1947–1995." *Journal of Politics* 60(1): 1–33.

Jones, Bryan D., Derek Epp, and Frank Baumgartner. 2019. "Democracy, Authoritarianism, and Policy Punctuations." *International Review of Public Policy* 1: 7–26. http://doi.org/10.4000/irpp.318.

Jones, Bryan D., Tracy Sulkin, and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions." *American Political Science Review* 97: 151–169.

Jones, Bryan D., Sean M. Theriault, and Michelle Whyman. 2019. *The Great Broadening* Chicago, IL: University of Chicago Press.

Jones, Bryan D., Trey Thomas, and Michelle Wolfe. 2014. "Policy Bubbles." *Policy Studies Journal* 42: 146–171.

Kingdon, John W. 1984. *Agendas, Alternatives, and Public Policies*. Boston, MA: Little, Brown. Koski, Chris, Saba Siddiki, Abdul-Akeem Sadiq, and Julia Carboni. 2016. "Representation in Collaborative Governance: A Case Study of a Food Policy Council." *The American Review of Public Administration* (SAGE Publications) 48: 359–373. doi:10.1177/0275074016678683.

Koski, Chris, and Samuel Workman. 2018. "Drawing Practical Lessons from Punctuated Equilibrium Theory." *Policy & Politics* 46: 293–308. doi:10.1332/030557318x15230061413778.

Loftis, Matt W., and Peter B. Mortensen. 2018. "Collaborating with the Machines: A Hybrid Method for Classifying Policy Documents." *Policy Studies Journal* (Wiley) 48: 184–206. doi:10.1111/psj.12245.

May, Peter J., Samuel Workman, and Bryan D. Jones. 2008. "Organizing Attention: Responses of the Bureaucracy to Agenda Disruption." *Journal of Public Administration Research and Theory* 18: 517–541. doi:10.1093/jopart/mun015.

Miller, Lisa L. 2007. "The Representational Biases of Federalism: Scope and Bias in the Political Process." *Perspectives on Politics* 5: 305–321.

Mortensen, P. B., and C. Green-Pedersen. 2014. "Institutional Effects of Changes in Political Attention: Explaining Organizational Changes in the Top Bureaucracy." *Journal of Public Administration Research and Theory* (Oxford University Press (OUP)) 25: 165–189. doi:10.1093/jopart/muu030.

Padgett, John F. 1980. "Bounded Rationality in Budgetary Research." *American Political Science Review* 74: 354–372.

Park, Angela Y. S., and Joshua Sapotichne. 2019. "Punctuated Equilibrium and Bureaucratic Autonomy in American City Governments." *Policy Studies Journal* (Wiley) 48: 896–925. doi:10.1111/psj.12333.

Robinson, Scott E., Floun'say Caver, Kenneth J. Meier, and Lawrence J. O'Toole Jr. 2007. "Explaining Policy Punctuations: Bureaucratization and Budget Change." *American Journal of Political Science* 51: 140–150.

Robinson, Scott E., C. M. Flink, and C. M. King. 2013. "Organizational History and Budgetary Punctuation." *Journal of Public Administration Research and Theory* (Oxford University Press (OUP)) 24: 459–471. doi:10.1093/jopart/mut035. Russell, Annelise. 2020. "Senate Representation on Twitter: National Policy Reputations for Constituent Communication." *Social Science Quarterly* (Wiley) 102: 301–323. doi:10.1111/ssqu.12904.

Sabatier, Paul A. 1987. "Knowledge, Policy-Oriented Learning, and Policy Change." *Knowledge* (SAGE Publications) 8: 649–692. doi:10.1177/016402598700 8004005.

Schattschneider, E. E. 1960. *The Semi-Sovereign People*. New York, NY: Holt, Rhinehart, and Winston.

Sebok Miklos and Boda Zsolt (eds). 2021. *Policy Agendas in Democracy, Autocracy and Hybrid Regimes: The Case of Hungary*. London: Palgrave.

Shannon, Claude E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27(4): 623–656.

Shpaizman, Ilana. 2016. "Identifying Relevant Cases of Conversion and Drift Using the Policy Agendas Project." *Policy Studies Journal* 45(3): 490–509.

Tosun, Jale, and Samuel Workman. 2017. "Struggle and Triumph in Fusing Policy Process and Comparative Research." In *Theories of the Policy Process*, edited by Christopher M. Weible and Paul A. Sabatier, 329–362. Boulder, CO: Westview Press. www.amazon.com/dp/0813350522.

Workman, Samuel. 2015. *The Dynamics of Bureaucracy in the US Government: How Congress and Federal Agencies Process Information and Solve Problems*. Cambridge: Cambridge University Press. www.amazon.com/author/samuelworkman.

Workman, Samuel, Bryan D. Jones, and Ashley E. Jochim. 2009. "Information Processing and Policy Dynamics." *Policy Studies Journal.* 37(1): 75-92. https://doi.org/10.1111/j.1541-0072.2008.00296.x

Workman, Samuel, Deven Carlson, Tracey Bark, and Elizabeth Bell. 2021. "Measuring Interest Group Agendas in Regulatory Proposals: A Method and the Case of US Education Policy." *Interest Groups & Advocacy*. https://doi.org/10.1057/s41309-021-00129-w

Workman, Samuel, JoBeth Shafran, and Tracey Bark. 2017. "Problem Definition and Information Provision by Federal Bureaucrats." *Cognitive Systems Research.* 43: 140-152. http://dx.doi.org/10.1016/j.cogsys.2016.11.002

Workman, Samuel, Scott E. Robinson, and Tracey Bark. 2021. "Bureaucratic Organization and Policy Change." *Institute for Policy Research and Public Affairs* Working Paper.

4 Methods for Applying Policy Feedback Theory

Mallory SoRelle and Jamila Michener

Introduction

In a political landscape structured by existing public policies, how can we understand the politics of policymaking without accounting for the dynamics created by preexisting programs? Policy feedback theory posits we cannot. Most theories of the policy process treat public policy as a final destination—the output of a progression shaped by other political factors. Scholars of policy feedback, by contrast, argue that the design, implementation, and benefits or burdens bestowed by public policy are capable of reshaping a variety of other political outcomes (Mettler and SoRelle 2018). The goal of policy feedback research is to ask how policies can influence subsequent politics,[1] and how that process ultimately affects future efforts at policy reform.

As Figure 4.1 depicts, policy feedback scholarship revolves primarily around four broad lines of inquiry: First, how do policies reshape *political agendas and the definition of policy problems* moving forward? Perhaps most obviously, once a policy is enacted, it influences agenda setting because it requires oversight and upkeep—either formally through reauthorization processes or informally through necessary updates, amendments, or technical corrections (Mettler 2016). Existing policies also induce a process of policy learning that may affect how lawmakers view and act to address an issue in the future. For example, scholars have demonstrated how policies addressing healthcare (Hacker 2002), childcare (Morgan 2006), higher education (Rose 2018), and consumer financial regulation (SoRelle 2020) set precedents that constrain what lawmakers deemed to be legitimate uses of government intervention in subsequent legislative efforts on those issues. Another critical way that existing policies can define future policymaking efforts is through the construction of target populations (Schneider and Ingram 1993) that generate norms—both positive and negative—about different beneficiary groups. These norms may influence how policymakers (or the public) evaluate the deservingness of different groups, with consequences for the generosity or punitive nature of future policy reforms.

A second major stream of inquiry considers how polices affect the *form of governance* for a particular issue or set of issues. Policies may build (or undermine) governmental capacities and standard operating procedures that affect
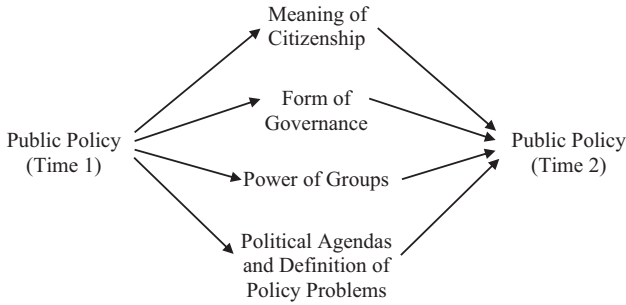
Figure content:

Meaning of
Citizenship

Form of
Governance

Public Policy
(Time 1)

Power of Groups

Political Agendas
and Definition of
Policy Problems

Public Policy
(Time 2)

Figure 4.1 Streams of policy feedback inquiry

lawmakers' choices about designs and administrative arrangements for future programs. For example, the creation of the Social Security Administration and its subsequent reputation as a capable and efficient manager of retirement benefits influenced the implementation of Medicare under its auspices (Derthick 1979). Skocpol and Finegold (1982) demonstrate how the distinct forms of governance employed for two New Deal programs—the Agricultural Adjustment Act and the National Industrial Recovery Act—shaped the disparate trajectories of those two policies. Scholars have also explored how the design of bureaucratic agencies can influence future political interventions through the process of path dependence (Pierson 1993). For example, Jacobs and King (2016) show how the initial design of the Federal Reserve helped the agency to expand and consolidate power in the regulation of financial markets, with significant consequences for future financial reform efforts.

While these first two streams of inquiry focus on policymaking institutions, the third and fourth streams address how programs can influence the political power of interest groups and ordinary individuals. With respect to the first query, feedback scholars have produced a wealth of research addressing how public policies influence the *power of groups*. Policies might, for example, create incentives for interest groups to form in the first place, or once established, shape their level of activity around or commitment of resources to a particular political agenda (Walker 1983; Patashnik 2008). For example, scholars demonstrate how the Social Security Act's provision of monetary benefits to older Americans sparked the formation of a powerful constituency and its associated public interest groups (Campbell 2002; Béland 2010). Thurston (2018), by contrast, illuminates how being denied benefits like housing finance that are conferred by new policy programs may motivate interest groups to coalesce and fight for access to those benefits.

Alternatively, the design of public policies can also preempt the formation of relevant coalitions. Goss (2013) explores how the liberal feminist foundations of equal rights laws enacted in the 1970s diminished the range of women's associational activity in politics, and SoRelle (2020) shows how

the marketizing provisions of consumer financial protections prevented the formation of broad-based borrower political mobilization in the wake of the 2008 financial crisis. Policies may also disempower communities by placing undue burdens on marginalized members, as can be the case for programs like TANF (Soss, Fording, and Schram 2011) and Medicaid (Michener 2018), or through carceral policies (Weaver and Lerman 2010; Davis 2020). Policies can foster partisan attachments with the creation of particular public programs, for example, the Affordable Care Act (ACA), that enable parties to mobilize voters who rely on them.

Finally, feedback scholars explore how public policies shape the *meaning of citizenship*. That is to say, once a policy is enacted, how does it structure the relationship between a government and those under its jurisdiction? Several scholars of policy feedback address this question quite literally: how do public policies shape membership, or citizenship, in a polity. For example, Bloemraad (2006) demonstrates how immigration policies in the United States and Canada differ in their effects on immigrant incorporation. There also exists a long tradition in feedback scholarship that charts how public policies expand or limit access to social citizenship and its benefits. For example, Skocpol (1992), Orloff (1993), Mettler (1998), Canaday (2009), Soss et al. (2011), Fox (2012), and Michener (2018) each present accounts of social welfare policy in the United States that demonstrate how those programs stratified access to benefits based on race and gender with consequences for beneficiaries' participation in the political sphere. Much of the recent work addressing the meaning of citizenship explores how public policies shape the way that people view themselves and others vis-à-vis the state. Specifically, policies can influence how the public construes the value and efficacy of their own citizenship in ways that can enhance (e.g., Campbell 2002; Mettler 2005), diminish (e.g., Soss 1999; Weaver and Lerman 2010), or produce mixed results (e.g., Chen 2013; Michener 2018; Hern 2019) for political engagement.

These four lines of inquiry implicate three main sets of political actors: governing institutions and agents, organized interests, and mass publics. They also center two broad mechanisms: resource and interpretive effects. The result is a rich tapestry of possible combinations for scholarly inquiry to probe how public policies (1) provide or limit access to resources and (2) construct positive or negative norms about self or government efficacy that can shape the capacities and incentives of institutions, groups, and individuals to make gains through the political process.

There is no single methodological template for scholars who wish to conduct policy feedback research. As a field, feedback scholarship and each of its major strands of inquiry bridges the divide between institutional and behavioral approaches to social science. Its origins lie in the realm of historical institutionalism, and its basic premise—that public policies are durable institutions capable of shaping a wide range of political outcomes—is consistent with this orientation. But much of the recent feedback work is inherently behavioral, exploring how beneficiary status as an individual (or group)

attribute shapes political attitudes, capacity, and engagement. The benefit of policy feedback's ecumenical methodological approach is that it offers scholars a great deal of flexibility in their empirical strategies. The challenge is that it invites methodological obstacles endemic to both institutional and behavioral frameworks. In the remainder of this chapter, we will consider how to conceptualize and design feedback studies, discuss the major empirical challenges facing policy feedback researchers, explore different methodological strategies employed by feedback scholars, and consider future directions for policy feedback inquiry.

Conceptualizing and Designing Policy Feedback Research

As scholars of policy feedback approach an expanding range of new and important inquiries, advancing knowledge in this arena will require continued and growing understanding of best practices with respect to research design and methods. A first-order task in this regard is to delineate the basic contours of theoretical conceptualization that form the bedrock of policy feedback scholarship. These conceptual elements shape research design, measurement, and analysis decisions in policy feedback studies. While they sometimes remain tacit, explicitly outlining them enables scholars to more critically and intentionally consider a fuller array of policy feedback questions and approaches. Table 4.1 emphasizes four key aspects of policy feedback: (1) the unit of analysis, (2) the contextual level, (3) the resources

Table 4.1 Key conceptual elements of policy feedback processes

Units of Analysis	Resource Effects	Interpretive Effects	Contextual Effects
Governing institutions/ actors/elites (e.g., elected legislators, bureaucrats, executives)	Infrastructure/ staff capacity; budget; expertise; professionalism	Social constructions of policy targets; political learning	National, state/ regional, or local governing institutions/ actors
Interest groups/ organizations (e.g., corporations, public interest groups)	Funding; administrative capacity	Coalitional assessments; future opportunities; difficulty/ possibility of change	National, state/ regional, and local organized interests
Mass publics (e.g., policy beneficiaries, communities)	Monetary resources; education; skills	Political efficacy (internal and external); political interest; political trust	Policy beneficiaries and other people affected by policy nested within communities, states, and nations

that shape the political process, and (4) the interpretive schemas and modes of learning that shape the political process.

Unit of Analysis

The unit of analysis in a given policy feedback study is the central actor being examined. As noted previously, early feedback research fell within the vein of historical institutionalism and was thus inclined toward an emphasis on governing institutions and actors as the unit of analysis (Derthick 1979; Skocpol and Finegold 1982; Pierson 1993; Mettler 1998; Thelen 1999; Hacker 2002). This analytical focus has continued in more recent work (Jacobs and King 2016) and remains a core component of policy feedback research. Importantly, the category of "governing institutions/actors" is capacious. Scholars working in this tradition have a wide range of actors and institutions to consider across governmental venues and contextual levels. The appropriate actor and institution for study is contingent on the research questions and theories at play. As a general practice, scholars focused on political institutions and elite political actors would do well to think critically about why particular actors and institutions are centered relative to others and to clearly articulate the theoretical and practical rationale for such choices.

Since the late 1990s, the central locus of policy feedback scholarship has echoed the discipline of political science more broadly, shifting toward individual-level behavioral approaches that emphasize how policies affect the political attitudes and actions of everyday people (mass publics). This line of research has spanned the gamut in terms of the specific populations targeted, with work focusing on how policies affect direct beneficiaries/ burdenficiaries (Soss 1999; Campbell 2002; Kumlin 2002; Mettler 2005; Weaver and Lerman 2010; Michener 2018), those in proximate contact with direct beneficiaries/burdenficiaries like family members (Walker 2014, 2020; White 2019; Walker, Roman, and Barreto 2020), those who live in communities with high concentrations of direct beneficiaries/burdenficiaries and are thereby indirectly affected (Burch 2013; Michener 2017), and the general population more broadly—whose attitudes may be shaped by salient policy changes (Soss and Schram 2007; Pacheco 2013).

Organized interests fall in the middle of the spectrum between governing institutions and ordinary people, and they have received relatively less attention from policy feedback scholars compared to the other two units of analysis. Still, an important and growing body of scholarship has emphasized organizations of various sorts, ranging from publicly subsidized after-school programs to public sector unions to women's groups to energy lobbies (Brown 2020; Skocpol 1992; Goss 2010; Anzia and Moe 2016; Hertel-Fernandez 2018; Goss, Barnes, and Rose 2019; Barnes 2020; Stokes 2020). The scope of this research reflects the many facets of policy feedback that involve organized interests, and the trajectory of the field suggests that an emphasis on feedback effects and organized interests will continue to grow.

Resource and Interpretive Effects

Moving across the columns in Table 4.1, the next two conceptual elements (resource and interpretive effects) are the core mechanisms understood to account for the processes by which policy feedback occurs. The underlying logic is that policy feedback happens when policy change triggers corresponding changes in resources like money, time, skills, and capacity (resource effects) and/or alters ideational schemas—the ways that people view and understand themselves, others, and the political world (interpretative effects). Considering resource and interpretive effects in light of varied units of analysis clarifies the relevance of each given the specific actors involved and pinpoints the different mechanisms at work across heterogeneous feedback processes. When governing institutions are the unit of analysis, for example, institutional features such as expertise, professionalism, staffing, and budgets are the most germane resources through which feedback processes might operate (Weaver 2010). Interpretive effects at this level may involve the way policies have altered political elites' social constructions of particular populations (Schneider and Ingram 1993; Dagan and Teles 2015) or the way policies have changed the political field, creating entirely new constituencies for elites to factor into the political calculi (Béland 2010).

Organized interests labor under different conditions than governing institutions, so we might expect distinct (even if overlapping) mechanisms with respect to resource and interpretive effects. Interest groups have less predictable funding streams, so funding is an especially important mechanism through which policies shape the politics of organizations (Francis 2019; Shanks and SoRelle 2021). Beyond funding, administrative capacity is another key resource. For example, when policies create new administrative challenges for organizations, they can drain groups of crucial resources that might have been directed toward other political ends. In the aftermath of the ACA, significant changes to Medicaid policies put the onus on state health advocacy organizations to fight for the adoption of newly available policy options (like Medicaid expansion), to oppose the enactment of other policies (Medicaid work reporting requirements), and to take on new work entirely (conducting outreach to enroll new Medicaid beneficiaries). These new burdens strained the resources and capacity of organizations and altered their political activities (Michener 2019b). At the same time, the interpretive repercussions of the ACA had varied organizational implications. On the one hand, state policy advocacy organizations were energized by the historic health coverage gains that the ACA promised. Subsequently, however, they were discouraged by the increasing polarization in the health policy political field and the ensuing difficulties posed for making progress on policy goals, forming viable coalitions, and having future windows of opportunity for potential political change (Michener 2019b).

Everyday people are also affected by overlapping but distinct feedback mechanisms. Monetary resources certainly matter, but civic skills and education are additional resources that have significant implications for feedback

processes (Bruch and Soss 2018). Crucially, most of the feedback scholarship focused on non-elites has emphasized the importance of interpretive effects, highlighting policies' influence on ideational factors like partisan polarization (Pacheco, Haselswerdt, and Michener 2020), political efficacy (Soss 1999), and political trust (Rocha, Knoll, and Wrinkle 2015; Bruch and Soss 2018; Davis 2020; Rosenthal 2021b).

Contextual Effects

The fourth column of Table 4.1 highlights the final core conceptual element undergirding policy feedback processes: context. Contextual effects differ analytically from both interpretive and resource effects insofar as they are not mechanisms explaining the occurrence of feedback effects but meso- and macro-level factors that explain the patterning of feedback effects across place (Michener 2018; Barnes 2020). Each of the units of analysis detailed above is nested within larger contextual units (countries, regions, states, cities, neighborhoods, etc.). These contextual realities shape the distribution and nature of policy feedback effects (Hern 2017; Michener 2018; Morel 2018; Grogan-Myers and Hatch 2019; Barnes 2020; Pacheco, Haselswerdt, and Michener 2020; Williamson 2020). For example, while much of the policy feedback literature considers how government service provision might influence political trust and engagement in advanced industrial democracies, Hern (2017) proposes and tests a contextually conscious scheme for feedback effects in what she calls low-capacity democracies. Attentiveness to context in policy feedback research does not substitute attentiveness to resource and interpretive mechanisms. Instead, an emphasis on context is complementary, generating richer multilevel analyses of feedback processes.

Insights on Conceptualizing Policy Feedback

The elements detailed in Table 4.1 point to critical insights for advancing the conceptual nuance of policy feedback scholarship. Not every study can contain all of the above-mentioned conceptual elements, but the overall schema laid out here can push scholars to think systematically about the logic for inclusion and exclusion of the various components. Considering these conceptual elements, scholars can ask critical questions about the design and content of their research: why focus on elites to the exclusion of organizations or ordinary people? Is it possible to account for the contexts in which policies are embedded? These and other questions can help scholars to expand their conceptual purviews in ways that can make policy feedback research more empirically sound and more practically useful.

Most policy feedback research focuses exclusively on one unit of analysis, contextual level, or type of mechanism. Going forward, more nuanced work might take a comprehensive approach, examining feedback processes across units of analysis, contexts, and categories of mechanisms. Especially rich studies can examine the interplay between these factors

(e.g., how intergovernmental political processes across contexts shape feed-back processes, how feedback outcomes at the "mass level" affect feedback processes among political elites, whether resource effects in turn shape inter-pretations of policies). Of course, not every study should be maximally com-prehensive along the conceptual registers detailed above. Throwing as much in a given study as possible simply for the sake of checking off conceptual boxes is not the goal. Instead, the rationale for emphasizing some concep-tual facets over others should be clear and well justified, even in articulating logical boundaries and limits.

The variety of conceptual pathways available to policy feedback scholars creates exciting opportunities to think creatively about data collection and analysis, but it also introduces some specific methodological challenges.

Challenges Conducting Policy Feedback Research

While policy feedback research traces its roots back to the pathbreaking work of twentieth-century scholars like E. E. Schattschneider (1935) and Theodore Lowi (1972), the field is still young. The bulk of its development has occurred since the mid-1990s, with the past two decades representing the breakthrough of policy feedback to more mainstream policy and polit-ical science scholarship. As with many research agendas—particularly those in their relative infancy—work in the field of policy feedback requires navi-gating a number of empirical challenges. The most significant obstacles to conducting policy feedback research thus far emanate from concerns over causal inference, data limitations, and problems of scope and perspective.

Capturing Causal Effects

One major goal scholars of policy feedback sometimes have is to demon-strate the cause and effect relationship between attributes, implementation, or use of a public policy and a particular political outcome. That is to say, we might be interested in demonstrating causal rather than correlational relationships between independent and dependent variables. The problem of causal inference is certainly not unique to policy feedback scholarship, but there are two issues in particular that make it especially challenging to cap-ture causal effects in a feedback framework: measuring policy exposure and addressing selection effects.

For scholars who want to understand how policy exposure structures future political preferences and behaviors, it is typically necessary to identify whether a person benefited from a particular program. This can be surpris-ingly difficult to do well. First, scholars relying on self-reported measures must contend with what Andrea Campbell (2012) described as the problem of recall in retrospective questioning. Unvalidated survey (or interview) responses are often unreliable when it comes to measuring whether a person participated in a particular program, especially when the requested time horizon is long. Response bias of this type can be mitigated by employing

specific techniques that are designed to enhance recall, like the use of short time horizons or event/autobiographical markers in questioning (see, e.g., Belli, Shay, and Stafford 2001), but recall problems can still manifest for research questions that do not lend themselves as well to such techniques.

Another issue in measuring exposure to policy programs stems from the fact that most people have multiple, ongoing interactions with government policies at any one time. But policy feedback work to date consists primarily of single-policy case studies (e.g., Campbell 2002; Mettler 2005). The result is that feedback scholars have not adequately grappled with the problem of multiple policy exposure in their research designs (though see Rosenthal 2021a).

Even when policy exposure can be appropriately measured, methodological challenges persist. One of the most pressing is that of endogeneity when measuring the causal relationship between a policy and a political outcome. Critics of policy feedback scholarship—particularly behavioral studies of mass publics—question whether the outcomes under examination can truly be tied to policy experience and not to other characteristics of a beneficiary or group. In statistical terms, the failure to address endogeneity, particularly that introduced by self-selection bias between recipients and non-recipients of government programs, is thus a perennial stumbling block (Mead 2004). So, how can scholars identify whether differences in preferences or participatory behavior are the result of policy usage versus some preexisting characteristics that influence which individuals elect to utilize a particular program? Increasingly, scholars are turning to statistical techniques like matching (Weaver and Lerman 2010; De Micheli 2018; Michener 2018), difference-in-difference (DID) (Lu 2014; Haselswerdt and Michener 2019), two-stage models (Rose 2018), and within program designs (Morgan and Campbell 2011). Another alternative to address the problem of selection bias in causal inference is to employ experimental or quasi-experimental approaches (e.g., Baicker and Finkelstein 2018; Clinton and Sances 2018; Lerman 2019; Kogan 2021), but the ability to randomly assign people to receive policy benefits is still relatively rare and often ethically dubious.

Data Limitations

Data limitations introduce a second broad category of methodological challenges facing feedback scholars. The ability to execute even the best research design is contingent on the availability of appropriate data. This issue can be especially problematic for scholars who wish to study historical cases or those with hard-to-reach populations. One of the biggest data limitations emerges for scholars who want to leverage existing surveys of program usage: the lack of political variables. There are a number of U.S.-based survey instruments that request data on program participation but do not include questions about political preferences or actions that are necessary for feedback scholarship (or the reverse). In fact, only two major data sets of program participation are well represented among feedback studies

(e.g., Weaver and Lerman 2010; Michener 2018; Bruch and Soss 2018): The National Longitudinal Study of Adolescent to Adult Health (Add Health), a nationally representative longitudinal study of American Youth, and the Fragile Families and Child Well-Being study, a panel study of disadvantaged parents and their children. Additional large surveys that are occasionally used by scholars of policy feedback include the National Educational Longitudinal Study (Condon 2015), the Black Youth Culture Survey (Cohen 2005; Barnes and Hope 2017), and the National Longitudinal Survey (Davis 2020). This puts scholars in the position frequently of needing to collect original survey data—a costly proposition that may limit who is able to conduct feedback scholarship.

In comparative contexts that embrace more universal social policy programs, scholars may have more latitude to employ national political survey instruments to capture feedback effects. For example, Guo and Ting (2015) rely on the 2005 China General Social Survey to capture employment sector differences in feedback effects generated by social insurance programs. But other data limitations emerge. In some cases, government data on policy provision may be unreliable (see Briggs 2012), challenging researchers to find alternative ways to capture the scope of policy implementation. And while the introduction of Global Barometer surveys offers new sources of both policy and political data that feedback scholars have begun to leverage, these data were not available in many locations until relatively recently, foreclosing prospects for more historical inquiry (see MacLean 2011).

A second data limitation emerges for those who wish to employ experimental approaches to capture feedback effects. The ability to randomly assign policy treatments, as described previously, can help scholars conduct sound causal research. But most government programs are not (and should not be) randomly assigned, limiting the number of policies that can be studied in this way. Some scholars have leveraged the differential rollout and implementation of social programs across time and place to approximate the empirical benefits of randomization (Lu 2014; Clinton and Sances 2018; Kogan 2021), but opportunities for such quasi-experimental approaches are often limited. Moreover, while some potential exists for scholars to partner with implementing agencies on the rollout of new programs, such partnerships are hard to develop—especially for scholars who lack the time and resources to invest in them. The random assignment of policy benefits also raises ethical questions that researchers must contend with that may constrain the ability to collect this type of data.

Scope and Perspective

The final limitation facing policy feedback scholars reflects both structural barriers and scholarly biases. How do we conduct research that explores the programs and politics that are most relevant to people's lived experiences? Policy feedback research to date generally focuses on the types of programs that are identified by elites—in both government and the academy—as

significant forms of government intervention. In doing so, feedback scholars overlook some of the most meaningful policies and policymaking institutions that shape the lived experiences of ordinary people—especially those who exist at the socioeconomic margins of a particular community. Michener, SoRelle, and Thurston (2020) propose a bottom-up framework for welfare state scholarship that could be applied to policy feedback to help remedy this lacuna. Similarly, the attributes of programs that scholars focus on frequently overlook the critical role that race, gender, and intersectional inequalities play in structuring feedback effects. Michener's racialized feedback framework (2019) is an attempt to help scholars navigate these existing biases for the issue of race, but similar efforts should be made to address other marginalities and their intersections.

Scholars of policy feedback are increasingly relying on more advanced and intentional methodological techniques and access to an array of rich data sets in order to overcome these challenges. The following section explores these in greater detail, highlighting common empirical tools and strategies used by feedback scholars to address each of the four major lines of inquiry.

Strategies for Empirical Analysis

Policy feedback scholarship has become a methodologically ecumenical space over the past two decades. While its roots lie in the largely qualitative field of historical institutionalism, feedback studies today employ a wide range of empirical strategies including historical analysis, interviews and ethnography, varied statistical techniques, experimental or quasi-experimental designs, and multi-method approaches. Notably, these empirical strategies often bridge qualitative and quantitative divides. The variety of both questions and approaches means there is no one-size-fits-all method to apply in policy feedback research. Instead, the following section offers a variety of examples that researchers might consult as templates from which to design their own studies. We describe why and how scholars employ these different methods and to which lines of policy feedback inquiry they apply each.

Historical Analysis

As previously discussed, the origins of policy feedback theory lie in the largely qualitative field of historical institutionalism. This field makes an apt starting point for policy feedback research because it contends that, as Orren and Skowronek explain, "all political change proceeds on a site, a prior political ground of practices, rules, leaders, and ideas, all of which are up and running" (2004: 20). For feedback scholars, existing policy programs are that site.

Historical analysis takes seriously the role of timing and sequencing, and it is particularly attuned to unraveling the processes of positive and negative path dependence that unfold between the enactment of a policy at time

one and the efforts for policy change at time two (Pierson 2004). Jessica Trounstine's *Segregation by Design: Local Politics and Inequality in American Cities* (2018) exemplifies this approach. Her research details how early local land-use policies (enacted at time one) shaped segregation, partisan polarization, and ultimately future patterns of public goods provision (enacted at time two) in the United States. Trounstine explains the causal challenges inherent in her research, noting how it can be difficult to get comprehensive data to (1) measure historic trends, (2) untangle issues of reverse causality, and (3) eliminate selection effects in feedback research. She explains how historical analysis can help to overcome these obstacles, weaving together archival records, original spatial data, and existing census and social survey data to "draw on the timing of events for evidence of causality" (45).

The use of historical analysis, which relies on a range of data including archival records, government documents, legislative histories, biographies, and surveys, has been employed by feedback scholars to address all four lines of inquiry. For example, Theda Skocpol, in her landmark study of U.S. social provision (1992), turns to qualitative historical methods to demonstrate how early efforts to enact civil war pensions ultimately shaped agenda setting and the forms of welfare provision that were politically viable in later decades. Skocpol relies on rich historical evidence culled from archives, legislative histories, magazines, etc. Esping-Andersen (1990) drew on a similar historical approach to explain how social welfare regimes structure politics in comparative perspective. In a more recent iteration of the historical approach, Leutert (2021) employs process tracing—drawing on three decades of qualitative and quantitative data—to identify the specific sequence of feedback effects that shaped the trajectory of market reforms in China.

Kristin Goss (2013) and Chloe Thurston (2018) each rely on historical analysis to demonstrate, as described earlier, how policies shaped the power of politically marginalized groups. In her study of women's collective action, Goss compiled an original data set of more than 10,000 women's groups' Congressional appearances between 1880 and 2000 for the centerpiece of her analysis. Thurston weaves together archival records from multiple public and private sources, including the National Archives, two presidential archives, the National Archives for Black Women's History, the National Association of Realtors Library, and personal papers from several special collections, to demonstrate how women's and civil rights groups mobilized for the expansion of credit. Scholars have also begun to use historical methods to understand how policies shaped the political engagement and preferences of mass publics. For example, Melanie Springer (2014) conducts an historical analysis of how voting laws shaped turnout patterns across the states over time, relying primarily on pooled time series cross-sectional models to analyze a data set of turnout compiled from several government sources. While each of these studies employs different data and empirical tools to address different lines of feedback inquiry, they share a common adherence to historical analysis.

Interviews and Ethnography

Conducting interviews or deep ethnographies is another methodological approach employed by feedback scholars. Interview data can be used to both generate and test hypotheses, and it is particularly helpful in tracing causal processes. Interviews can be conducted with elites—including elected policymakers, bureaucrats, service providers, and advocates—as well as with ordinary people—including beneficiaries or potential beneficiaries. They are an especially critical tool for scholars practicing a bottom-up approach to feedback research—that is to say, those who seek to center the voices of everyday people (Michener, SoRelle, and Thurston 2020). Interview and ethnographic methods have been most commonly employed by scholars to address feedback effects for mass publics, but they are also used to study feedback effects for the power of groups and the form of governance.

A number of scholars exemplify the use of interviews and ethnography to understand how marginalized communities' experiences with policy administration shape the political behavior and attitudes of community members. For example, Barnes and Henly (2018) utilize data from 85 qualitative interviews of childcare subsidy recipients in New York and Illinois to understand how their experiences with burdensome administrative features of the program shaped their attitudes about political efficacy and their future political engagement both in making benefit claims and in broader participatory practices. Similarly, Michener (2018) draws on semi-structured interviews with Medicaid beneficiaries and policy stakeholders conducted across multiple states to understand how program participation in different contexts shaped the political lives of recipients.

While each of these examples uses interview data to understand the effect of participating in a specific government program, Sally Nuamah employs a combination of ethnographic and interview methods to understand how policies that limit access to resources—in this case, the school closure process—prompt feedback effects on citizen participation. To untangle how negative experiences engaging with school closure proceedings diminished political efficacy among black community members, Nuamah (2021) conducted over 100 semi-structured interviews of individuals, community leaders, and policymakers and 60 ethnographic observations of community meetings in Chicago and Philadelphia between 2012 and 2017.

These scholars are primarily probing how policies shape the meaning of citizenship for mass publics, but others employ interview methods to understand how public policies can influence the form and capacity of governance and the power of groups. With respect to the first, Soss et al.'s in-depth study (2011) of the feedback dynamics of public assistance in the United States combines three years of ethnography in Florida—observing welfare sanction trainings, intake sessions, and regional staff meetings—with interviews of Florida state and regional officials and TANF case managers to understand how the disciplinary turn in public assistance policy affects welfare bureaucratic decision-making and administration. With respect to the

second, SoRelle's study (2020) of the politics of consumer financial protection relies on semi-structured interviews with the leadership of four major national consumer advocacy groups to help understand how the administrative environment for financial regulation shaped both their insider and outsider lobbying strategies in the lead up to and aftermath of the 2008 financial crisis.

Case selection and recruitment are major considerations for researchers who turn to interview and ethnographic methods in feedback scholarship, and investigators employ a number of different strategies depending on their goal. For example, Soss et al. focus on Florida because the state embodies all the criteria they identify as defining evidence of the neoliberal paternalist turn in poverty governance. As they explain of their case selection,

> The key question is not whether particular states are typical; it is whether their distinctive features make them more or less useful for studying how neoliberal paternalism proceeds in practice. Florida … [is] a leading-edge case in the disciplinary turn.
>
> (2011: 141)

By comparison, Michener's study of Medicaid employed a sequential process of case selection, using information from early interviews to shape the recruitment and substance of later interviews. She collected data until interviews reached "saturation" in the new information they revealed (see Small 2009). The outcome was 61 interviews conducted across 13 states. Michener also worked to ensure that interview sites varied with respect to the demographic characteristics of the area and the Medicaid policy designs employed (2018: Appendix A).

Quantitative Analysis

Of course, many policy feedback scholars rely on statistical methods of analysis. These are most prominent in studies of the behavioral elements of feedback—especially consideration of how policies shape the preferences and behaviors of individuals or mass publics. Because of the empirical challenges described earlier in the chapter, many researchers who are trying to establish evidence of individual-level causal relationships between policy experience and political outcomes eschew standard cross-sectional models in favor of a variety of alternatives.

An increasingly common, albeit resource intensive, option for feedback scholars is the use of panel or time series cross-sectional data. This can be especially helpful for those who are interested in capturing the effects of policy change. For example, Morgan and Campbell (2011) conducted a three-wave panel study to understand how enrollment in Medicare programs created by the 2003 Medicare Modernization Act changed both attitudes and political behavior among seniors. Similarly, Mettler and Jacobs launched a panel data collection effort to capture changing preferences for the ACA in the first

decade of its implementation (Jacobs and Mettler 2018; Jacobs, Mettler, and Zhu 2019). While each of these represents original data collection projects, other scholars have found existing panel data that provide sufficient information on the usage of a variety of policy benefits to allow scholars to capture the feedback effects of these programs. For example, Soss et al. (2011) use data from the Youth Development Study, a panel survey of Minnesota public school children, in their assessment of modern U.S. poverty governance.

The combination of longitudinal and geographic variation is an especially promising quantitative approach that leverages multiple sources of heterogeneity. For example, Pacheco, Haselswerdt, and Michener (2020) pool survey data from 2009 to 2016 and use a three-quarter moving average (e.g., pooling individual surveys three months at a time) to estimate quarterly opinion toward the ACA at the state level. They combine a time series approach with the use of multilevel regression and poststratification (MRP)—a small area estimation technique—to estimate state-level opinions toward the ACA (Park, Gelman, and Bafumi 2004, 2006).

Time series and panel data can be especially useful for studying policy feedback effects that may take time to come to fruition. Feedback scholars have no underlying, generalizable assumptions about the time frame in which feedback effects are likely to emerge. And, indeed, some studies have demonstrated remarkably fast windows between policy implementation and subsequent shifts in political behavior (e.g., Chen 2013; Stokes 2016; Clinton and Sances 2018). But there are also a number of policies, for example, Social Security and the ACA, that for a variety of reasons (e.g., low initial public approval, problems with early rollout) may take longer for feedbacks to emerge. In these cases, access to panel or time series data may be necessary to capture the evolution of positive or negative feedbacks (e.g., Jacobs, Mettler, and Zhu 2019).

Another strategy scholars use in quantitative studies of policy feedback is the introduction of a two-stage model that relies on instrumental variables to help mitigate the problems of endogeneity and selection bias. Deondra Rose exemplifies this strategy in her work on the effect of higher education policy on women's political participation. Rose employs two-stage regression analysis, wherein the first stage uses logistic regression to produce an instrument to predict higher education policy uptake and the second stage relies on the resulting propensity scores to predict political outcomes (2018: Appendix C). Mettler and Welch's (2004) analysis of the impact of the G.I. Bill on political participation offers another example where a two-stage model first produces an instrument to predict which veterans might chose to use G.I. benefits before using that instrument to predict veteran political participation. In each case, the inclusion of an instrumental variable is designed to account for unobserved factors that may influence whether an individual chooses to use a program that might also reasonably shape that person's political behavior.

Statistical matching accomplishes a similar goal, and other feedback scholars have turned to that approach to test the relationship between program experience and political outcomes. For example, Weaver and

Lerman's investigation of how contact with the criminal justice system influences people's attitudes toward government and political engagement employs matching to account for the possibility that contact with the criminal justice system is not randomly distributed, and thus individuals who have carceral contact may be inherently different from those who do not draw the attention of police powers (Weaver and Lerman 2010). Relying on panel data from both Add Health and Fragile Families, Weaver and Lerman begin by conducting statistical analysis of the data that employs controls for demographic traits, other forms of contact with the state, and the propensity for criminal behavior (based on self-reported offenses). They conduct a second set of analyses to corroborate these results, using genetic matching that, as with two-stage modeling, uses a propensity score to estimate the probability of receiving carceral contact. They match respondents from the Add Health data who report illegal drug use but no carceral contact with those who report illegal drug use and also carceral contact. In her analysis of the effect of Medicaid on beneficiaries' political participation, Michener employs coarsened exact matching to a similar end. She then uses seemingly unrelated regression to further account for self-selection bias (Michener 2018). In the comparative context, De Micheli (2018) uses coarsened exact matching to address the endogeneity of race and class in his study of the racialized feedback effects of Brazilian conditional cash transfer programs. Im and Meng (2016) employ propensity score matching to capture the effects of four welfare policies—pensions, educational subsidies, healthcare, and minimum livelihood assistance—on public preferences for government intervention in China.

Experiments and Quasi-Experiments

In addition to statistical techniques, several scholars are also beginning to employ more experimental and quasi-experimental methods to help untangle the relationship between policies and public perceptions. The use of random assignment allows for the direct comparison of average treatment effect between treatment and control groups, helping to account for endogeneity and establish causation (see Druckman 2011). These experimental approaches have been used primarily to explore individual-level effects of policy on the meaning of citizenship.

   As described previously, some scholars have leveraged timing intervals or geographic differences in the rollout of a particular policy to approximate the conditions of randomization. For example, in their study of how Medicaid shapes political participation, Clinton and Sances (2018) took advantage of the geographical discontinuity created by a 2012 Supreme Court case allowing states to decide whether to expand Medicaid under the auspices of the ACA. As a result of the case, the diverging choices to expand Medicaid made by neighboring states allowed Clinton and Sances to analyze and compare the rates of Medicaid expansion and subsequent increase of political engagement in border counties. Amy Lerman also made use of geographic

(and age) discontinuity designs in her study of citizen evaluations of public versus private service provision (2019). She exploits differences in public versus private waste management in demographically similar neighboring towns to demonstrate how reputation-motivated reasoning leads people to erroneously attribute high-quality public services to private providers—notably increasing support for privatization in the process. Jowei Chen (2013) exploits variation in the rollout of Federal Emergency Management Agency (FEMA) disaster relief to demonstrate how receiving distributive government benefits increases voter turnout among members of the incumbent party, while decreasing turnout for the challenger's party.

Scholars also use DID methods to capitalize on the policy rollout process to approximate experimental conditions. For example, Xiaobo Lu (2014) employs DID to capture the effects of a policy to abolish school fees on support for government financing of compulsory education in China. Lu uses data from two surveys—one conducted before and one after the implementation of the policy—to measure attitudes before and after "treatment." Moreover, Lu leverages the timed rollout of the policy between the two surveys to capture differences in "intensity" of the treatment duration.

Scholars have also turned to survey experiments to test the dynamics of policy feedback on individual attitudes and behaviors. While survey experiments may not be able to examine the direct relationship between the treatment of a particular program and the resulting political outcome, they can simulate differences in policy design or framing on reported preferences and practices. For example, Mettler (2011) uses an online survey experiment to test the effect of specific types of policy information on people's attitudes toward government social programs. Similarly, Faricy and Ellis (2014) employ a series of framing experiments to demonstrate how policy design shapes people's support for tax versus direct expenditure programs. SoRelle (2020) uses survey experiments to see whether making the state's role in financial regulation visible to individuals increases the likelihood of political engagement to support antipredatory lending reform.

Multiple Methods

While each of these individual empirical strategies can be used to study policy feedback effects, perhaps the most common approach in recent years has been the adoption of multi-method research designs. Drawing on a mix of empirical tools is often the best strategy for scholars who are attempting to address multiple lines of inquiry or modes of conceptualization, whose work spans both historical and current time periods, who cannot demonstrate causal patterns with only one type of data, or who face data limitations due to their study population. One of the best early examples of this multi-method approach is Suzanne Mettler's study *Soldiers to Citizens: The G.I. Bill and the Making of the Greatest Generation* (2005). Her book, which explores how G.I. benefits created positive resource and interpretive effects that

boosted the participation of recipients, begins with historical analysis of the implementation and expansion of G.I. benefits before incorporating original survey and interview data from a national sample of veterans.

More recent examples of this multi-method strategy include work from Michener (2018), Rose (2018), and SoRelle (2020). As has already been described, Michener's study of how state Medicaid programs generate differential political effects among beneficiaries puts both qualitative interview methods and quantitative statistical techniques to work to tell her story. Rose bridges historical analysis with the quantitative statistical analysis described previously. She leverages historical analysis to show how the adoption of gender parity language in federal financial aid policies led to increased educational attainment among women. Rose then turns to two-stage modeling of survey data to show how receiving federal aid influenced students' political engagement. SoRelle uses historical analysis to demonstrate how early consumer credit policies influenced policymakers' later choices about how to regulate consumer financing. She then turns to quantitative analysis of archival data, original survey data, and survey experiments to demonstrate how those policies reduced political efficacy and participation among consumer advocacy groups and individual borrowers.

These examples represent only a small number of the myriad studies of policy feedback effects, but they demonstrate the breadth of data sources and methodological approaches used to understand how policies shape politics. Policy feedback theory offers scholars the opportunity to ask many different types of questions and to answer those questions in many interesting ways, as befit a scholar's empirical skills and needs.

Future Directions for Policy Feedback Methods

Notwithstanding the extensive corpus of scholarship cited throughout this chapter, policy feedback research is still burgeoning, with fertile and untilled terrain that feedback scholars have yet to explore. The developmental trajectory of feedback studies is unpredictable to some extent, because it depends on how policy processes unfold in the real world. At the same time, an assessment of the field to date indicates important points of departure for further building and expanding the methodological (both empirical and conceptual) toolkit of feedback scholars. At least three future directions are especially worthy of discussion.

First, if we consider the units of analysis noted in Table 4.1 in relation to the general thrust of the feedback literature, it is clear that the lion's share of feedback scholarship focuses on individual political behavior and attitudes. As the field expands its purview to more thoroughly integrate organizations, this conceptual move should engender the collection of relevant new data. Local organizations (at the municipal, county, or neighborhood levels) have been particularly underattended in the feedback landscape. Systematic collection of data on local organizations and their implications for politics is a promising frontier for the field (Michener 2020).

A second important policy feedback frontier lies in taking seriously cumulative and overlapping experiences with policy (Mettler and Stonecash 2008; Rosenthal 2021a; Shanks-Booth and Mettler 2019). Doing so presents a data availability challenge, since very few surveys allow for a comprehensive accounting of government program usage while also containing political outcome variables. A turn to multiple policies also presents more fundamental empirical challenges. In survey-based quantitative analyses, the use of one government program is sometimes highly correlated with the use of another (for example, roughly 80 percent of beneficiaries of the Supplemental Nutrition Assistance Program for Women, Infants and Children are simultaneously beneficiaries of Medicaid). Given this patterning, identifying the "effect" of any program or set of programs is a tall order, inferentially speaking. Moreover, even where there are opportunities for randomization, they rarely (if ever) allow for experimental or quasi-random tests of the effects of *multiple* policies on political outcomes. In this arena, qualitative work may be particularly useful. Policy beneficiaries can often express in their own words precisely how one policy is different from others, even if the populations' those policies serve seem similar.

A third and final future consideration in research on policy feedback involves assessing policy feedback as an intentional political strategy. This topic sits in the cross hairs of several of the conceptual and empirical dilemmas articulated throughout this chapter. A handful of scholars have implicated policy feedback as an intentional strategic calculus (Galvin and Thurston 2017; Schneider and Ingram 2019; Hackett 2020). This is a crucial perspective that speaks to the pragmatic political uses of policy feedback research and theories (Hacker and Pierson 2019; Michener 2019b). Yet, the empirical challenge of gaining knowledge about intention is significant. Qualitative observation of and interviews with policymakers can point toward intentions, but they are not likely to fully reveal them. Beyond political elites, feedback scholars have much to learn about the ways that organizations intentionally attempt to shape policies with an eye toward future iterations of the political game. National, state, and local organizations may incorporate this kind of thinking in their strategizing without calling it "policy feedback." Scholars have yet to fully tackle the task of identifying whether and when this is happening and measuring its effects. It is exceedingly difficult to trace the empirical nodes in the process of (1) strategically planning for policy feedback, (2) taking steps to catalyze that process, and (3) finally actualizing (or failing to) the intended outcomes of that process.[2] Studying the entire feedback loop in this way (not just slices of it) requires multilevel, multi-institutional, multi-actor, longitudinal analyses that will often necessitate multiple methods. Yet, it is precisely such complex and multifaceted research that will most effectively enable us to grapple with policy feedback processes and their political implications. In this sense, one of the most significant methodological quandaries that policy feedback researchers will continue to face is the imperative to balance conducting research that faithfully captures the nuance of the policy world while also

systematically applying a range of rigorous and appropriate methods to the task of doing so.

Notes

1 Policy feedback differs from policy evaluation. While both fields explore certain effects of public policies, policy feedback studies focus explicitly on political outcomes. That is to say, scholars of policy feedback evaluate how public policies influence a range of political variables (e.g., public opinion, voter turnout) and not, for example, programmatic outcomes, efficacy, or efficiency.
2 On a related point, there is also a dearth of literature on the absence of feedback effects (although, see Patashnik and Zelizer 2013).

References

Anzia, Sarah F., and Terry M. Moe. 2016. "Do Politicians Use Policy to Make Politics? The Case of Public-Sector Labor Laws." *American Political Science Review* 110(4): 763–777.

Baicker, Katherine and Amy Finkelstein. 2018. "The Impact of Medicaid Expansion on Voter Participation: Evidence from the Oregon Health Insurance Experiment." *Quarterly Journal of Political Science* 14(4): 383–400.

Barnes, Carolyn Y. 2020. *State of Empowerment*. Ann Arbor, MI: University of Michigan Press.

Barnes, Carolyn Y., and Henly, J. R. 2018. "'They Are Underpaid and Understaffed': How Clients Interpret Encounters with Street-Level Bureaucrats." *Journal of Public Administration Research and Theory* 28(2): 165–181.

Barnes, Carolyn Y., and Elan C. Hope. 2017. "Means-Tested Public Assistance Programs and Adolescent Political Socialization." *Journal of Youth and Adolescence* 46(7): 1611–1621.

Béland, Daniel. 2010. "Reconsidering Policy Feedback: How Policies Affect Politics." *Administration & Society* 42(5): 568–590.

Belli, Robert F., William L. Shay, and Frank P. Stafford. 2001. "Event History Calendars and Question List Surveys: A Direct Comparison of Interviewing Methods." *Public Opinion Quarterly* 65(1): 45–74.

Bloemraad, Irene. 2006. *Becoming a Citizen: Incorporating Immigrants and Refugees in the United States and Canada*. Berkeley: University of California Press.Briggs, Ryan C. 2012. "Electrifying the Base? Aid and Incumbent Advantage in Ghana." *The Journal of Modern African Studies* 50(4): 603–624.

Brown, Heath. 2020. "How Do Policy Organizations Frame Issues and Shape Identity? Exploring the Case of School Choice." *Administration & Society* 52(7): 1038–1068.

Bruch, Sarah K., and Joe Soss. 2018. "Schooling as a Formative Political Experience: Authority Relations and the Education of Citizens." *Perspectives on Politics* 16(1): 36–57.

Burch, Traci. 2013. *Trading Democracy for Justice: Criminal Convictions and the Decline of Neighborhood Political Participation*. Chicago, IL: University of Chicago Press.

Campbell, Andrea Louise. 2002. "Self-Interest, Social Security, and the Distinctive Participation Patterns of Senior Citizens." *American Political Science Review* 96(3): 565–574.

———. 2012. "Policy Makes Mass Politics." *Annual Review of Political Science* 15: 333–351.

Canaday, Margot. 2009. *The Straight State: Sexuality and Citizenship in Twentieth-Century America*. Princeton, NJ: Princeton University Press.

Chen, Jowei. 2013. "Voter Partisanship and the Effect of Distributive Spending on Political Participation." *American Journal of Political Science* 57(1): 200–217.

Clinton, Joshua D., and Michael W. Sances. 2018. "The Politics of Policy: The Initial Mass Political Effects of Medicaid Expansion in the States." *American Political Science Review* 112(1): 167–85.

Cohen, Cathy J. 2005. *Black Youth Culture Survey*. Chicago, IL: Black Youth Project.

Condon, Meghan. 2015. "Voice Lessons: Rethinking the Relationship between Education and Political Participation." *Political Behavior* 37(4): 819–843.

Dagan, David, and Steven M. Teles. 2015. "The Social Construction of Policy Feedback: Incarceration, Conservatism, and Ideological Change." *Studies in American Political Development* 29(2): 127.

Davis, Brandon R. 2020. "Testing Mechanisms: Carceral Contact and Political Participation." *Social Science Quarterly* 101(2): 909–924.

De Micheli, David. 2018. "The Racialized Effects of Social Programs in Brazil." *Latin American Politics and Society* 60(1): 52–75.

Derthick, Martha. 1979. *Policymaking for Social Security*. Washington, DC: Brookings Institution.

Druckman, James N. 2011. *Cambridge Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.

Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.

Faricy, Christopher, and Ellis, Christopher, 2014. "Public Attitudes toward Social Spending in the United States: The Differences between Direct Spending and Tax Expenditures." *Political Behavior* 36(1): 53–76.

Fox, Cybelle. 2012. *Three Worlds of Relief: Race, Immigration, and the American Welfare State from the Progressive Era to the New Deal*. Princeton, NJ: Princeton University Press.

Francis, Megan Ming. 2019. "The Price of Civil Rights: Black lives, White Funding, and Movement Capture." *Law & Society Review* 53(1): 275–309.

Galvin, Daniel J., and Chloe N. Thurston. 2017. "The Democrats' Misplaced Faith in Policy Feedback." *The Forum* 15(2): 333–343.

Goss, Kristin A. 2010. "Civil Society and Civic Engagement: Towards a Multi-level Theory of Policy Feedbacks." *Journal of Civil Society* 6(2): 119–143.

———. 2013. *The Paradox of Gender Equality: How American Women's Groups Gained and Lost Their Public Voice*. Ann Arbor, MI: University of Michigan Press.

Goss, Kristin A., Carolyn Barnes, and Deondra Rose. 2019. "Bringing Organizations Back In: Multilevel Feedback Effects on Individual Civic Inclusion." *Policy Studies Journal* 47(2): 451–470.

Grogan-Myers, Patrick, and Megan E. Hatch. 2019. "City Services by Design: Policy Feedback, Social Construction, and Inequality." *State and Local Government Review* 51(1): 68–77.

Guo, Hua, and Kwok-fai Ting. 2015. "Policy Feedback in Transitional China: The Sectoral Divide and Electoral Participation." *Social Science Research* 54: 233–245.

Hacker, Jacob. 2002. *The Divided Welfare State: The Battle over Public and Private Social Benefits in the United States*. New York, NY: Cambridge University Press.

Hacker, Jacob, and Paul Pierson. 2019. "Policy Feedback in an Age of Polarization." *The ANNALS of the American Academy of Political and Social Science*, *685*(1): 8–28.

Hackett, Ursula. 2020. *America's Voucher Politics: How Elites Learned to Hide the State*. Cambridge: Cambridge University Press.

Haselswerdt, Jake, and Jamila Michener. 2019. "Disenrolled: Retrenchment and Voting in Health Policy." *Journal of Health Politics, Policy and Law* 44(3): 423–454.

Hern, Erin. 2017. "Better Than Nothing: How Policies Influence Political Participation in Low-Capacity Democracies." *Governance* 30(4): 583–600.

———. 2019. *Developing States, Shaping Citizenship: Service Delivery and Political Participation in Zambia*. Ann Arbor, MI: University of Michigan Press.

Hertel-Fernandez, A. 2018. "Policy Feedback as Political Weapon: Conservative Advocacy and the Demobilization of the Public Sector Labor Movement." *Perspectives on Politics* 16(2): 364–379.

Im, Dong-Kyun, and Tianguang Meng. 2016. "The Policy–Opinion Nexus: The Impact of Social Protection Programs on Welfare Policy Preferences in China." *International Journal of Public Opinion Research* 28(2): 241–268.

Jacobs, L.R., and Desmond King. 2016. *Fed Power: How Finance Wins*. New York, NY: Oxford University Press.

Jacobs, L. R., and S. Mettler. 2018. "When and How New Policy Creates New Politics: Examining the Feedback Effects of the Affordable Care Act on Public Opinion." *Perspectives on Politics* 16(2): 345–363.

Jacobs, L.R., S. Mettler, and L. Zhu. 2019. "Affordable Care Act Moving to New Stage of Public Acceptance." *Journal of Health Politics, Policy and Law* 44(6): 911–917.

Kogan, Vladimir. 2021. "Do Welfare Benefits Pay Electoral Dividends? Evidence from the National Food Stamp Program Rollout." *The Journal of Politics* 83(1), 58–70.

Kumlin, Staffan. 2002. *The Personal and the Political: How Personal Welfare State Experiences Affect Political Trust and Ideology*. Department of Political Science Statsvetenskapliga Institutionen.

Lerman, Amy E. 2019. *Good Enough for Government Work: The Public Reputation Crisis in America (and What We Can Do to Fix It)*. Chicago, IL: University of Chicago Press.

Leutert, Wendy. 2021. "Innovation through Iteration: Policy Feedback Loops in China's Economic Reform." *World Development* 138:1–11.

Lowi, Theodore J. 1972. "Four Systems of Policy, Politics, and Choice." *Public Administration Review* 32(4): 298–310.

Lu, Xiaobo. 2014. "Social Policy and Regime Legitimacy: The Effects of Education Reform in China." *American Political Science Review* 108(2): 423–437.

MacLean, Lauren. 2011. "State Retrenchment and the Exercise of Citizenship in Africa." *Comparative Political Studies* 44(9): 1238–1266.

Mead, Lawrence M. 2004. "The Great Passivity." *Perspectives on Politics* 2(4): 671–675.

Mettler, Suzanne. 1998. *Dividing Citizens: Gender and Federalism in New Deal Public Policy*. Ithaca, NY: Cornell University Press.

———. 2005. *Soldiers to Citizens: The G.I. Bill and the Making of the Greatest Generation*. New York, NY: Oxford University Press.

———. 2011. The Submerged State: How Invisible Government Policies Undermine American Democracy. Chicago, IL: University of Chicago Press.

———. 2016. "The Policyscape and the Challenges of Contemporary Politics to Policy Maintenance." *Perspectives on Politics* 14(2): 369–390.

Mettler, Suzanne, and Mallory SoRelle. 2018. "Policy Feedback Theory." In *Theories of the Policy Process*, 4th Edition, edited by Christopher M. Weible and Paul A. Sabatier. London: Routledge. 151–181.

Mettler, Suzanne, and Jeffrey M. Stonecash. 2008. "Government Program Usage and Political Voice." *Social Science Quarterly* 89(2): 273–293.

Mettler, Suzanne, and Eric Welch. 2004. "Civic Generation: Policy Feedback Effects of the GI Bill on Political Involvement over the Life Course." *British Journal of Political Science* 34(3): 497–518.

Michener, Jamila D. 2017. "People, Places, Power: Medicaid Concentration and Local Political Participation." *Journal of Health Politics, Policy and Law* 42(5): 865–900.

———. 2018. *Fragmented Democracy: Medicaid, Federalism, and Unequal Politics*. Cambridge: Cambridge University Press.

———. 2019a. "Policy Feedback in a Racialized Polity." *Policy Studies Journal* 47(2): 423–450.

———. 2019b. "Medicaid and the Policy Feedback Foundations for Universal Healthcare." *The ANNALS of the American Academy of Political and Social Science* 685(1): 116–134.

———. 2020. "Power from the Margins: Grassroots Mobilization and Urban Expansions of Civil Legal Rights." *Urban Affairs Review* 56(5): 1390–1422.

Michener, Jamila D, SoRelle Mallory, and Chloe Thurston. 2020. "From the Margins to the Center: A Bottom-Up Approach to Welfare State Scholarship." *Perspectives on Politics*, 1-16. doi:10.1017/S153759272000359X

Morel, D. 2018. *Takeover: Race, Education, and American Democracy*. New York, NY: Oxford University Press.

Morgan, Kimberly. 2006. *Working Mothers and the Welfare State: Religion and the Politics of Work-Family Policies in Western Europe and the United States*. Palo Alto, CA: Stanford University Press.

Morgan, Kimberly, and Andrea Louise Campbell. 2011. *The Delegated Welfare State: Medicare, Markets, and the Governance of Social Policy*. New York, NY: Oxford University Press.

Nuamah, Sally A. 2021. "The Cost of Participating while Poor and Black: Toward a Theory of Collective Participatory Debt." *Perspectives on Politics* 19 (4), 1115–1130.

Orloff, Ann Shola. 1993. *Gender and the Social Rights of Citizenship: The Comparative Analysis of Gender Relations and Welfare States*. Madison, WI: University of Wisconsin.

Orren, Karen, and Skowronek, Stephen. 2004. *The Search for American Political Development*. Cambridge: Cambridge University Press.

Pacheco, Julianna. 2013. "Attitudinal Policy Feedback and Public Opinion: The Impact of Smoking Bans on Attitudes toward Smokers, Secondhand Smoke, and Anti-Smoking Policies." *Public Opinion Quarterly* 77(3): 714–734.

Pacheco, Julianna, Jake Haselswerdt, and Jamila Michener. 2020. *RSF: The Russell Sage Foundation Journal of the Social Sciences* 6(2): 114–130.

Park, D. K., Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls." *Political Analysis* 12(4): 375–385.

———. 2006. "State Level Opinions from National Surveys: Poststratification Using Multilevel Logistic Regression." In *Public Opinion in State Politics,* edited by Jeffrey E. Cohen. Stanford, CA: Stanford University Press.

Patashnik, Eric M. 2008. *Reforms at Risk: What Happens After Major Policy Changes Are Enacted*. Princeton, NJ: Princeton University Press.

Patashnik, Eric M., and Julian Zelizer. 2013. "The Struggle to Remake Politics: Liberal Reform and the Limits of Policy Feedback in the Contemporary American State." *Perspectives on Politics* 11(4): 1071–1087.

Pierson, Paul. 1993. "When Effect becomes Cause: Policy Feedback and Political Change." *World Politics* 45 (4): 595-628.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

Rocha, Rene R., Benjamin R. Knoll, and Robert D. Wrinkle. 2015. "Immigration Enforcement and the Redistribution of Political Trust." *The Journal of Politics* 77(4): 901–913.

Rose, Deondra. 2018. *Citizens by Degree: Higher Education Policy and the Changing Gender Dynamics of American Citizenship*. New York, NY: Oxford University Press.

Rosenthal, Aaron. 2021a. "Conflicting Messages: Multiple Policy Experiences and Political Participation." *Policy Studies Journal* 49(2): 616–639.

Rosenthal, Aaron. 2021b. "Submerged for Some? Government Visibility, Race, and American Political Trust." *Perspectives on Politics 19(4):* 1098–1114

Schattschneider, E. E. 1935. *Politics, Pressure, and the Tariff*. New York, NY: Prentice Hall.

Schneider, Anne, and Helen Ingram. 1993. "Social Construction of Target Populations: Implications for Politics and Policy." *American Political Science Review* 87 (2): 334–347.

———. 2019. "Social Constructions, Anticipatory Feedback Strategies, and Deceptive Public Policy." *Policy Studies Journal* 47(2): 206-236.

Shanks, Delphia, and Mallory SoRelle. 2021. "The Paradox of Policy Advocacy: Philanthropic Foundations, Public Interest Groups, and Second-Order Policy Feedback Effects." *Interest Groups & Advocacy 10(2): 137–157*.

Shanks-Booth, Delphia, and Suzanne Mettler. 2019. "The Paradox of the Earned Income Tax Credit: Appreciating Benefits but not their Source." *Policy Studies Journal* 47(2): 300-323.

Skocpol, T. 1992. *Protecting Soldiers and Mothers: The Politics of Social Provision in the United States, 1870s-1920s*. Cambridge, MA: Belknap Press of Harvard University Press.

Skocpol, T., and Finegold, K. 1982. "State Capacity and Economic Intervention in the Early New Deal." *Political Science Quarterly* 97(2): 255–278.

Small, Mario Luis. 2009. "How Many Cases Do I Need?' On Science and the Logic of Case Selection in Field-based Research." *Ethnography* 10(1): 5–38.

SoRelle, Mallory E. 2020. *Democracy Declined: The Failed Politics of Consumer Financial Protection*. Chicago, IL: University of Chicago Press.

Soss, Joe. 1999. "Lessons of Welfare: Policy Design, Political Learning, and Political Action." *American Political Science Review* 93 (2): 363–380.

Soss, Joe, Richard C. Fording, and Sanford F. Schram. 2011. *Disciplining the Poor: Neoliberal Paternalism and the Persistent Power of Race*. Chicago, IL: University of Chicago Press.

Soss, Joe, and Sanford F. Schram. 2007. "A Public Transformed? Welfare Reform as Policy Feedback." *American Political Science Review* 101(1): 111–127.

Springer, Melanie Jean. 2014. How the States Shaped the Nation: American Electoral Institutions and Voter Turnout, 1920-2000. Chicago, IL: University of Chicago Press.

Stokes, Leah Cardamore. 2016. "Electoral Backlash against Climate Policy: A Natural Experiment on Retrospective Voting and Local Resistance to Public Policy." *American Journal of Political Science* 60(4): 958–974.

———. 2020. *Short Circuiting Policy: Interest groups and the Battle over Clean Energy and Climate Policy in the American States*. New York, NY: Oxford University Press.

Thelen, Kathleen. 1999. "Historical Institutionalism in Comparative Politics." *Annual Review of Political Science* 2(1): 369–404.

Thurston, Chloe N. 2018. *At the Boundaries of Homeownership: Credit, Discrimination, and the American State*. Cambridge: Cambridge University Press.

Trounstine, Jessica. 2018. *Segregation by Design: Local Politics and Inequality in American Cities*. Cambridge: Cambridge University Press.

Walker, Hannah L. 2014. "Extending the Effects of the Carceral State: Proximal Contact, Political Participation, and Race." *Political Research Quarterly* 67(4): 809–822.

———. 2020. *Mobilized by Injustice: Criminal Justice Contact, Political Participation, and Race*. New York, NY: Oxford University Press.

Walker, Hannah, Marcel Roman, and Matt Barreto. 2020. "The Ripple Effect: The Political Consequences of Proximal Contact with Immigration Enforcement." *Journal of Race, Ethnicity and Politics* 5(3): 537–572.

Walker, Jack L. 1983. "The Origins and Maintenance of Interest Groups in America." *American Political Science Review* 77(2): 390–406.

Weaver, Kent R. 2010. "Paths and Forks or Chutes and Ladders: Negative Feedbacks and Policy Regime Change." *Journal of Public Policy* 30: 137–162.

Weaver, Vesla M., and Amy E. Lerman. 2010. "Political Consequences of the Carceral State." *American Political Science Review* 104(4): 817–833.

White, Ariel. 2019. "Family Matters? Voting Behavior in Households with Criminal Justice Contact." *American Political Science Review* 113(2): 607–613.

Williamson, Abigail Fisher. 2020. "Intergovernmental Policy Feedback and Urban Responses to Immigrants." *PS: Political Science & Politics* 53(1): 20–24.

5 Advocacy Coalition Framework

Advice on Applications and Methods

Adam Douglas Henry, Karin Ingold, Daniel Nohrstedt, and Christopher M. Weible

Introduction

The Advocacy Coalition Framework (ACF) is a widely applied theoretical framework that has been used to understand and explain the dynamics of the policy process. Since Paul Sabatier and Hank Jenkins-Smith originally developed the ACF in the 1980s (Sabatier, 1988; Sabatier & Jenkins-Smith, 1993), the ACF has blossomed into a vibrant research program with applications to many different policy issues. It has been regularly applied in countries in North America and Western Europe and increasingly to countries in Asia, Africa, and South America (e.g., Weible, Sabatier, & McQueen, 2009; Pierce, Peterson, Jones, Garrad, & Vu, 2017; Li & Weible, 2019; Calmon & Araujo, 2021). The ACF has traditionally been applied in areas of high conflict frequently involving scientific and technical controversies; however, recent work also spans issues with varying degrees and types of conflict.

One of the engines of the ACF's success as a framework has been scholars' emphasis on conceptual clarity and hypothesis testing, critique, and, sometimes, revision. This approach is enabled by strong theoretical conceptualization, operationalization, and measurement of key variables. The ACF has been applied using a variety of different research methods. Hence, there is not one single way to apply the framework; it can support numerous research designs, forms of data, and means of analysis and modeling. This creates challenges for both experienced and new researchers planning to apply the framework. Any useful methodological approach must be sufficiently well articulated, transparent, and replicable to facilitate cumulative, intersubjective learning and theoretical advancement.

The purpose of this chapter is to explicate a range of methods of data gathering and analysis that apply to the study of core concepts embedded within the ACF, including subsystems, policy actors, beliefs, coalitions, learning, and policy change. Our goal is to provide scholars with guidance on how to apply the ACF to make descriptive sense of one or more policy process cases, to test hypotheses, to advance policy process theory, and for other purposes. We review prior approaches and discuss best practices. Given the common types of data sources available for analyzing policy change, coalitions, and learning, we summarize exemplary studies in Table 5A.2

and indicate the advantages and disadvantages of different data sources in Table 5A.3.

Conceptual Overview

Within many policy processes, change is slow-moving and often mired in conflict. The ACF provides an understanding of why this is the case, with a recognition that major policy change tends to be the exception rather than the norm and that minor (or incremental) policy change is far more common. This is explained by the behavior of policy actors, who organize themselves into advocacy coalitions that compete to translate their core beliefs into public policy.

The ACF views public policy as a representation of beliefs held by coalitions. These beliefs concern the fundamental goals of public policy, the proper role of government in managing societal problems, the definition and priority of problems to be addressed, and the design and selection of instruments for addressing those problems (see the description of ACF's model of belief systems later in this chapter). When one views public policy dynamics in terms of coalitions competing to translate their beliefs into public policy, then it follows that major policy change requires a shift in the beliefs of those in power. Beliefs, however, tend to be very stable and resistant to change even in the face of evidence that contradicts them. Thus, major policy change is generally only possible through shifting power to a different coalition with different policy core beliefs through shocks that undermine support for existing policies or that provide opportunities to capitalize on opening for change. However, while belief systems show rigidity, they also show some receptivity to information. Learning – in the form of enduring change in policy beliefs – can occur over time among coalition allies with a pronounced bias toward reinforcing prior beliefs and only occasionally with bigger shifts occurring between coalition opponents. Thus, while paths to major policy change often require power shifts, they can also follow from learning within and among coalitions.

As a theoretical framework, the ACF has three conceptual focal points: coalitions, policy change, and learning. These concepts are sometimes studied together in "whole-framework" applications of the ACF, but often scholars will focus their attention on only one of these concepts at a time. This focused attention can help to develop and elaborate our understanding of policy processes; however, it is important to be mindful of the relations that these concepts have with each other. Supporting these three theoretical emphases are cross-cutting categories of policy subsystems, policy actors, and belief systems, which we describe below (see also Table 5A.4).

Cross-cutting Concepts

The theoretical pillars of coalitions, policy change, and learning are tied together by three synthetic concepts that require serious treatment in any

empirical application of the framework. These include policy subsystems, policy actors, and belief systems.

Policy Subsystem

The ACF's primary unit of analysis is the policy subsystem. The constituent elements of policy subsystems include a geographic scope, a topical area, and the array of policy actors involved (see description of policy actors below).[1]

Topical Scope

A subsystem can best be perceived through an understanding of an issue such as climate change, migration, education, or the impacts of the management of a river basin. For example, we might start by gathering an initial historical narrative, conducting interviews, or both to understand policy processes for a particular issue, including actors, events, and controversies. This includes an exploration of who is or is not involved (see also *policy actors* below). It might also include analyzing news media, social media, miscellaneous documents (such as public policies), and public hearings (e.g., Carlson et al., 2019). As we gain confidence in our understanding of the characteristics of the policy subsystem, we might then focus on achieving our research objectives, possibly related to exploring or testing assumptions concerning coalitions, learning, and policy change, all of which usually involve a long-term focus with observations spanning a decade or more.

Geographic Scope

Issues are linked to the policy process with an explicit geographic scope, such as one specific decisional level, or in one jurisdiction, area (like river catchment), or community. Note that polity and political styles can influence geographic scope, and explain why, for instance, in one country an issue is tackled at the national level, and in another country at the subnational level (see Weible, Heikkila, Ingold, & Fischer, 2016, when comparing "fracking" politics in 12 countries around the globe).

 The ACF is applied to issues where a subsystem exists – a precondition that is not satisfied for many policy issues. Policy subsystems emerge because formulating and implementing public policies and achieving desirable outcomes requires both specializations among policy actors and dedication of governmental resources through a diversity of institutional structures (e.g., as might be found in various administrative arrangements).[2] Two factors distinguish a subsystem from the "policy primeval soup" of potential subsystems (Kingdon, 1984). The first factor is the identification of a core topical area (e.g., harmful air emissions) that is both durable and recognized by actors who have sufficient stake in the issue to form coalitions to pursue their preferred policies over substantial periods (usually a decade or more). The second factor is institutional structures (e.g., government-sponsored

programs) that exist to justify the expenditure of coalition resources to pursue policy objectives. One example of a subsystem is the forest governance policy subsystem of Papua New Guinea that operates at the national level and involves people inside and outside of government seeking to influence its development (Babon et al., 2014).

Policy Actors

Policy actors is a term used to describe those individuals and organizations engaged either directly or indirectly in a policy subsystem. The list of who or what organizations might be policy actors is nearly limitless and depends on the characteristics of the policy venue, policy subsystem, and political system. For example, past research has found that policy actors include scientists and researchers, elected and career bureaucrats, elected government officials, journalists, business and corporate representatives, nonprofit leaders, citizens, and more.

Two common criteria for identifying policy actors include reputation and professional engagement. Not every individual engaged in politics and who participates within or around a policy subsystem qualifies as a relevant policy actor. However, engagement and reputation helps to distinguish between subsystem participants and more peripheral policy actors who have neither the position nor the ideological motivation to participate.

Reputation means that other policy actors or experts see one organization or individual as influential or important in a given policy process. Reputational power helps to identify those actors who dispose of resources and power to impact policy change decisively (Knoke, Pappi, Broadbent, & Tsujinaka, 1996).

Policy actors spend a large part of their professional time engaged in the subsystem and participate in subsystem activities regularly over time. Regular participation helps form relationships with other actors. However, actors with less regular participation may still be relevant but mainly when included in coalitions. For example, a policy actor involved in the short-term or intermittently and who brings substantial resources can be relevant; this, in turn, drives efforts to recruit members to an advocacy coalition (Jenkins-Smith, 1990, see Table 5A.1 for different categories of policy actors). Resources and techniques useful for identifying policy actors often include documents (e.g., social and news media, reports, minutes from policy venues) and interviews.

Belief Systems

Belief systems are a conceptual focal point of the ACF. The term "beliefs" has long been used as a covering term that includes many types of cognitions of interest in various social science fields, including preferences, values, and beliefs in the classical sense of the perception that two phenomena are causally related. According to the ACF, the important distinction between these myriad cognitions is their scope relative to the subsystem. There are different

ways to collect information about policy actors' beliefs, e.g., interviews, surveys, public documents, and news/social media material.

A Hierarchical Model of Beliefs

The ACF applies a three-tiered hierarchy of belief systems: *Deep core beliefs* are fundamental normative orientations or worldviews; *policy core beliefs* are the translation of deep core beliefs to one specific policy subsystem; and *secondary beliefs* are instrumental means for realizing the policy-core (see Table 5.1). While belief systems are expected to show stability, a spectrum of change is posited from secondary beliefs being the most likely to change to deep core beliefs being the least likely to change.

It is not always clear whether a specific belief belongs to the policy core or secondary belief categories. For example, a "ban on fracking" might belong to an actor's core belief in how to shape policymaking in the unconventional oil and gas subsystem (see again Weible et al., 2016). Other actors (probably in other countries or contexts) see a ban as one rather "technical" type of state intervention and would thus be categorized as a secondary belief.[3]

Beliefs may be measured using several established research methodologies; however, there are two major approaches to belief measurement.[4]

Belief measurement through unsolicited written or verbal statements. Beliefs may be measured through analyzing policy actors' unsolicited written or verbal statements. The raw data may include beliefs as expressed in published literature (including technical reports, organizational websites, and other "grey

Table 5.1 Belief systems in the ACF

	Definition	Examples
Deep core beliefs	Fundamental normative orientations	• Political ideologies • Normative values (e.g., religious beliefs) • Identities • Cultural orientations • Basic priorities (e.g., freedom vs. security) • Views of human nature
Policy core beliefs	Normative and empirical beliefs concerning policy subsystems	• General goals for a policy subsystem • Positions on general policy solutions and policy instruments • Problem severity and cause • Role of government in subsystem affairs (e.g., vs. markets)
Secondary beliefs	Instrumental beliefs or beliefs about a subset of a policy subsystem	• Instrumental means for achieving policy-core ends (or goals) • Relative weight of various causal mechanisms of problems • Solutions and problems associated with part of a policy subsystem

literature" sources), media reports of policy issues, or transcripts of open-ended interviews with policy participants outside of a research study. The set of relevant beliefs and actors' positions on these beliefs would need to be inferred through systematic coding of the raw data. This was the approach used in some of the original ACF studies (Jenkins-Smith, St. Clair, & Woods, 1991), and recent methodological innovations have expanded our ability to effectively understand belief systems through coding expressed beliefs. For instance, Discourse Network Analysis (Kukkonen, Ylä-Anttila, & Broadbent, 2017) is an emerging method to extract belief systems, characterized as networks, from collections of documents.

Belief measurement through solicited self-reports. An alternative approach to belief measurement relies on direct solicitation of beliefs, usually by interacting directly with policy actors as research participants. This category includes belief measurement using surveys and structured interviews. This approach will generally involve creating a measurement instrument that presumes a set of relevant beliefs that fully describes the belief system and then assigning a position on each belief for every research participant. For instance, policy actors may be invited to respond to a survey that asks them to rate their level of agreement or disagreement – often on an ordinal Likert scale – to a series of statements representing discrete beliefs.

Three Meta-Theoretical Key Concepts

Advocacy Coalitions

Coalitions are the ACF's most studied phenomena (Weible et al., 2009; Pierce et al., 2017). Identifying and studying advocacy coalitions requires, at a minimum, data on policy actors' belief systems and, depending on the scope and focus of the research, data on coordination, resources, and strategies over time. Advocacy coalitions are important to study in themselves because they represent the political behavior of policy actors repeatedly engaging in policy issues over time, focusing on the ways they join forces to influence change or maintain stasis. Thus, they become the vehicles in which policy actors make sense of their belief systems, develop their perceptions of allies and opponents, and coordinate political behavior. As policy actors engage in coalitions, they simultaneously contribute to the evolution of a subsystem through defining and redefining problems and developing solutions that might reshape, for example, the institutions structuring that subsystem. As researchers, we study coalitions because they can provide an indirect indicator of the intensity of conflict and, thus, suggest the likelihood of policy change. We also study coalitions to explore how belief systems remain constant or shift over time between individual policy actors and groups of policy actors in coalitions.

Coalitions are not formal entities. They are usually informal or implicit alliances of policy actors somehow engaging to influence the course and

coverage of a policy subsystem. Advocacy coalitions are formed by policy actors when there is policy conflict or disagreement; the ACF defines an advocacy coalition as a group of policy actors with shared beliefs and who coordinate to influence policy. Additionally, advocacy coalitions are also characterized in terms of their resources and level of stability (Henry, 2011; Weible, Ingold, Nohrstedt, Henry, & Jenkins-Smith, 2020). At a minimum, measuring a coalition requires identifying policy actors and shared beliefs (see below and Table 5A.4).

Shared Beliefs

The ACF refers to belief systems as the "glue" that holds coalitions together. A necessary condition for a group of actors to be members of the same coalition is that they share a common set of core beliefs. Most studies of advocacy coalitions, therefore, start with identifying policy actors' belief systems (see Table 5.1; McDougall, 2006; Sotirov & Winkel, 2016) and then cluster policy actors based on their shared beliefs. The exact technique for clustering policy actors into coalitions varies based on the type of data. We outline two common approaches that rely on quantitative data sources.

Belief clustering. The first approach to identifying belief coalitions uses data on beliefs for all policy actors to find natural centers within this distribution. The various centers of the distribution are viewed as belief positions that characterize the central tendency or idealized version of a coalition's belief system. This technique uses data gathered from individual policy actors (since beliefs are fundamentally an attribute of policy actors), however can be conceptually decoupled from the actors themselves.

Belief clustering may employ a variety of dimension-reduction methods, including k-means clustering, principal components analysis, and factor analysis (Ansell, Reckhow, & Kelly, 2009; Schmid, Sewerin, & Schmidt, 2020). Moreover, it is important to be mindful that some degree of researcher judgment is needed to choose an appropriate method of belief aggregation. Variations in judgement may pose a threat to the reliability of findings.

Beliefs as relations. An alternative method to the identification of belief coalitions is to represent beliefs as a relational variable, where policy actors are related in terms of their similarity of difference in beliefs. This yields a representation of shared beliefs in network terms, from which groups of actors with shared beliefs may be identified using any method of cohesive subgroup detection from network analysis. These may include relatively simple methods of subgroup detection such as clique analysis, to more sophisticated methods such as modularity and community detection.

Importantly, there are two major approaches to the construction of a shared belief network, from which belief coalitions may be identified. The first is to infer belief similarity from the independently-measured belief attributes of policy actors. Thus, for instance, two policy actors giving similar responses to belief items on a survey would be assumed to be "close" within

a shared belief network. A distinct approach would use network measurement methods to ask policy actors directly to nominate other policy actors with whom they have shared or divergent beliefs (Fischer, Ingold, Sciraini, & Varone, 2016).

Researchers rarely directly ask policy actors about their coalition membership. This for two reasons: first, much preliminary case knowledge would be required to ask about coalition membership; and second, advocacy coalitions, as noted above, are most often informally structured, where participation might not seem obvious to their members.

Coordination

The ACF argues that advocacy coalition members engage in a non-trivial degree of coordination based on shared beliefs. Coordination can be restricted to the weak interaction among coalition members, such as monitoring and responding to each other's behaviors, to strong interactions, such as developing joint plans (Zafonte & Sabatier, 1998; Calanni, Siddiki, Weible, & Leach, 2014; Aamodt & Stensdal, 2017; de Medeiros & Corrêa Gomes, 2019; Heikkila, Berardo, Weible, & Yi, 2019; see Table 5.2).

Often, coalitions are identified by asking policy actors with whom they interact on policy issues; this yields a measurement of policy networks. Researchers usually rely on three techniques for network measurement: (1) a roster of other policy actors to identify their network partners; (2) an opportunity to name network partners through free recall; and (3) a hybrid approach using a combination of a roster and free recall (see Henry, Lubell, & McCoy, 2012).

Recent studies also consider activities on social media and mass media as proxies for coordination and strategy (Elgin, 2015), for example, by documenting various joint activities by policy actors to influence public opinion (Nohrstedt & Olofsson, 2016). It is important to note that coordination between coalitions may also provide useful information about the subsystem, and whether the subsystem tends to be more or less adversarial or cooperative (see Table 5A.1).

Coalition Resources and Strategies

When resources and strategies have the focus of analysis, they have been mostly measured as self-reported by policy actors as found in surveys (e.g., Elgin & Weible, 2013; Weible & Heikkila, 2016) or been inferred from publicly available information (such as websites, annual or financial reports).

Distinct from resources, strategies embrace activities (rather than capacities) to influence policy processes and their outputs (Nohrstedt, 2011; Montefrio, 2014). One common strategy is the engagement in coordination and joint efforts to influence decision-making. Other strategies may be imagined and measured through media or document analysis, as well as surveys (see Table 5.2).

Table 5.2 Coordination, resources, and strategies of advocacy coalitions

	Definition	*Examples*
Strong coordination	Activities agreed upon and acknowledged by coalition actors	• Formulation and implementation of a common plan • Sharing resources
Weak coordination	Activities that are in sync toward achieving a common goal but are not jointly agreed upon	• Monitoring the use of resources and deployment of strategies and altering behavior accordingly • Understanding the positions and niches of allies and behaving in a complementary manner
Resources	Accessible capacity of policy actors to influence policy processes	• Supportive members of the public • Number of stable allies • Access to elected officials or those with authority • Money • Personnel • Scientific and technical information • Leadership
Strategies	Activities of policy actors to influence policy processes	• Coordination • Public attention and presence in the media • Lobbying • Co-signing amicus curiae briefs (or similar alliances)

Sources: Fenger and Klok (2001); Smith (2000); Weible and Ingold (2018); Weible et al. (2020).

Research Foci and Challenges in the Study of Coalitions

ACF research commonly focuses on two aspects of advocacy coalitions. The first deals with identifying, describing, and understanding the creation, structure, and evolution of advocacy coalitions. An illustration is provided by work using discourse network analysis (Leifeld, 2013; Rennkamp, Haunss, Wongsa, Ortega, & Casamadrid, 2017). Through the systematic coding of actors and their beliefs, researchers can draw the ideological or semantic development of one specific policy discourse and identify coalitions of actors based on shared beliefs. Some studies emphasize the role of single actors within coalitions, such as experts (Ingold & Gschwend, 2014), parties (Afonso, 2014), or international organizations (Kukkonen, Ylä-Anttila, Swarnakar, Broadbent, Lahsen, & Stoddart, 2018). These studies might explore the level of coalition polarization before, during, and after a policy change. For example, Koebele, Bultema, and Weible (2020) explored coalition structures before and after a policy change in the Lake Tahoe Basin in the United States. Similarly, studies have explored coalitions and their stability and defection

longitudinally around multiple instances of policy change (see, e.g., Jenkins-Smith et al., 1991).

The second aspect relates coalitions, their internal and external power (often linked to characteristics of a coalition, including their resources, in comparison to their opposition), and conflict structures to policy change (and vice versa). For example, Nohrstedt (2008, 2011) adopts an encompassing research design and emphasizes the impact of (external) events on subsystem and coalition responses and eventually on policy change and stasis. Generally, scholars relate coalitions to the emergence and development to the introduction of new policies (Ingold, 2011), as well as the revision of existing policies (Harrinkari, Katila, & Karppinen, 2016). Coalitions are also, but less frequently, related to processes of policy implementation (Barnes, van Laerhoven, & Driessen, 2016). For instance, Gralepois et al. (2016) compare flood defense strategies among six European countries and explain differences therein through institutional and structural factors as well as advocacy coalitions. This is just one example that combines the ACF with other frameworks to identify different factors of change.

Policy Change

One of the theoretical emphases in the ACF is in the study of policy change (or stasis). The emphasis on policy changes brings methodological challenges related to description (where and how to identify policy change) and explanation (how to explain policy change). There are many ways in which ACF studies have studied, and can study, policy change. The ACF posits four pathways to policy change in the form of external shocks, internal shocks, learning, and negotiated agreements (see definitions and hypothesis, Jenkins-Smith, Nohrstedt, Weible, & Ingold, 2018).

Along with these pathways, the ACF also posits that policy change can occur with the change in the governing coalition – hence, a transition from one belief system to another (ibid). Theoretically, policy change does not always follow all external or internal shocks, instances of learning, or negotiated agreements. We expect, however, that at least one of these pathways will precede any instance of policy change, and sometimes, these pathways can interact with each other. This variance in policy change remains an important study topic within the policy sciences more generally where the ACF can offer theoretical insights to guide empirical research (Dunlop & Radaelli, 2017; Moyson, Scholten, & Weible, 2017; Pierce, Peterson, & Hicks, 2020). Broadly speaking, policy change means an alteration in public policy content or design. As described further below, any ACF research studying public policy change must account for the type of public policy and the policy venue wherein that policy change occurred (e.g., legislature or rulemaking agency) as well as the scope of policy change focusing on the magnitude of change, from minor to major (Sabatier & Jenkins-Smith, 1999, p. 147).

Type of Public Policy and Policy Venue

The ACF uses "public policy" as an umbrella term that refers to any decision or non-decision of government (or similar authority).[5] Any decision to adopt or reject a public policy occurs in policy venues (e.g., legislature, government agency, judiciary, executive) at any level of government or the supranational level transcending nations.[6] Subcategorizations of public policies include laws, regulations, programs, legal decrees, executive orders, and so on. Their concrete form depends upon the country- or constituency-specific polity, political styles, and culture (e.g., laws versus acts, decrees versus ordinances).

Policy change may include revisions in policy core components of government programs, termination of programs, or launching of new programs. For example, Smith (2000) studied policy change as administrative and procedural reform in an industrial pollution policy subsystem in the United Kingdom. For all forms of public policy and policy change, researchers should collect documents for analysis and make comparisons through time. This could involve finding the policy documents on government websites or archives. Clarity concerning the characterization of the type of public policy is important to ensure reliability and encourage comparative research within the ACF research program. And this is similar for the characterization of the policy venue.

Rarely studied under the ACF is the characteristics of the policy venue in which the policy change occurred. Any categorization of public policy involves a designated policy venue, and all policy venues consist of their institutional structures and processes, such as with multi-stakeholder policy venues characterized by rules favoring consensus and conflict mitigation (see Koebele, 2020). Thus, policy venues affect how policies change. Also, policy venues affect what ideas are even considered by keeping some ideas off their agendas and accepting others. If an idea were translated into public policy, the policy venue would also influence the content and design of the adopted public policy. For example, all policy venues have boundaries to their authorities, and the policy decisions that emerge will reflect that authority. As one illustration, the Supreme Court of the United States lacks authority to tax or enforce their policies, and their court decisions, as policy change, reflect this boundary.

Minor–Major Policy Change

Any adoption of public policy represents an instance of policy change. A policy change might lead to actual changes in outcomes of any kind or it might not lead to any changes at all. When exploring policy change, we are concerned with whether the policy change, as written in the content or design of the policy, constitutes marginal or substantial change in the structure of the policy subsystem. Outcomes of policy change are not the explanatory focus of the ACF, except when outcomes become subject of

conflict and provide the impetus for coalitions to revise their beliefs and strategies (Sabatier & Jenkins-Smith, 1999).

The study of public policy has long explored variations in the trajectory and magnitude of policy change, from incremental to punctuated change (Lindblom, 1959; Hall, 1993; Baumgartner & Jones, 2010). The ACF contributes to these traditions with its distinction between minor policy change and major policy change.

The definitions of minor and major change derive from the premise that policy can be conceptualized in the same way as belief systems. Major policy change refers to a change in the priorities or goals of the policy subsystem, often articulated through government programs. We often identify such change as a change in the policy core of the policy subsystem. Minor change refers to change in either the means to achieve priorities or goals, or a minor subset of the policy subsystem. We often refer to minor change as a change in the secondary beliefs (see Table 5.1) of the policy subsystem. Determining whether a policy change is major or minor usually requires multiple sources of data and methods of analysis, such as document analysis of the written content of the change combined with a description of the policy subsystem from interviews or other secondary sources.

The distinction between major and minor policy change is a theoretical simplification to guide empirical analysis of policy change and stability. In reality, the scope and magnitude of policy change are often contested and a source of debate and conflict between advocacy coalitions. This is partially due to their different beliefs; since members of different coalitions have different beliefs about the seriousness and causes of problems, they often make different interpretations of the scope, sufficiency, and effectiveness of policy change for addressing those problems.

The number of analytical techniques for explaining policy change far exceeds the space in this chapter. However, we highlight two common techniques:

Process tracing. Process tracing is a prominent method of within-case analysis used to draw an inference based on causal mechanisms and the temporal sequence of events (George & Bennett, 2005; Collier, 2011). This technique is most suitable for single-case instances of policy change often involving in-depth interviews or analysis of documents, news media, and social media. For example, Heinmiller (2016) identified through process tracing two advocacy coalitions in Alberta's water quality subsystem and explained the inclusion of eco-friendly instruments through the increased "soft" power by one pro-environment coalition.

Qualitative comparative analysis (QCA). QCA is a method for systematic comparison of a small number of cases. It allows for "combinatory" causality (or equifinality), i.e., different combinations of causal conditions can lead to one particular outcome (Ragin, 1987; Rihoux & Ragin, 2009; Fischer & Maggetti, 2017). An example of this type of analysis is Fischer's (2015a) work on eleven Swiss policy subsystems, where he explains the effect of Europeanization on change and coalition structures.

Research Foci and Challenges in the Study of Policy Change

We highlight two approaches by which ACF scholars have directly described or explained policy change. The first approach involves the study of a singular instance of policy change for a policy subsystem. This usually occurs using techniques akin to single-case studies. For example, Khayatzadeh-Mahani, Breton, Ruckert, and Labonté (2017) conduct an in-depth case study of the inability to ban shisha smoking in public places in Iran. The researchers interviewed 24 policy actors and analyzed documents to assess the lack of learning, the failed attempt of an exploitive coalition to capitalize on events, and the overall institutional structure in Iran's autocratic state.

The second approach involves the comparative analysis of similar or the same type of policy change across similar policy subsystems. For example, Sotirov and Winkel (2016) explored policy change in forest policy subsystems in Germany and Bulgaria. Relying primarily on 73 interviews, these researchers explored the mechanisms of policy change in both countries with an emphasis on the role of different cultural biases (drawn from Cultural Theory). Policy change remains a popular focus in previous ACF applications (Pierce et al., 2017), which provide numerous examples of studies investigating one or several of the four pathways.

Like any study involving the ACF, the strength of its focus on policy change is in understanding the case (or cases) themselves.[7] This requires a keen understanding of the contextual setting, which affects all steps above. Projects and studies set on explaining policy change through the lens of the ACF also have to take steps to properly define and empirically measure key drivers of policy change, including learning, negotiated agreements, and internal and external events. Finally, since these hypothesized pathways are based on assumptions concerning the structure, behavior, and relationships of coalitions, any analysis into policy change also requires going through the steps outlined above for documenting subsystems and coalitions.

Learning

Learning is a core concept within the ACF, yet the least studied (Pierce et al., 2017) and most difficult to conceptualize given the great diversity of approaches that exist in the social sciences for the study of this concept. In the ACF, the term policy-oriented learning has been long defined as

> enduring alterations of thought or behavioral intentions that result from experience and which are concerned with the attainment or revision of the precepts of the belief system of individuals or of collectives (such as advocacy coalitions).
> (Sabatier & Jenkins-Smith, 1993, as cited in Pierce et al., 2017)

Policy-oriented learning is distinguished from other forms of learning considered across the social sciences in that the object of learning (i.e., what

is learned) is policy relevant. More specifically, policy-oriented learning results in a shift in beliefs – whether at the level of secondary, policy core, or deep core beliefs. Policy participants are viewed as fundamentally Bayesian in their approach to the assimilation of new evidence. That is, actors are viewed as having a set of policy-relevant beliefs that form priors about the world, and new pieces of information are interpreted in terms of what they suggest about the quality of these priors. Thus, presuming that a measurable set of beliefs are available, policy-oriented learning is evidenced by a change in these beliefs over time.

There are two essential conceptual attributes to learning as it is studied in the ACF: the process of learning, meaning how information is assimilated to produce learning outcomes, and the outcomes themselves. Actual examples of research on learning are heavily skewed toward the observation of learning outcomes.

Processes of Learning

The first approach emphasizes that learning is a process of belief or behavioral change. This process involves potentially complex cognitive processes of information seeking, assimilation, and adjustment as internal belief systems respond to information and behaviors are adjusted accordingly. While learning is indeed a process, it is rarely studied in this way empirically. This is because the study of learning as a process requires research designs that are difficult to implement or infeasible in many study contexts, such as experimental design that requires controlled research settings and participants drawn from relatively inaccessible populations (policy actors).

Outcomes of Learning

The second essential approach is to study learning through the observation of outcomes that are expected when learning takes place. This is by far the most common approach to the study of learning. This approach will treat the learning process as a "black box," where the processes may be assumed or ignored altogether. In this approach, observations of learning amount to changes in learning outcomes over time, and ideally in response to some stimulus. These studies might observe a process of belief change, for example, and might assume that learning occurs. In the following, we outline four essential approaches that have been used to measure learning outcomes.

Self-reported learning. These studies approach learning through self-reported statements of policy actors, usually in surveys or through interviews. This usually occurs through interviews and surveys. For example, Leach, Weible, Vince, Siddiki, and Calanni (2014) asked policy actors whether they had any change in their knowledge from participating in a collaborative process. Similarly, Moyson et al. (2017) and Pattison (2018) explored self-reported learning in similar ways, including questions about reinforced beliefs.

Observed changes in beliefs (directly measured). These studies report on changes in beliefs, often in a survey, but do not ask a "learning" question directly. For example, Weible et al. (2009) studied changes in beliefs over time. Similarly, Henry, Dietz, and Sweeney (2020) studied learning as changes in organizational beliefs about the environmental risk that were preceded by the formation of information-sharing networks between organizations.

Observed changes in beliefs (indirectly measured). These studies focus on reported changes in beliefs by a policy actor or sets of policy actors, often through the observation of statements in social/news media or legislative hearing data. These can also be described as shifts in frames or narratives as found in analytical debates. The strengths of these approaches are the availability of the data and the longitudinal aspect of the data.

Correlation between policy change and the accumulation of information at the subsystem level. This is the most common approach to study learning using case study and/or document analysis. In this approach, learning is conceptualized at the subsystem level and is viewed as a driver of major policy change. Studies using this approach are usually set on assessing or testing ACF's hypothesis about learning as a potential pathway to policy change. When major policy change is correlated with the accumulation of information about an issue, and not attributable to alternative pathways to policy change such as internal or external shocks, then the change is attributed to policy-oriented learning. Learning of this sort was identified by Cairney (2007) in UK tobacco policy, Kingiri (2011) in Kenyan biotech policy, and Weber, Driessen, Schueler, and Runhaar (2013) in Dutch noise policy.

Research Foci and the Challenges of Learning Research in ACF

ACF research on learning is often applied to the study of within- versus cross-coalition learning. The distinction between these types of learning is whether the stimuli to learning support the beliefs of a member of a competing coalition or the belief of one's coalition. Thus, whether a particular instance of belief change would count as within- or cross-coalition learning depends on the belief in question and the priors of the learning agent.

This sometimes-vague distinction underscores the need for clear conceptualization of learning and the different components of the learning process. Learning is a pathway to belief change, and there exists multiple pathways that may explain belief change within the ACF. For instance, learning may depend upon interpretation of raw data and experience from one's environment (referred to as individual learning), or learning may be driven by the adoption of the beliefs or behaviors of one's social contacts (referred to as social learning). Moreover, the adoption of new beliefs is likely influenced by the content of these beliefs relative to one's priors, as well as the relevance of these beliefs to ongoing policy debates.

Developing a clear conceptualization of learning within the ACF, and the conditions under which learning occurs, is a persistent challenge. While

most ACF learning research studies belief change as a proxy for learning, future work should pursue a better understanding of the process of learning. This will include, among other phenomena, an understanding of how policy actors seek out certain information sources and use this information to update beliefs following a process of biased assimilation. Biased assimilation (the interpretation of evidence in a way that supports one's priors) has long been assumed to be a driver of belief stability; however, it is likely that the effects of biased assimilation will vary depending on policy contexts, empirical versus normative content of beliefs, as well as the position of the learned belief in the three-tier hierarchy. While survey and archival work may support research on the process of learning, future work should also consider the use of experimental methods to further study these processes.

Conclusion

This chapter describes how the ACF is applied to the study of coalitions, learning, and policy change. We partitioned the framework into its theoretical foci and then described its lexicon and research strategies. Of course, the ACF's theoretical foci overlap and are often studied in tandem. For example, we usually need to know something about coalitions to understand policy change. Therefore, we do not offer strategies as concrete procedures that must be followed all the time in a literal way. Instead, we provide certain model strategies and stress the importance of understanding the spirit of the core theoretical ideas and the logic that has led to certain common empirical strategies in ACF research. Future work should maintain a connection with the core ideas, but may follow new strategies for research based on the researchers' best judgment as well as the characteristics of the cases to be studied.

We approach the empirical application of the ACF with humility. We offer ideas for applying the framework, based on questions frequently asked of us, as well as observations of extant ACF scholarship. Many unanswered questions and unachieved objectives remain. Progress requires continued dialogue and exchange among researchers. We therefore conclude this chapter with a set of key methodological challenges to advance the ACF through the collective efforts of policy scholars.

1   **Develop a common interview protocol.** There are many ways to collect data in ACF research, however one of the most common approach is through the use of interviews and surveys. Thus, a key need is a generic interview protocol suitable for analyzing coalitions, learning, and policy change phenomena. Such an interview protocol should be generic for its portability but also adaptable enough to fit various contexts and research objectives. Application of a common protocol would provide a compilation of best practices beneficial to each researcher, but would also provide an opportunity to conduct cross-case, structured and focused comparisons (cf. George & Bennett, 2005).

2 **Continue to develop methods for analyzing text**. As found across the sciences, textual analysis has become a major form of analyzing human thought and behavior. For the ACF, a common approach has been deployed for analyzing text, especially with news media (e.g., Heikkila et al., 2019) and in using Discourse Network Analysis (e.g., Leifeld, 2013). These techniques also follow a lasting tradition in the ACF to measure belief systems through textual analysis, particularly of legislative hearings and testimonies (Sabatier & Jenkins-Smith, 1993, appendix; Zafonte & Sabatier, 2004; Carlson et al., 2019).

3 **Launch small- to moderate-sized teams for comparative analyses.** The ACF tendency for single-case study research requires researchers to know their context, which inhibits a large number of cases conducted comparatively. This can be overcome through launching small- to moderate-sized teams asking similar questions on similar topics in different locales. For example, the edited volume of Weible et al. (2016) showcases eight case studies for eight countries all focusing on the nature of coalitions in the topical area of unconventional oil and gas development (i.e., "fracking"). In doing so, they were able to understand better how country-level characteristics condition politics in this area along with the composition and behaviors of coalitions.

4 **Continue to work on modeling and statistical analysis as well as data generation and sharing.** We still have a long way to go in terms of standardizing appropriate statistical techniques for different types of data. A synthesis of available techniques and a review of techniques used in the past is needed, and at the same time we must be mindful that looking to the past may artificially constrain our choices for the future. We need to continue to build on new approaches to data generation. The advent of machine learning methods and coding approaches has created new opportunities for ACF-based research. This also lends itself to making ACF datasets publicly available.

5 **Consider partial applications and plausibility probes.** Scholars and students might find the ACF too complex to handle, given its conceptual richness and theoretical breadth. The aspiration to study the policy process while simultaneously considering coalitions, learning, and policy change adds to the challenge. However, important advances of the ACF also depend on partial applications of particular aspects of the framework, for instance, by testing one or a few hypotheses. Provisional assessments of cases based on an initial sampling of accessible data are also needed to advance the framework in important ways. Such contributions are needed to enable insight into new cases and contexts and to guide more ambitious data collection efforts.

We will conclude with a comment on methodological pluralism. There is a broad array of useful and acceptable approaches for applying the ACF. This chapter provides an overview of these approaches, with the hope that interested researchers will take the ideas and citations mentioned in this

chapter and explore and build on them as they conduct their research. However, to continue to build the ACF research program, our methods must be as transparent as possible, our instruments (e.g., interview protocols and surveys) must be made public, and our overall approach must meet a strong standard of intersubjective reliability. Only by meeting these expectations can we continue to learn from our mistakes and successes.

Notes

1 The geographic and institutional "space" that a policy subsystem occupies is a subset of the broader macro polity of the governing jurisdiction (e.g., country-level policy venues, such as those involving the exercise of legislative, executive, and judiciary authority). Policy subsystems are often associated with the administrative or implementation structures in regulatory processes or with the institutions charged with public service delivery (Redford, 1969), though the critical feature for subsystems is that these institutional structures have the capacity for authoritative policymaking (e.g., in rulemaking processes).

2 Jenkins-Smith et al. (2018) describe at length the properties of policy subsystems. This includes (i) the complex interactions involving everything from their biophysical or institutional conditions to the culture and belief systems of the policy actors therein; (ii) the demarcation of policy actors involved from those individuals not involved; (iii) the tendency for policy subsystems to overlap and nest with other policy subsystems; (iv) authority possibly found in administration, policy venue, or both; and (v) a tendency for both change and stasis.

3 In his 1998 publication, Sabatier (1998) presents a table with exemplary but general deep core belief dimensions and highlights how they can be translated into concrete policy subsystems to identify policy core and secondary beliefs, respectively. Very often, this approach is then triangulated with country-specific institutions (e.g., centralized versus decentralized policymaking) or subsystem-specific ideological concepts (e.g., sustainability; self-sufficiency; see, for example, Fischer, 2015b; Markard, Suter, & Ingold, 2016) that also enter the belief systems that policy actors can potentially "activate."

4 Finally, scholars applying the ACF should be aware of certain persistent challenges in belief measurement. Here we highlight two challenges:

   1 Overcoming path dependency in belief measurement. Many belief items have been replicated by ACF scholars, over time and across many different policy issues. While this allows for much-needed comparative analysis, the reason for replication of items is often because we lack a solid ex ante knowledge of the relevant beliefs within a subsystem. Replicating items previously used allows for an entry point in research design but also creates a path dependency problem where assumptions about belief salience propagate throughout the research program. Coupling inductive and deductive approaches to belief measurement can help to overcome this challenge.

   2 The causal relationships between individual beliefs within a belief system are important but rarely studied. Some beliefs likely constrain and shape other beliefs; however, most methods of beliefs analysis treat beliefs as discrete and independent but co-varying units. Methods of structural equation modeling or neural networks are promising approaches for understanding the full complexity of belief systems.

Rather than viewing these challenges as barriers to research, we should view these challenges as opportunities for methodological innovation.

5 Public policies are also comprised of "institutional rules" (see ACF's flow diagram in Jenkins-Smith et al., 2018) that refer to policy content or designs. These institutional rules structure implementation that then results in impacts or feedback, shaping the overall structure of the policy subsystem. Studying these institutional rules might illuminate the politics of the public policy's formation or the detailed structure of the public policy's implementation.

6 Jenkins-Smith, Silva, Gupta, and Ripberger (2014, p. 486), for example, define "Policies – seen as the aggregate sets of rules, incentives, sanctions, subsidies, taxes, and other instruments – are measured against belief systems."

7 Of course, the type of analysis conducted and the specific analytical methods utilized for collecting and analyzing data depend heavily on the objectives. For example, if we wanted to explore the correspondence of policy actors' belief systems and the content of public policy, we would have to measure those beliefs and analyze the text of the public policy and then assess their overlap and consistency. As is typically the case, analyzing policy change usually requires taking a long-term time perspective, over a decade or more. Thus, researchers recommend exploring the historical developments preceding the policy change over extended periods.

References

Aamodt, Solveig, and Iselin Stensdal. 2017. "Seizing Policy Windows: Policy Influence of Climate Advocacy Coalitions in Brazil, China, and India, 2000–2015." *Global Environmental Change* 46: 114–125.

Afonso, Alexandre. 2014. "How to Please Voters without Alienating Friends? Parties, Organised Interests and Advocacy Coalitions in Swiss Immigration Policy." *Comparative European Politics* 12(6): 568–583.

Ansell, Chris, Sarah Reckhow, and Andrew Kelly. 2009. "How to Reform a Reform Coalition: Outreach, Agenda Expansion, and Brokerage in Urban School Reform." *Policy Studies Journal* 37(4): 717–743.

Babon, Andrea, Daniel McIntyre, Gae Y. Gowae, Caleb Gallemore, Rachel Carmenta, Monica Di Gregorio, and Maria Brockhaus. 2014. "Advocacy Coalitions, REDD+, and Forest Governance in Papua New Guinea: How Likely is Transformational Change?" *Ecology and Society* 19(3): 16.

Barnes, Clare, Frank Van Laerhoven, and Peter P.J. Driessen. 2016. "Advocating for Change? How a Civil Society-Led Coalition Influences the Implementation of the Forest Rights Act in India." *World Development* 84: 162–175.

Baumgartner, Frank R., and Bryan D. Jones. 2010. *Agendas and Instability in American Politics*. Chicago, IL: University of Chicago Press.

Berardo, Ramiro. 2009. "Generalized Trust in Multi-organizational Policy Arenas: Studying Its Emergence from a Network Perspective." *Political Research Quarterly* 62(1): 178–189.

Brandenberger, Laurence, Karin Ingold, Manuel Fischer, Isabelle Schläpfer, and Philip Leifeld. 2020. "Boundary Spanning Through Engagement of Policy Actors in Multiple Issues." *Policy Studies Journal*. Online first.

Cairney, Paul. 2007. "A 'Multiple Lenses' Approach to Policy Change: The Case of Tobacco Policy in the UK." *British Politics* 2(1): 45–68.

Cairney, Paul, Karin Ingold, and Manuel Fischer. 2018. "Fracking in the UK and Switzerland: Why Differences in Policymaking Systems Don't Always Produce Different Outputs and Outcomes." *Policy & Politics* 46(1): 125–147.

Calanni, John C., Saba Siddiki, Christopher M. Weible, and William D. Leach. 2015. "Explaining Coordination in Collaborative Partnerships and Clarifying the Scope of the Belief Homophily Hypothesis." *Journal of Public Administration Research and Theory* 25(3): 901–927.

Calmon, Paulo, and Suely Mara Vaz Guimarães de Araújo. 2021. "Coalitions and Ideas in Conflict: Public Policies on the Path to Sustainable Development." *Revista de Administração Pùblica*. Forthcoming.

Carlson, Deven, Joseph Ripberger, Wesley Wehde, Hank Jenkins-Smith, Carol Silva, Kuhika Gupta, Benjamin Jones, and Robert Berrens. 2019. "Benefit–Cost Analysis, Policy Impacts, and Congressional Hearings." *Journal of Benefit-Cost Analysis* 10(1): 65–94.

Collier, David. 2011. "Understanding Process Tracing." *Political Science & Politics* 44(4): 823–830. doi:10.1017/s1049096511001429.de Medeiros, Estela Alves, and Ricardo Corrêa Gomes. 2019. "Advocacy Coalitions and Negotiation Strategies in the Revision of the Forest Code." *Revista de Administração Pùblica* 53(1). https://doi.org/10.1590/0034-7612173987

Dunlop, Claire A. and Claudio M. Radaelli. 2017. "Learning in the Bath-Tub: The Micro and Macro Dimensions of the Causal Relationship between Learning and Policy Change." *Policy and Society* 36(2): 304–319.

Elgin, Dallas J. 2015. "Utilizing Hyperlink Network Analysis to Examine Climate Change Supporters and Opponents." *Review of Policy Research* 32(2): 226–245.

Elgin, Dallas J. and Christopher M. Weible. 2013. "A Stakeholder Analysis of Colorado Climate and Energy Issues Using Policy Analytical Capacity and the Advocacy Coalition Framework." *Review of Policy Research* 30(1): 114–133.Fenger, Menno, and Pieter-Jan Klok. 2001. "Interdependency, Beliefs, and Coalition Behavior: A Contribution to the Advocacy Coalition Framework." *Policy Sciences* 34(2): 157–170.

Fischer, Manuel. 2015a. "Institutions and Coalitions in Policy Processes: A Cross-Sectoral Comparison." *Journal of Public Policy* 35(2): 245–268.

Fischer, Manuel. 2015b. "Collaboration Patterns, External Shocks and Uncertainty: Swiss Nuclear Energy Politics before and after Fukushima." *Energy Policy* 86: 520–528.

Fischer, Manuel, Karin Ingold, Pascal Sciarini, and Frédéric Varone. 2016. "Dealing with Bad Guys: Actor- and Process-Level Determinants of the 'Devil Shift' in Policy Making." *Journal of Public Policy* 26(2): 309–334.

Fischer, Manuel, and Martino Maggetti. 2017. "Qualitative Comparative Analysis and the Study of Policy Processes." *Journal of Comparative Policy Analysis*: Research and Practice 19(4): 345–361.

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

Gralepois, Mathilde, Corinne Larrue, Mark Wiering, Ann Crabbé, Sue Tapsell, Hannelore Mees, Kristina Ek, and Malgorzata Szwed. 2016. "Is Flood Defense Changing in Nature? Shifts in the Flood Defense Strategy in Six European Countries." *Ecology and Society* 21(4): 37.

Hall, Peter A. 1993. "Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain." *Comparative Politics* 25(3): 275–296.

Harrinkari, Teemu, Pia Katila, and Heimo Karppinen. 2016. "Stakeholder Coalitions in Forest Politics: Revision of Finnish Forest Act." *Forest Policy and Economics* 67: 30–37.

Heikkila, Tanya, Ramiro Berardo, Christopher M. Weible, and Hongtao Yi. 2019. "A Comparative View of Advocacy Coalitions: Exploring Shale Development Politics in the United States, Argentina, and China." *Journal of Comparative Policy Analysis: Research and Practice* 21(2): 151–166.

Heinmiller, B. Timothy (2016). *Water Policy Reform in Southern Alberta: An Advocacy Coalition Approach*. Toronto, CA: University of Toronto Press.

Henry, Adam Douglas. 2011. "Ideology, Power, and the Structure of Policy Networks." *Policy Studies Journal* 39(3): 361–383.

Henry, Adam Douglas, Thomas Dietz, and Robin L. Sweeney. 2020. "Coevolution of Networks and Beliefs in U.S. Environmental Risk Policy." *Policy Studies Journal*. Early view psj.12407. https://doi.org/10.1111/psj.12407

Henry, Adam Douglas, Mark Lubell, and Michael McCoy. 2012. "Survey-Based Measurement of Public Management and Policy Networks." *Journal of Policy Analysis and Management* 31(2): 432–452.

Ingold, Karin. 2011. "Network Structures within Policy Processes: Coalitions, Power, and Brokerage in Swiss Climate Policy." *Policy Studies Journal* 39(3): 435–459.Ingold, Karin, Manuel Fischer, and Paul Cairney. 2017. "Drivers for Policy Agreement in Nascent Subsystems: An Application of the Advocacy Coalition Framework to Fracking Policy in Switzerland and the UK." *Policy Studies Journal* 45(3): 442–463.

Ingold, Karin and Muriel Gschwend. 2014. "Science in Policy-Making: Neutral Experts or Strategic Policy-Makers?" *West European Politics* 37(5): 993–1018.

Jenkins-Smith, Hank C. 1990. *Democratic Politics and Policy Analysis*. Pacific Grove, CA: Brooks/Cole Publishing.

Jenkins-Smith, Hank C., Daniel Nohrstedt, Christopher M. Weible, and Karin Ingold. 2018. "The Advocacy Coalition Framework: An Overview of the Research Program." In Christopher M. Weible and Paul Sabatier (Eds.) *Theories of the Policy Process*. 4th edition. London: Routledge (pp. 135–171).Jenkins-Smith, Hank, Carol L. Silva, Kuhika Gupta, and Joseph T. Ripberger. 2014. "Belief System Continuity and Change in Policy Advocacy Coalitions: Using Cultural Theory to Specify Belief Systems, Coalitions, and Sources of Change." *Policy Studies Journal* 42(4): 484–508.

Jenkins-Smith, Hank C., Gilbert K. St. Clair, and Brian Woods. 1991. "Explaining Change in Policy Subsystems: Analysis of Coalition Stability and Defection over Time." *American Journal of Political Science* 35(4): 851–880.

Khayatzadeh-Mahani, Akram, Eric Breton, Arne Ruckert, and Ronald Labonté. 2017. "Banning Shisha Smoking in Public Places in Iran: An Advocacy Coalition Framework Perspective on Policy Process and Change." *Health Policy and Planning* 32(6): 835–846.

Kingdon, John W. 1984. *Agendas, Alternatives, and Public Policies*. Boston, MA: Little Brown. doi:10.2307/3323801.

Kingiri, Ann Njoki. 2011. "Conflicting Advocacy Coalitions in an Evolving Modern Biotechnology Regulatory Subsystem: Policy Learning and Influencing Kenya's Regulatory Policy Process." *Science and Public Policy* 38(3): 199–211.

Knoke, David, Franz Urban Pappi, Jeffrey Broadbent, and Yutaka Tsujinaka. 1996. *Comparing Policy Networks*. Cambridge: Cambridge University Press.

Koebele, Elizabeth A. 2020. "Cross-Coalition Coordination in Collaborative Environmental Governance Processes." *Policy Studies Journal* 48(3): 727–753.

Koebele, Elizabeth A., Stephanie Bultema, and Christopher M. Weible 2020. "Modeling Environmental Governance in the Lake Tahoe Basin: A Multiplex Network Approach." In Manuel Fischer and Karin Ingold (Eds.) *Networks in Water Governance.* Cham, Switzerland: Palgrave Macmillan (pp. 173–202).

Kukkonen, Anna, Tuomas Ylä-Anttila, and Jeffrey Broadbent. 2017. "Advocacy Coalitions, Beliefs and Climate Change Policy in the United States." *Public Administration* 95(3): 713–729.

Kukkonen, Anna, Tuomas Ylä-Anttila, Pradip Swarnakar, Jeffrey Broadbent, Myanna Lahsen, and Mark CJ Stoddart. 2018. "International Organizations, Advocacy Coalitions, and Domestication of Global Norms: Debates on Climate Change in Canada, the US, Brazil, and India." *Environmental Science & Policy* 81: 54–62.

Leach, William D., Christopher M. Weible, Scott R. Vince, Saba N. Siddiki, and John C. Calanni. 2014. "Fostering Learning through Collaboration: Knowledge Acquisition and Belief Change in Marine Aquaculture Partnerships." *Journal of Public Administration Research and Theory* 24(3): 591–622.

Leifeld, Philip. 2013. "Reconceptualizing Major Policy Change in the Advocacy Coalition Framework: A Discourse Network Analysis of German Pension Politics." *Policy Studies Journal* 41(3): 169–198.

Li, Wei, and Christopher M. Weible. 2019. "China's Policy Processes and the Advocacy Coalition Framework." *Policy Studies Journal.* Online first. doi:10.1111/psj.12369.

Lindblom, Charles E. 1959. "The Science of 'Muddling Through'." *Public Administration Review* 19(2): 79–88.

Markard, Jochen, Marco Suter, and Karin Ingold. 2016. "Socio-Technical Transitions and Policy Change – Advocacy Coalitions in Swiss Energy Policy." *Environmental Innovation and Societal Transitions* 18: 215–237. doi:10.1016/j.eist.2015.05.003.

McDougall, Lori. 2006. "Power and Politics in the Global Health Landscape: Beliefs, Competition and Negotiation among Global Advocacy Coalitions in the Policy-Making Process." *International Journal of Health Policy and Management* 5(5): 309–320.

Montefrio, Marvin Joseph F. 2014. "State versus Indigenous Peoples' Rights: Comparative Analysis of Stable System Parameters, Policy Constraints and the Process of Delegitimation." *Journal of Comparative Policy Analysis: Research and Practice* 16(4): 335–355.

Moyson, Stephane. 2017. "Cognition and Policy Change: The Consistency of Policy Learning in the Advocacy Coalition Framework." *Policy and Society* 36(2): 320–344.

Moyson, Stephane, Peter Scholten, and Christopher M. Weible. 2017. "Policy Learning and Policy Change: Theorizing Their Relations from Different Perspectives." *Policy and Society* 36(2): 161–177.

Nohrstedt, Daniel. 2008. "The Politics of Crisis Policymaking: Chernobyl and Swedish Nuclear Energy Policy." *Policy Studies Journal* 36(2): 257–278.

Nohrstedt, Daniel. 2011. "Shifting Resources and Venues Producing Policy Change in Contested Subsystems: A Case Study of Swedish Signals Intelligence Policy." *Policy Studies Journal* 39(3): 461–484.

Nohrstedt, Daniel and Kristin Olofsson. 2016. "Advocacy Coalition Politics and Strategies on Hydraulic Fracturing in Sweden." In Christopher M. Weible, Tanya Heikkila, Karin Ingold, and Manuel Fischer (Eds.) *Policy Debates on Hydraulic Fracturing: Comparing Coalition Politics in North America and Europe.* New York, NY: Springer (pp. 147–175).

Pattison, Andrew. 2018. "Factors Shaping Policy Learning: A Study of Policy Actors in Subnational Climate and Energy Issues." *Review of Policy Research* 35(4): 535–563.

Pierce, Jonathan J., Holly L. Peterson, and Katherine C. Hicks. 2020. "Policy Change: An Advocacy Coalition Framework Perspective." *Policy Studies Journal* 48(1): 64–86.

Pierce, Jonathan J., Holly L. Peterson, Michael D. Jones, Samantha P. Garrard, and Theresa Vu. 2017. "There and Back Again: A Tale of the Advocacy Coalition Framework." *Policy Studies Journal* 45(S1): S13–S46. Ragin, Charles C. 1987. *The Comparative Method. Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press.

Redford, Emmette S. 1969. *Democracy in the Administrative State*. New York, NY: Oxford University Press.

Rennkamp, Britta, Sebastian Haunss, Kridtiyaporn Wongsa, Araceli Ortega, and Erika Casamadrid. 2017. Competing Coalitions: The Politics of Renewable Energy and Fossil Fuels in Mexico, South Africa and Thailand." *Energy Research & Social Science* 34: 214–223.

Rihoux, Benoît, and Charles C. Ragin. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. London: SAGE Publications.

Ripberger, Joseph T., Kuhika Gupta, Carol L. Silva, and Hank C. Jenkins-Smith. 2014. "Cultural Theory and the Measurement of Deep Core Beliefs within the Advocacy Coalition Framework." *Policy Studies Journal* 42(4): 509–527.

Sabatier, Paul A. 1988. "An Advocacy Coalition Model of Policy Change and the Role of Policy-Oriented Learning Therein." *Policy Sciences* 21: 129–168.

Sabatier, Paul A. 1998. "The Advocacy Coalition Framework: Revisions and Relevance for Europe." *Journal of European Policy* 51: 98–130.

Sabatier, Paul A. and Hank C. Jenkins-Smith. 1993. *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder, CO: Westview Press.

Sabatier, Paul A., and Hank C. Jenkins-Smith. 1999. "The Advocacy Coalition Framework: An Assessment." In Paul A. Sabatier (Ed.) *Theories of the Policy Process*. Boulder, CO: Westview Press (pp. 117–166).

Schmid, Nicolas, Sebastian Sewerin, and Tobias S. Schmidt. 2020. "Explaining Advocacy Coalition Change with Policy Feedback." *Policy Studies Journal* 48(4): 1109–1134.

Smith, Adrian. 2000. "Policy Networks and Advocacy Coalitions: Explaining Policy Change and Stability in UK Industrial Pollution Policy?" *Environment and Planning C: Government and Policy* 18(1): 95–114.

Sotirov, Metodi and Georg Winkel. 2016. "Toward a Cognitive Theory of Shifting Coalitions and Policy Change: Linking the Advocacy Coalition Framework and Cultural Theory." *Policy Sciences* 49: 125–154.

Stritch, Andrew. 2015. "The Advocacy Coalition Framework and Nascent Subsystems: Trade Union Disclosure Policy in Canada." *Policy Studies Journal* 43(4): 437–455.

Valman, Matilda. 2016. "Beliefs and Behavior in International Policy Making: Explanations to Longitudinal Changes in the Governance of the Baltic Sea." *Maritime Studies* 15(12): 1–16. https://doi.org/10.1186/s40152-016-0050-4

Vieira, Diego Mota. 2020. "The Discourse and Coordination among Advocacy Coalitions: The Case of Belo Monte." *RAUSP Management Journal* 55(1): 86–99. http://dx.doi.org/10.1108/rausp-10-2018-0096

Wagner, Paul, and Tuomas Ylä-Anttila. 2018. "Who Got Their Way? Advocacy Coalitions and the Irish Climate Change Law." *Environmental Politics* 27(5): 872–891.

Weber, Miriam, Peter PJ Driessen, Ben J. Schueler, and Hens AC Runhaar. 2013. "Variation and Stability in Dutch Noise Policy: An Analysis of Dominant Advocacy Coalitions." *Journal of Environmental Planning and Management* 56(7): 953–981.

Weible, Christopher M. and Tanya Heikkila. 2016. "Comparing the Politics of Hydraulic Fracturing in New York, Colorado, and Texas." *Review of Policy Research* 33(3): 232–250.

Weible, Christopher M., Tanya Heikkila, Karin Ingold, and Manuel Fischer (Eds.). 2016 *Policy Debates on Hydraulic Fracturing: Comparing Coalition Politics in North America and Europe.* New York, NY: Springer.

Weible, Christopher M. and Karin Ingold. 2018. "Why Advocacy Coalitions Matter and Practical Insights about Them." *Policy & Politics* 46(2): 325–343.

Weible, Christopher M., Karin Ingold, Daniel Nohrstedt, Adam Douglas Henry, and Hank C. Jenkins-Smith. 2020. "Sharpening Advocacy Coalitions." *Policy Studies Journal* 48(4): 1054–1081.

Weible, Christopher M., Paul A. Sabatier, and Kelly McQueen. 2009. "Themes and Variations: Taking Stock of the Advocacy Coalition Framework." *Policy Studies Journal* 37(1): 121–140.Zafonte, Matthew and Paul A. Sabatier. 1998. "Shared Beliefs and Imposed Interdependencies as Determinants of Ally Networks in Overlapping Subsystems." *Journal of Theoretical Politics* 10(4): 473–505.

Zafonte, Matthew and Paul A. Sabatier. 2004. "Short-Term versus Long-Term Coalitions in the Policy Process: Automotive Pollution Control, 1963–1989." *Policy Studies Journal* 32(1): 75–107.

Table 5A.1 Categories of policy actors and coalitions and their definitions

		Definitions
Policy actor categories	Latent/isolate/ potential policy actors	Policy actors not engaged who might have stake in the issue.
	Principal and auxiliary actors	Principal actors are regularly engaged central to the subsystem; auxiliary policy actors are intermittently engaged, operate on the periphery of the subsystem, or both.
	Purposive/material groups	Purposive groups are organizations motivated by a value-based mission (e.g., a nonprofit environmental group); material groups are organizations motivated by profit (e.g., a corporation or business).
	Policy brokers	A policy actor motivated to help opponents reach compromise or agreements.
	Policy entrepreneur	A policy actor motivated in shaping policy change or stasis decisively.

Table 5A.1 Cont.

		Definitions
Advocacy coalition categories	Adversarial coalitions	Coalitions that compete over beliefs and worldviews, with low cross–coalition coordination and high across–coalition conflict.
	Cooperative coalitions	Coalitions that still compete over beliefs and worldviews, but that engage in substantive cross–coalition coordination, with low across–coalition conflict.
	Dominant/minority coalition	One coalition is dominant if it impacts policy (change or stasis) considerably, and this normally over time. Minority coalitions are less successful in their translation of beliefs into policy. Dominant/minority often (but not always!) also correlates with size, resources, and strategy.
	Ephemeral or coalition of convenience	Trust and former contacts shape the observable coordination patterns among actors more than beliefs (Berardo, 2009). Such "coalitions of convenience" (Stritch, 2015; Cairney Ingold, & Fischer, 2018) were most often found in "nascent subsystem" where clear-cut structures of beliefs and coordination have yet to form, but actors still see an advantage in joining forces.
	Potential coalition	Potential coalitions are the ones, where actors have some ideological agreement but did not yet engage in subsystem activities, or there is just the empirical proof of it lacking.

Source: Weible et al. (2020).

Table 5A.2a Methods of data gathering (one key example of recent or relevant illustration)

Method	Advocacy Coalitions			Learning	Policy Change
	Shared Beliefs	*Coordination*	*Resources and Strategies*		
News and social media	Kukkonen et al. (2017) used Discourse Network Analysis of the news media to document climate-related beliefs.	Koebele et al. (2020) analyzed coordination through news media.			
Surveys	Ripberger, Gupta, Silva, and Jenkins-Smith (2014) used surveys to measure deep core beliefs via Cultural Theory.	Henry (2011) used surveys to identify collaborative ties in a regional planning subsystem.	Elgin and Weible (2013) used surveys to measure the political resources in a climate and energy subsystem.	Leach et al. (2014) analyzed learning as gains in knowledge in marine aquaculture subsystems.	
Interviews	Sotirov and Winkel (2016) conducted 73 interviews in forest policy subsystems in Germany and Bulgaria to measure beliefs via Cultural Theory.	Ingold (2011) interviewed 54 senior representatives of 34 organizations in Swiss climate policy to gather data about their coordination strategies.		Moyson et al. (2017) studying self-reported learning.	Strong for process tracing techniques (e.g., Heinmiller, 2016).

| Document analysis (e.g., public policies/ government and nongovernment reports) | Besides media analysis, Brandenberger et al. (2020) link actors to issues and policy preferences based on parliamentary minutes in Swiss water policy. | To identify coalition coordination in the governance of the Baltic Sea, Valman (2016) codes official documents and deduces coordination from reservation patterns of state actors in the Helsinki Commission. | Vieira (2020) codes congressional public hearings and deduces strategies of coalition members in the Belo Monte case. | Gralepois et al. (2016) discuss about changes in Dutch flood defense (combination with focus group research and interviews). |

Note: Blank fields do not mean that these combinations are not possible or do not exist, but that there is any example at the authors knowledge.

Table 5A.2b Methods of data analysis (one key example of recent or relevant illustration)

Coalitions	Belief Clustering	Beliefs as Relations	Weak/Strong Coordination	Presence/Absence of Coordination
	Schmid et al. 2020 use Discourse Network Analysis to investigate modularity analysis	Ingold et al. (2017) compare fracking politics in Switzerland and the UK and operationalize advocacy coalitions based on ally and enemy relations and blockmodeling techniques.	Zafonte and Sabatier (1998) analyzed weak and strong coordination using ally network data based on survey results.	Wagner and Ylä-Anttila (2018) investigate advocacy coalitions around the Irish Climate Change Law and identify coordination (before belief) clusters based on survey data.
Policy Change	**Explaining Different Degrees of Policy Change** Fischer (2015a) compares different subsystems and types of policy change applying QCA.	**Minor or No Change** Nohrstedt (2008) linking Chernobyl to Swedish energy policy and documenting pathways to minor and no change.	**Causality – Drivers for Policy Change** Sotirov and Winkel (2016) exploring policy change in German and Bulgarian forest policy.	
Learning	**Investigating Process of Learning** Moyson (2017) studying the process of learning over time related to Belgian network industries and their liberalization.	**Learning and Policy Change** Kingiri (2011) studying biotech policy in Kenya.	**Lack of Learning** Khayatzadeh-Mahani et al. (2017) studying shisha ban attempts in Iran.	

Table 5A.3 Advantages and disadvantages of different data sources

	Advantage	*Shortcomings*
News and social media	Analysis over time; data accessibility	Media ownership; pays attention to salient issues and most active/interesting actors only. Only some subsystem actors (particular type) are active on Twitter, FB, and so on. Difficult to identify core beliefs and stable patterns as well as resources; often indirect coordination through hyperlink analysis or so. Combination of inductive and deductive approach necessary.
Surveys	Complete data	Social desirability; data availability; response rate issues; snapshot only.
(Expert) interviews	Good for actor identification or data validation	Difficult to get data on complete subsystem actors, their beliefs, etc. Quality and completeness of data very much dependent on selection and availability of key informants.
Document analysis (e.g., public policies/ government and nongovernment reports)	Systematic data gathering strategy, but only during the "active" policy negotiation (more difficult for policy implementation for example)	Difficult to get complete data; problem on data accessibility; often only indirect data on coordination (through joint venue participation). Better documentation during phases of agenda-setting, policy formulation; less so during implementation and evaluation.

Table 5A.4 Concepts, attributes, and measurements

Reference Category	Foundational Theoretical Concept	Attributes	Measurement Alternatives	Generally Measured/Identified Using …
Cross–cutting concepts	**Policy subsystem**	**Topical scope**	*Historical narrative of an issue*	Expert interviews, document, or media analysis
			Media or public attention of an issue	Document or media analysis
			Linking an issue to its policy process	Expert interviews, document analysis
		Geographic scope	*Decisional level; jurisdiction; catchment area; community*	Expert interviews, document analysis
		Policy actors	*See below*	
	Policy actors	**Professional engagement**	*Decisional and positional approaches Network analysis*	Document (or media) analysis
		Reputation and influence	*Reputational approach*	Expert interviews, surveys (name generator)
	Belief systems	**Scope of belief (deep core, policy core, secondary beliefs)**	*Belief measurement through unsolicited written or verbal statements*	Document or media analysis (e.g., Discourse Network Analysis)
			Belief measurement through solicited self-reports	Surveys, interviews
Three meta-theoretical key concepts	**Coalitions**	**Shared beliefs**	*Belief clustering*	Data gathering: survey, interviews, media or document analysis; data analysis: clustering techniques

		Beliefs as relations	Data gathering: survey, interviews, document or media analysis; data transformation into network matrices (belief similarity or distance); data analysis: clustering techniques, blockmodeling, network models such as Exponential Random Graph Models (ERGM) or Stochastic Actor Oriented Models (SAOM)
	Coordination	Weak or strong coordination interaction	Survey, interviews, less document or media analysis
		Presence or absence of coordination interaction	Survey, interviews, less document or media analysis
	Coalition resources	Money, personnel, members, allies, important contacts, information and knowledge, leadership	Document analysis (typically annual or financial reports), interviews, organizational websites
	Coalition strategies	Coordination (see above), lobbying, co-signing documents, public attention, media presence	Document analysis (typically annual or financial reports), interviews, organizational websites
Policy change	**Type of public policy and policy venue**	Collection of policy documents	Document analysis
		Study of actors' beliefs and potential translation into policy	Observational methods; document analysis; key informant interviews

Table 5A.4 Cont.

Reference Category	Foundational Theoretical Concept	Attributes	Measurement Alternatives	Generally Measured/Identified Using …
		Nature and Magnitude of Policy change (e.g., major vs. minor)	*Process tracing*	Document analysis; expert interviews
			Qualitative comparative analysis	Document analysis; interviews; expert interviews (mainly for validation)
	Learning	**Process of learning**	*Causal effect of information stimuli on belief change*	Surveys; experimental research
		Outcome of learning	*Self-reported learning*	Surveys; interviews
			Observed changes in beliefs (directly and indirectly measured)	Surveys; hearings; media data
			Correlation between change and information	Case study and document analysis

6 Conducting Narrative Policy Framework Research

From Theory to Methods

*Michael D. Jones, Mark K. McBeth,
Elizabeth A. Shanahan, Aaron Smith-Walter,
and Geoboo Song*

Introduction

The Narrative Policy Framework (NPF) is a comprehensive approach to studying public policy that aspires to capture the policy process at multiple levels of analysis, across time, and within any context. As may have already been intimated from the name, the NPF assumes narratives are likely important to such an endeavor. As such, the NPF's central goal is to assess the role of narratives within the policy process. Because the NPF's approach to public policy is comprehensive, it encompasses a vast number of potential facets and concepts within the policy process that can be coupled with an equally vast number of research design permutations. Therefore, the scope of the NPF is admittedly large and implementing it can be daunting for those just coming to the framework. There have been several publications over the years expounding upon the theory and concepts of the NPF (Jones, Shanahan, and McBeth, 2014; McBeth, 2014; Shanahan, Jones, McBeth, and Radaelli, 2018), potential normative (e.g., Jones and McBeth, 2020; Sievers and Jones, 2020) or communication applications (Crow and Jones, 2018; Jones and Crow 2017), and nonscience-based applications (e.g., Gray and Jones, 2016), among other expositions. While we do, of course, address theory and concepts in this chapter, we advise readers interested in nuanced aspects of NPF theory to consult relevant publications. Here, however, our goal is to expand upon previous efforts (Shanahan, Jones, and McBeth, 2018) to guide readers in the nuts and bolts of actually constructing a science-based NPF study.

This chapter proceeds by first asking the reader to consider if the NPF is the right approach for them by assessing their research goals in terms of the NPF's theoretical assumptions. We then work through developing NPF research questions and hypotheses, determining levels of analysis, and concept operationalization. Subsequent sections address aspects of research design, including type, appropriate methodologies, and data, as well as various analytic techniques employed by NPF researchers. We close with a brief discussion of our take on the future of NPF research. Figure 6.1 illustrates the path our guide takes. We recognize that while our approach here is linear,
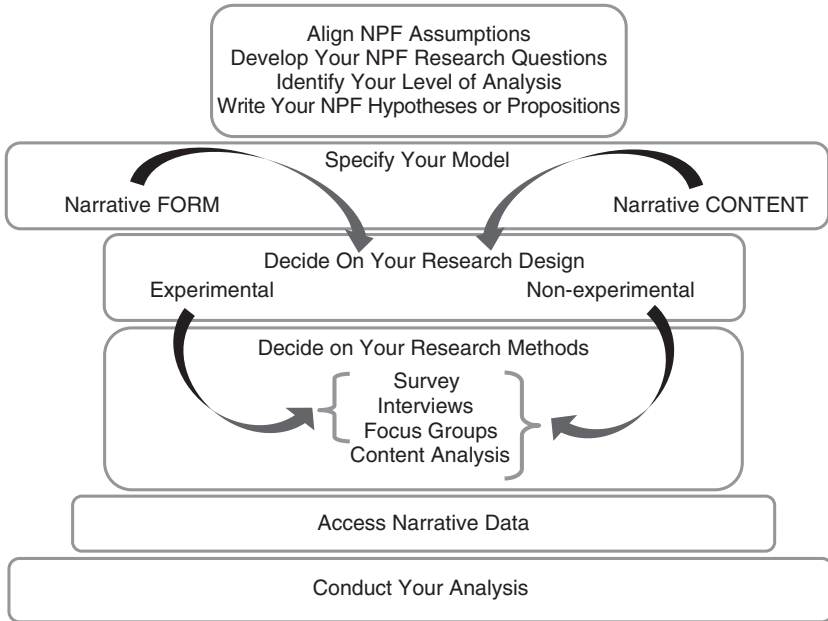
Figure 6.1 How to conduct NPF research

that the path of research design is frequently not. So, please feel free to flow from one category to another as your process dictates.

To provide the most current or seminal works related to conducting an NPF study, existing NPF research was collected and consulted. Ninety-two studies were identified by the authors between the years 2004 and 2019. As the NPF was named in 2010 (Jones and McBeth, 2010), most of the NPF applications referenced in this chapter were published after said naming. However, since the NPF's beginnings clearly precede 2010 (see McBeth, 2014), several select studies from the mid-2000s are also referenced (e.g., McBeth, Shanahan, and Jones, 2005; McBeth, Shanahan, Arnell, and Hathaway, 2007; Shanahan, McBeth, Hathaway, and Arnell, 2008). As a potential aid to researchers, these articles are categorized by NPF concepts and included as appendices, along with a meso-level NPF research design flow chart example (Appendices A and B).[1]

Deciding If the NPF Is the Right Framework for Your Research

The NPF situates narrative as the centerpiece for understanding the policy process. However, narrative has historically been an elusive concept within public policy and exactly how narratives operate within the policy process has been variously justified based on different ontological and epistemological

assumptions about how the world works. As such, considering the alignment of the five NPF assumptions detailed below with your research assumptions will assist you in deciding whether the NPF is the appropriate theoretical lens (see Jones, 2018, p. 727):

1   *Social Construction of Policy Realities:* The NPF accepts that there is a world independent of human perception. However, for public policy what is important about the world is more about what people perceive it to be and what that means to them, rather than determining what actually "is." Thus, the NPF assumes that policy realities are best understood through people's collective and individual social constructions.
2   *Bounded Relativity:* The meanings that people ascribe to various objects or processes related to public policy will vary considerably, but that variation is not boundless, nor is it random. Rather, individuals seek meanings from existing possibilities derived from their already systematized ways of understanding the world, such as their identity or culture. These systems create interpretive boundaries within which a limited number of possibilities exist.
3   *Narratives Have Generalizable Components:* The NPF takes a structuralist approach to narrative, which means that narratives are objects in the world that have specific and identifiable features (e.g., narrative elements, narrative strategies) that can be counted, allowing mathematical and statistical operations to be performed.
4   *Three Levels of Analysis:* The NPF assumes narratives scale and thus can be examined at three interacting levels: the micro (individual), meso (group), and macro (cultural and institutional).
5   *Homo narrans:* An amalgamation of scientific findings and theorization across multiple academic fields (see Shanahan, Jones, McBeth, and Radaelli, 2018, pp. 181–183), the NPF assumes that emotion precedes reason and that it is the affect-imbued stories people tell each other and themselves that drive cognition, communication, and decision-making.[2]

The most common misstep in research that invokes the NPF but fails to effectively implement the framework is a failure to calibrate the assumptions of the NPF to the goals of the research and the orientations of researchers. Simply put, "the NPF is not a one-size fits all for research projects centered on narrative" (Shanahan, Jones, and McBeth, 2018, p. 333). During the early stages of research design, carefully consider each assumption to determine if your project conforms to NPF assumptions. If not, it is probably not an NPF study. For example, your approach to research might assume reality is socially constructed and that narrative is central to understanding that reality. But if you also assume narrative is completely contextual and unique (i.e., *sui generis* and non-generalizable), the NPF is not for you. In this case, you would meet assumptions #1 and #5 but clearly violate assumptions #2 and #3 (and probably #4). Examining the fit of NPF assumptions with your research is as much about you as the researcher as it is about your research design.

Crafting NPF Research Question(s)

All good research is guided by one or more research questions. An NPF study is no exception. However, what distinguishes an NPF study is the focus on the role of narrative in the policy process. NPF research questions typically fall in one of two broad categories: questions that are driven by a concern for a specific policy issue and questions that are driven by a primary want to advance theory. The usual way to tell the difference is that with the latter, the policy area is a means to test the theory and with the former, the theory is a means to test something relevant to the policy area. In both cases, the theoretical contribution of the study to the NPF will need to be addressed, although the priorities of such research clearly differ. Table 6.1 provides several generic research questions at different levels of analysis (discussed more in the next section).

Determining Level of Analysis: Micro, Meso, or Macro

One of the first issues that an NPF researcher will need to consider—usually concurrently with assumption alignment and research question development—is determining the level of analysis. Whereas the *unit* of

Table 6.1 Level of analysis in NPF research questions

Level of Analysis	Generic Research Question	Examples
Micro	What influence do narratives have on individual preferences and cognitions? What influence do narratives have on individual decision-making?	McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; Shanahan et al., 2019; Zanocco, Song, and Jones, 2018 Gray and Jones, 2016; Guenter and Shanahan, 2020; McMorris, Zanocco, and Jones, 2018
Meso	How do groups construct policy narratives? Do policy narratives shape policy outcomes? How do groups use images in their policy narratives? What variations in policy narratives differences are associated with partisan control of government?	Merry, 2020; Schwartz, 2019 Dupuis, 2018; McMorris et al., 2018; Laufer and Jones, 2021 Boscarino, 2020; McBeth et al., 2012 Chang and Koebele, 2020; Peterson, 2018
Macro	What are the conditions under which macro-level narratives develop and change?	Boscarino, 2020; Ney, 2014; Peterson, 2018; Veselková and Beblavy, 2014

Note: Table updated and based on Shanahan et al. (2018, p. 334).

analysis refers to the specific units from which you gather your observational data (e.g., coded narrative statements such as interest group tweets), the *level* of analysis sets the scope of your research.

As described in Shanahan, Jones, and McBeth (2018, p. 334), "each level of analysis in the NPF provides a vantage point for examining the role of narratives in the policy process." Being clear with your level of analysis means that you are clear about the population of your study: individuals (micro), groups (meso), or institutions and/or culture (macro).

- A rule of thumb is that if you are studying how narratives influence individuals' beliefs, cognitions, preferences, and decisions (e.g., Jones, 2014a) or how individuals produce narratives (Colville, 2019), you are operating at the micro-level of analysis. These types of studies have tended to rely on approaches such as experimental designs (e.g., Guenther and Shanahan et al., 2020) and surveys (e.g., Knackmuhs, Farmer, and Knapp, 2020; Laufer and Jones, 2021).
- Meso-level studies tend to focus on how groups strategically produce narrative and to what extent these policy narratives shape policy outcomes. At this level of analysis, a common approach is to conduct content analysis of social artifacts such as public consumption documents (e.g., Smith–Walter et al., 2016), parliamentary or congressional testimonies (e.g., O'Leary, Borland, Stockwell, and MacDonald, 2017), media accounts (e.g., Gupta, Ripberger, and Wehde, 2018), social media tweets (Merry, 2016), or similar narratives made available by groups and advocacy coalitions.
- Finally, if you study how grand policy narratives develop and change across institutions (Peterson, 2018), society, and cultural norms (Boscarino, 2020), you are studying the macro-level. Macro-level NPF may be focused on the policy narrative changes reflected in institutional and societal norms, across historical events and culture (Ney, 2014), and typically across long periods of time.

The NPF does not assume the levels of analyses are independent. Rather, levels are understood to be interconnected, with connections also being worthy of their own research questions. For example, macro-level policy narrative disruptions are likely to have downstream effects on meso-level policy narratives. Or, shifts in micro-level narrative policy persuasion or attention (see Peterson, 2018, pp. 834–837) may affect how groups construct meso-level policy narratives. To date, NPF studies seem to land at one level of analysis or another; however, future studies accounting for the interplay between levels of analysis are both encouraged and needed (e.g., McMorris et al., 2018).

Articulating NPF Hypotheses and Expectations

The aim of this chapter is to provide guidance for NPF research that adheres to the "clear enough to be wrong" standard (Jones and McBeth, 2010),

which means that studies following this chapter's guidance should aspire to meet the scientific requirements of validity, reliability, replicability, and falsifiability.[3] As such, articulating hypotheses and expectations are standard practice in such endeavors. Indeed, much of the NPF research referenced in this chapter is hypothesis driven (e.g., Brewer 2020; Gottlieb, Oehninger, and Arnold, 2018; O'Donovan, 2018) but some include expectations (e.g., Cline, 2015; Heikkila, Weible, and Pierce, 2014; Merry, 2016) or propositions (e.g., Lybecker, McBeth, and Stoutenborough, 2016). All approaches are appropriate for a "clear enough to be wrong" NPF study, so long as they calibrate to NPF assumptions and aspire to scientific standards (see King, Keohane, and Verba, 1994, for additional guidance). The NPF has recorded several hypotheses at the micro and meso-levels of analysis (Shanahan, Jones et al., 2017), but given the rapid maturing of the NPF, new hypotheses and expectations are regularly posited. See Table 6.1 for references to several exemplar studies with hypotheses at each level of analysis. However, we want to be clear that researchers should not at all feel bound to existing hypotheses.

Specifying Your Model: Operationalizing NPF Concepts

NPF theory specifying the relationships between concepts is the scaffolding of the NPF. Consequently, valid concept operationalization is foundational to any NPF study and also critical to achieving reliability within and between NPF research models. Below, we describe current thinking about concepts, as well as citations that help provide a roadmap for the interested researcher. We understand that calibrating concepts to particular contexts as well as aspirations for greater precision and validity will lead to important conceptual innovations over time. We encourage such concept iterations, as they are critical to developing better scientific understandings of policy narratives.

The Policy Narrative

Given the principal aim of the NPF to understand the power of narratives in the policy process, policy narratives lie at the heart of the framework. A clear operational definition of what constitutes a policy narrative is thus critical, as this choice will have a cascading effect on future decisions that follow as you set up your study. While the NPF articulates a definition of policy narrative, we understand that narrative scholarship asserts different definitional criteria (e.g., Herman, 2009; Shenhav, 2015). Such alternate definitions are not strictly prohibited, and while not common, some have provided interesting alternatives (see Weible et al., 2016). However, because the NPF focuses on narratives in policy domains, a trend in NPF research has led to what has become orthodoxy in terms of defining a policy narrative. The minimum requirements for a text (understood as *anything* potentially containing a narrative) to be considered a policy narrative (vs. nonnarrative) is the presence of at least one character and the presence of a policy reference (Shanahan et al., 2018). When specifying your model, a researcher must be

absolutely clear about how policy narratives are defined, as this ensures you are working with narrative data. Moreover, failure to define policy narratives would be an obvious violation of assumption #3, discussed earlier in this chapter.

Drilling down into the policy narrative, the NPF takes a structuralist approach to narrative (see Herman, 2009, pp. 23–36) by designating two components, within which all narrative objects can be classified: narrative form (also sometimes referred to as narrative structure) and narrative content. Manifesting NPF assumption #3, narrative form refers to the generalizable structures of narrative that are theorized to exist across space and time, regardless of context. Within narrative form, the NPF identifies several elements, including setting, characters, plot, and moral of the story. Narrative content refers to the subject matter of the policy narrative (i.e., what the story is about). Manifesting NPF assumption #2 and to deal with the relative nature of narrative content, the NPF has conceptualized narrative content as two elements consisting of beliefs and strategies. As with other aspects of the NPF, we encourage science-based innovation upon and within NPF policy narrative elements.

Policy Narrative Form

Setting

Policy narrative settings situate character(s) in a time and place, thus focusing attention on "where and when the action is taking place" (Shanahan, Raile et al., 2018, p. 928). Policy settings have typically been defined in terms of geography (Knackmuhs, Farmer, and Knapp, 2018; O'Leary, Borland, Stockwell, and MacDonald, 2017), legal and/or constitutional frameworks (Boscarino, 2018; Jones, Fløttum, and Gjerstad, 2017; Smith-Walter et al., 2016), evidence (Eide Kjargard, and Söreide, 2014; Radaelli, Dunlop, and Fritsch, 2013; Schlaufer, 2018; Smith-Walter et al., 2016), resources (Husmann, 2015; Mosley and Gibson, 2017), and demographics (Kirkpatrick, 2017; McMorris, Zanocco, and Jones, 2018). Recent research has begun to refine the initially ambiguous characterization of setting with additional concepts such as proximity (Lawlor and Crow, 2018; Merry, 2018), issue frames (Shanahan, Raile et al., 2018), governance arrangements (Weiss, 2018), and cultural contexts and assumptions (Dupuis, 2018; Huda, 2019; Ney, 2014).

Characters

Like any good story, the bread and butter of policy narratives are the characters which populate the tale. The NPF initially recognized *heroes, villains,* and *victims* as key characters who were of particular interest to policy scholars. The *victim* is the person, people, or value that is being (or will be) harmed. The *hero* of the story is the agent of alleviation, the character who can solve the problem and bring relief to victims. Finally, the *villain* is the

actor who is responsible for harming the victim or standing against the hero's action to address the problem. However, these are not the *only* characters that NPF scholarship has explored, nor the only ones that could be discovered in future studies. Other character types invoked within NPF research include *beneficiaries* (Huda, 2018; O'Donovan, 2018, Weible et al., 2016), *allies* (e.g., Boscarino, 2020; Merry, 2016; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010), *opponents* (Merry, 2016), and "shadow characters" who are not referred to by name but whose existence is implied by the plot of the narrative (O'Leary et al., 2017).

Plot

Plots link characters to each other and organize their actions and interactions within the setting. Most NPF studies have operationalized plot utilizing Deborah Stone's (2012) story lines (e.g., story of decline, story of stymied progress) (e.g., Cristoforetti and Querton, 2019; Dupuis, 2018; Jones, 2014b; McBeth et al., 2017; McBeth et al., 2013; O'Donovan, 2018; Schlaufer, 2018; Shanahan et al., 2013; Veselková, 2014). Other operationalizations include beginning–middle–end (e.g., Boscarino, 2020; Honeck, 2018; Nisbett, 2017), intention of the villain character (Crow et al., 2017), threats and opportunities (O'Leary et al., 2017), stories of disorientation (Dupuis, 2018), and plots conditioned by problem definition (Crow, Lawhon et al., 2017; Kear and Wells, 2014; Price, 2019).

Moral of the Story

The actions of the characters in a narrative are often intended to lead to some outcome or solution. This moral of the story gives the characters' actions and motivations purpose. In the context of policy narratives, the moral of the story is frequently operationalized as the policy solution (e.g., Ertas and McKnight, 2020; McGough, Bedell, and Tinkler, 2018; McMorris, Zanocco, and Jones, 2018; Shanahan et al., 2013), a moral lesson (e.g., Clemons et al., 2012; Jones, 2014a; Schwartz, 2019), or a call to action (e.g., Beck, 2018; Brewer, 2019; Jones, 2018; Jones and McBeth, 2010).

Policy Narrative Content

Narrative form provides the structural building blocks of policy narratives. Narrative content gives policy narratives meaning. As one might expect, narrative content has considerably more variation than narrative form, as what a story is about can vary along an almost innumerable number of dimensions. When you interact the story with the narrator and the audience, you get even more interpretive variation, prompting many scholars who have studied public policy through narrative to conclude that generalization is impossible (e.g., Miller, 2015). This phenomenon is referred to as narrative relativity within the NPF (see Jones et al., 2014, pp. 4–5). Similar to

Healy (2017), we reject this critique. To identify systematized ways of measuring meaning-making within policy narratives, the NPF has employed two helpful categories: beliefs and strategies.

Beliefs and Strategies

To mediate the problem of narrative relativity, the NPF recommends using well-understood belief systems, where researchers have identified systemized ways by which people assign meaning to objects and processes they encounter in the world, which of course includes narrative objects. This sounds more complicated than it is. For example, the content of policy debates (e.g., immigration debates in Europe, Australia, and the United States) can be understood by measurements of nonrandom variation in policy beliefs such as the individualism and collectivism worldviews (Arieli and Sagiv, 2018). An NPF study could investigate whether the policy narratives or individuals with an individualism belief assign a different meaning to immigration than policy narratives or individuals with a collectivism belief. Beliefs counter narrative relativity for the NPF because while beliefs vary within policy debates, they do so in predictable ways. Within the NPF, a host of belief systems have proven useful in capturing narrative content in this way, including ideology (Arnold, 2018; Chang and Koebele, 2020; Peterson, 2018; Shanahan et al., 2011), cultural theory (Jorgensen, Song, and Jones, 2018; McMorris, Zanocco, and Jones, 2018; Ney, 2014), Old and New West (McBeth et al., 2005; Shanahan, McBeth, Hathaway, and Arnell, 2008), George Lakoff's moral politics theory (Clemons, McBeth, and Kusko, 2012; Knackmuhs, Farmer, and Knapp, 2020), the advocacy coalition framework's policy beliefs (Kear and Wells, 2014; Mosely and Gibson, 2017; O'Donovan, 2018), federalism (McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; Smith-Walter, 2018), conception of good citizenship (McBeth, Lybecker, and Husmann, 2014), and several systems of categorization associated with environmental attitudes (Beck, 2018; Knackmuhs, Farmer, and Knapp, 2018; Shanahan, Jones, McBeth, and Lane, 2013).

Similar to beliefs, strategies provide ways of understanding meaning beyond narrative relativity by leveraging the strategic deployment patterns of policy narratives by policy actors. Several policy narrative strategies have been commonly examined within the NPF. For example, the *devil and angel shift*, where policy actors are portrayed as nefarious or valiant, is a regularly examined strategy, (e.g., Gottlieb, Oehninger, and Arnold, 2018; Heikkila, Weible, and Pierce, 2014; Merry, 2019). Informed by the classic work of E.E. Schattschneider (1960), *scope of conflict* strategies are designed to call actors to the cause or dissuade them from entering the narrative fray (Gupta, Ripberger, and Collins, 2014; Shanahan, Jones, McBeth, and Lane, 2013; Stephan, 2020). *Causal mechanisms* are strategies used by policy actors to assign causality and blame within a policy narrative (Price, 2019; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, McBeth, and Lane, 2013). While not as common as the previously mentioned strategies, other strategies

have also been examined such as the *use of symbols* or *policy surrogates* (Kear and Wells, 2014; McBeth, Shanahan, Arnell, and Hathaway 2007; Veselková, 2014), the *use of science and evidence* (Crow and Lawlor, 2016; Knackmuhs, Farmer, and Knapp, 2018), the *use of analogy* (Boscarino, 2018), the *impotent shift* (Brewer, 2019), and *thematic coherence and strength* (Lebel and Lebel, 2018).

Narrativity: Form + Content

Finally, we address an NPF concept that can straddle both policy narrative content and form. Narrativity refers to the extent in which a policy narrative contains more or less policy narrative elements: setting, characters, plot, moral, beliefs, and strategies (Crow and Berggren, 2014; McBeth et al., 2012; Merry, 2016). Often referred to as a narrativity index, where some additive calculation is made to assess how many narrative elements are present, this index is usually employed to determine how complete or incomplete a policy narrative is and to what effect (Boscarino, 2020; Crow and Lawlor, 2016; Huda, 2018).

Research Designs

Research designs are the overall plan that one develops to address and provide answers to research questions. Typically, research questions lead to one or more hypotheses or propositions from which emanate strategies to acquire data and test those hypotheses or propositions. Below, we divide NPF research designs into two broad categories: experimental and nonexperimental.

Experimental Designs

In the most basic of terms, experiments work by examining the change in "something" based upon another "something," while making sure other "somethings" do not interfere. In more technical terms, we observe the variation in the dependent variable based upon an experimental intervention, while holding other independent variables constant. This is considered the gold standard of scientific research designs. Within the NPF, experimental designs have been employed at both micro (e.g., Zanocco, Song, and Jones, 2018) and meso (e.g., McMorris et al., 2018) levels of analysis, although micro-experimental applications are far more ubiquitous.

While NPF experimental designs vary, there are definitely some commonalities. Whether within-subjects or between-subjects designs, most NPF experiments start with a policy narrative stimulus that is introduced to experimental subjects, conditions are controlled for, and then variation in some dependent variable, such as policy preferences or assessment of risk, is observed before and after the experimental treatment (or with and without treatments in the case of between-subjects designs). The narrative experimental intervention within these research designs will usually hold some portion of the policy narrative itself constant, while manipulating narrative

elements, such as the characters (e.g., Husmann, 2015; Jones, 2014a; Shanahan et al., 2019) or the moral of the story (e.g., Jones, Fløttum, and Gjerstad, 2017). Most typically these types of designs have occurred within the context of surveys leveraging representative samples (e.g., Guenther and Shanahan, 2020; Zanocco et al., 2018), so frequently survey items will represent control variables, such as age and education, which are then included in multivariate models used to analyze related experimental outcomes. Samples that are truly randomized and representative of relevant populations could in principle remove the need for such controls, as one would be able to assume the orthogonality of those variables. To our knowledge, these purer forms of experimental designs have yet to materialize within NPF research. For an exemplar NPF experimental study, see Jorgensen, Song, and Jones (2018).

Nonexperimental Designs

As mentioned above, the experiment is the gold standard of science. Sometimes, however, our research designs fall more in the silver, bronze, or tin categories of scientific research designs. This can happen for a myriad of reasons, including the research may not be amenable to an experimental design (e.g., what policy narratives are being deployed by interest group X?), or there is a logistic inability to conduct pre- and post-assessments, or perhaps the resources necessary for an experiment are simply unattainable, and so forth. In these situations, there are many viable research designs that do an admirable job at meeting the needs of their research questions.

A common research design within NPF studies is the case study. A case study can be employed when the primary goal of a study is to describe the policy narratives within a particular context, such as a policy subsystem (Merry, 2020). However, a stronger scientific version of the case study design is to compare multiple cases along dimensions of similarity and difference (see Yin, 2014). To date, comparative case studies have had a relatively minor presence in the NPF (although, see Choiruzzad, 2019; Dupuis, 2018 Gottlieb et al., 2018, for notable exceptions). However, additional studies utilizing this design can play an important role in refining the NPF's theorized relationships by working toward *analytic generalization* (Yin, 2003, p. 32) by comparing findings from a case study to the outcomes hypothesized by theory (see Chang and Koebele, 2020, for an excellent example of this). While there is a tendency to assume that NPF case studies use primarily qualitative data collection techniques such as interviews and the interpretation of secondary sources, NPF case studies employ a host of methodologies ranging from content analysis (Lybecker et al., 2015), to interviews (Gray and Jones, 2016), to participant observation (Mosley and Gibson, 2017), to document review (O'Donovan, 2018).

Other nonexperimental NPF research designs include cross-sectional, panel, and time-series designs. Perhaps, an easy way to understand these types of designs is in terms of how and when dependent and independent variables are measured. Cross-sectional studies take a snapshot of policy

narratives, looking at them at one point in time (e.g., Lybecker, McBeth, and Kusko, 2013). Panel studies would examine repeated measurements with the same respondents at different points in time. Time-series designs, as the name indicates, measure policy narratives over time (e.g., Peterson, 2019).

Research Methods

Survey Methods

Survey methods are regularly employed in NPF research and primarily used at the micro-level of analysis, most often in conjunction with experimental designs (e.g., Husmann, 2015; Jones, 2014a; Shanahan et al., 2014; Shanahan, McBeth, and Hathaway, 2011). In such experimental designs, researchers use the survey to present a narrative experimental treatment consisting of some combination of narrative elements (e.g., characters, causal mechanism, policy beliefs) within the survey and then evaluate how the narrative influences dependent variables of interest (e.g., affective response, risk perception, policy views, or policy choice). Sampling is crucial to the ability to generalize results. As always, surveys that draw a random sample of a population are optimal (e.g., Boscarino, 2018). However, cost considerations can frequently push random samples out of reach for researchers who are typically operating on tight budgets. As a result, researchers often employ more affordable options. For example, representative samples from survey companies such as YouGov, Qualtrics, or Survey Sampling Inc. are regularly employed by NPF researchers (Jones and Song, 2014). Less commonly, surveys are also conducted with convenience samples such as university students (e.g., Ertas, 2015; Husmann, 2015) or Mechanical Turk (Clemons, McBeth, Peterson, and Palmer, 2019). These more affordable options have their drawbacks but can be appropriate under the right circumstances, with the right kinds of research questions. Researchers should always articulate the limitations of their sampling techniques and frame their discussions of findings and generalizations accordingly.

Content Analysis

Utilized in formative meso-level NPF studies (McBeth, Shanahan, and Jones, 2005; McBeth, Shanahan, Arnell, and Hathaway, 2007; Shanahan, McBeth, Hathaway, and Arnell, 2008), content analysis is currently the most commonly applied NPF meso-level methodology. Similar to Shanahan, Jones, and McBeth (2018), we offer the following seven points to consider when conducting an NPF content analysis.

The Codebook

All researchers are encouraged to publish their coding instruments with published NPF research. As such, there are now many exemplar NPF codebooks that researchers can partially or fully replicate for their own

research designs (e.g., Boscarino, 2020; Chang and Koebele, 2020; Shanahan et al., 2013). NPF codebooks are traditionally included as appendices in published research.

Census vs. Sample

Some sources of narrative data (documents, speeches, tweets) are limited enough that you do not have to draw a sample but instead can conduct a census. For example, if you have 100 documents for an interest group in the year 2020 and your goal is to code how the interest group strategically used narrative in 2020, you could code the entire 100 documents. At other times, however, limited resources and a large number of narrative sources (such as when a longitudinal analysis is being conducted) will require you to draw a random sample. For either sampling strategy, it is important to detail your sampling strategy and justify the number of sources content analyzed.

Unit of Coding Analysis

Decisions need to be made about the unit of analysis for coding your policy narratives. For more traditional coding of narrative texts, questions to consider are whether your coding unit will be a sentence, paragraph, document, or collection of documents. Where you land on these coding decisions will have critically important implications for the kinds of questions your data can address and what types of interpretations are appropriate for those data. Historically, many NPF studies have simply counted narrative elements such as heroes and villains within documents (the coding unit) and aggregated counts (e.g., McBeth et al., 2005; Merry, 2016; Shanahan et al., 2013). Other studies have deviated from this historical approach and sought to improve standardization and reliability. For example, Smith-Walter et al. (2016) code for the presence or absence of narrative elements at the paragraph level, which they argue improves reliability. Other studies have gone the direction of scaling up, to capture more of the aggregate policy narrative across documents. For example, McBeth and Lybecker (2018) code for evidence as a percentage of word count for all documents in their population, allowing them to compare the use of evidence in different venues. In short, there is no single "correct" unit of coding, but whatever choice a researcher makes, there will be guidelines to follow within the content analyses literature and limitations and benefits for each choice, specifically in how that choice intersects with NPF theory. In determining your coding unit of analysis, we suggest consulting both relevant NPF research, as well as literature providing guidance on how best to do content analysis.

Human vs. Automated Coding

Given the essential importance of context within narratives, NPF studies have primarily relied on the use of trained human coders (e.g., Merry,

2016). Consider the following: "Donald Trump has saved (or destroyed) the U.S. economy and he has benefited (or hurt) average Americans." In a narrative, context is everything. Trump is a hero in one case and a villain in the other. "Average Americans" could be either beneficiaries or victims. Most automated approaches have trouble reliably accounting for the nuance inherent in narrative context. However, there have been some forays into solving the puzzle of automated narrative coding (e.g., Crow and Wolton, 2020), though the use of automated coding is still in the development stage.

Use Independent Human Coders

Our experience is that policy narrative content coding is an iterative process that requires practice. NPF studies normally proceed with human coding as follows. First, a codebook is developed to guide coding. Second, coders are then trained on the operational definitions of concepts to be coded. Third, trained coders independently code a small number (around 5–10) of narrative data (textual or visual). The purpose of this initial coding is to test the extent to which coders are similarly applying the operational definitions to the narrative data. Fourth, the extent to which coders are reliably coding the narrative data is assessed (e.g., simple agreement or Dean Freelon's ReCal reliability calculator http://dfreelon.org/utils/recalfront/). Fifth, coders meet for reconciliation sessions to reconcile codes. At this point, disagreements in codes are normal, and detailed discussions over operational definitions will result in tweaks to the codebook's specified decision rules. This process is crucial for reliable content analysis and takes time.

Reliability Testing

After the coding and reconciliations are complete, intercoder reliability is calculated. While many NPF studies (e.g., McBeth and Lybecker, 2018) use percent agreement to assess reliability, McHugh (2012) argues that chance agreements are not accounted for with percent agreement reliability and that more rigorous tests should be used. Lombard, Snyder-Duch, and Bracken (2002) provide a summary of such tests including Scott's pi, Krippendorff's alpha, and Cohen's kappa coefficient. These more rigorous tests are now the norm within NPF scholarship (e.g., Boscarino, 2020; Smith-Walter et al., 2016). In fact, providing both percent agreement and these more advanced measures is recommended (e.g., Merry, 2016).

Coding Visual Narratives

The study of narrative within the NPF has primarily focused on text. Yet, we know that visual narratives exist, and such images are increasingly important in contemporary policy issues. The NPF has been leveraged to study these visual narratives, showing both the promise of such an approach

and the inherent difficulties of coding them. The first such NPF study, by McBeth, Shanahan, Anderson, and Rose (2012), conducted a content analysis of YouTube videos of a Greater Yellowstone interest group, the Buffalo Field Campaign (BFC). The policy narratives within the analyzed videos were shown to be strong predictors of public attention. This study and those that have followed (Lybecker, McBeth, Brewer, and De Sy, 2018; Lybecker, McBeth, Husmann, and Pelikan, 2015) have demonstrated that the reliability of NPF video coding is consistent with the reliability of textual policy narrative coding, so long as independent coders participate in the reconciliation sessions described above.

Other NPF studies have examined still frame visual images as narratives. For example, Boscarino (2020) content analyzed images from tweets of five environmental groups involved in the Dakota Pipeline controversy between 2016 and 2017. Her findings demonstrate that tweets with images were higher in narrativity than tweets that did not have images and that Twitter users are more likely to pay attention to tweets with images. Finally, Guenther and Shanahan (2021) conducted a survey experiment related to bat management in Australia and found that narratives with accompanying images are linked to emotional reactions that influence risk perception and support for management policies.

Interviews and Focus Groups

Interviews are a classic means by which to obtain social scientific data and the NPF has no shortage of studies leveraging these classic tools. Within the NPF, interviews are typically semi-structured, where some questions are scripted, but allowances are made for the emergence of new data during the interview. The structured aspect of the interview allows researchers to ask questions related to specific NPF elements (e.g., "So, who is the bad guy here?"), which speaks directly to NPF theoretical concepts of interest and allows for targeted data acquisition (Peltomaa, Hildén, and Huttunen, 2016). The nonstructured aspects of the interview allow for inductive aspects of a policy narrative to emerge, which is often relevant to providing the nuances of narrative content as well as other emergent information and concepts. Within the NPF, interviews have been used as primary data sources themselves (Shanahan et al., 2018) as well as to inform other elements of a research design, such as survey and experimental treatment development (Laufer and Jones, 2021; McMorris et al., 2018).

While not as common as interviews, focus groups have also been used in NPF research. For example, Smith-Walter, Jones, Shanahan, and Peterson (2019) used narrative data collected from focus groups to explore differences between character usage in the construction of problems and solutions related to campaign finance reform in the United States. Four focus groups were conducted, each constituted by individuals of similar cultures but distinct from the other focus groups (Kahan, 2012). Participants in each group were presented with the exact same campaign finance information, which

they then proceeded to make sense of as a group. Transcripts were content analyzed, and networks analyzed (see, section on social network analysis [SNA], below) to examine the emergence of distinctive worldview group-based narratives. Findings showed that each group produced a unique campaign finance narrative that was anchored to their cultural worldviews.

Policy Narrative Data

Because we are *Homo narrans*, we are continually surrounded by narrative data, which means narrative data are virtually everywhere, but that does not mean everything is a policy narrative. Some visual and textual data are nonnarrative. So how best to acquire narrative data? We find approaching narrative data by level of analysis a useful way to grapple with this facet of research design. A working assumption throughout the following sections is that you have settled on a clear definition of policy narrative. Without it, it is hard to imagine what collecting viable NPF policy narrative data would even look like.

Micro-Level Narrative Data

Micro-level NPF analyses focus on the relationship between policy narratives and individuals. Focusing on such relationships can lead to a myriad of forms of data. A researcher might solicit narrative data through a survey to assess aggregate understandings of policy narratives, policy preferences (Knackmuhs, Farmer, and Knapp, 2020), or public opinion (Jorgensen, Song, and Jones, 2018). Interviews might be used to inquire about the policy narratives of elites or stakeholders (Gray and Jones, 2016) or focus groups to assess the narrative understanding of everyday individuals, but in a more in-depth way than a survey might provide (Smith-Walter et al., 2019).

As we argued earlier, one of the most potent types of narrative data within NPF studies is the experimental variety. These types of data are typically derivative or relevant to a specific policy issue or area. Researchers have approached coming to terms with the operative policy narratives and assembling treatments in a variety of ways, including reviewing extant literature (Jones and Song, 2014; Zanocco et al., 2018), interviewing stakeholders (McMorris et al., 2018), through surveying the public (Jorgensen, Song, and Jones, 2018), and natural language processing (Shanahan et al., 2019). In reviewing these studies, we offer a bit of advice. If the goal is to have the policy narrative represent actual operating narratives in the policy area, researchers should try to get as close to the source as possible. That is, talk to people (e.g., interviews, focus groups) or procure proposed policy documents and work toward using source-centric policy language.

Meso-Level Narrative Data

Meso-level NPF data focus on policy narrative related to groups or advocacy coalitions. Those groups are constituted by individuals (e.g., public officials,

members of media, citizens, etc.) but can be assessed by the narratives of those individuals or through policy narratives of the groups or coalitions themselves. Many of these narratives are simply sitting out there in the public domain and are thus freely available via group or organizational websites (McBeth et al., 2005), published newsletters (McBeth et al., 2007), blogs, speeches, legislative records (McBeth and Lybecker, 2018), parliamentary inquiries, court rulings, tweets (Boscarino, 2020; Gupta, Ripberger, and Wehde, 2018; Merry, 2016), Facebook posts, traditional media sources (Peltomaa, 2018), and so on. We suggest consulting some of the NPF meso-level studies referenced in this chapter for some ideas on where to find data, but availability is literally as vast as the internet itself. That does not mean, however, that the NPF researcher should infer that all policy narrative data are public narrative data. There may very well be good cause to acquire data that is not publicly available, such as the policy narratives of specific elite policy actors (e.g., Gray and Jones, 2016) whose stories carry considerable weight, as well as the stories of those who have been ignored or deliberately excluded from the policy process (e.g., Sievers and Jones, 2020).

Once you have decided on the sources of your policy narratives, you will need to determine what time period of data will be collected. Time periods within NPF studies can vary substantially, and there is no "set" time frame for all studies. You might examine 15 months of data (Boscarino, 2020); you might examine over 50 years (Peterson, 2018). Time period considerations are determined by the context of the policy issue or area you are examining and the nature of your research question(s). You will then follow this choice with a determination about whether to use a census or sample of the population, as we discussed earlier in the chapter in the section on content analysis.

Once the above has been worked through, the next step is to determine which of the codable items (e.g., a newsletter or tweet) are narratives and which are not. This is done by applying your definition of policy narrative, which typically in the NPF will mean that a character is present and that there is a reference to the policy. In the interest of replication and transparency, document your rules of inclusion and exclusion.

Macro-Level Narrative Data

Even after ten years since the naming of the NPF, there have only been a few macro-level studies. Within these studies, the macro-level has been characterized as institutions (Peterson, 2018), culture (Ney, 2014), and even as a kind of approach in its own right (Knox, 2013). Data for such studies include the content of State of the Union Addresses relating to the presentation of environmental messages (Peterson, 2018) and written documents postulating the most desirable types of cultural relationships to undergird societal functioning (Ney, 2014), among others. However and despite the noble attempts at specifying the macro-level, in our estimation, existing macro-level studies have not yet quite captured the intent of this level of

analysis.[4] While the NPF's short-hand description of macro-level narratives has indeed been institutional and cultural narratives, in *Theories of the Policy Process*, we provide a more detailed depiction that references Danforth (2016, p. 584) who described macro-level narratives as "communal, historical narratives that are expansive enough to explain a variety of human events across time and place." In fact, similar to Lyotard's (1984) conception of meta narrative, we called macro-level narratives grand narratives or narratives conceived as the society-wide cultural and institutional stories that under-write our other stories at the meso and micro levels. In short, these are the big stories people use to justify and make sense of the smaller stories in their lives. We expect data for macro-level studies to cover large swathes of time, permeate many institutions, and house many of the meso and micro narratives we see elsewhere in a given polity. Such a grand conception is notoriously hard to measure, but we believe macro-level NPF studies will come from narrative treatments of historical events over time (e.g., Büthe, 2002), through the examination of historical archives and the like.

Analyzing Narrative Data

What Analytic Technique to Use?

Many NPF researchers rely heavily on statistical data analysis to test hypotheses derived from their research questions. In doing so, they utilize various statistics and related analytical methods including, but not confined to, simple descriptive statistics (frequencies, probability density, central tendency measures, dispersion measures, etc.), the examination of group mean difference (two sample *t*-test, analysis of variance [ANOVA], etc.), relational analysis (crosstab with chi-square test, correlation analysis, etc.), more predictive tools (regression analysis, causal mediation analysis, Bayesian posterior simulation, topic modeling, etc.), and applied network analysis. Of course, you, as a researcher, need to choose appropriate statistics and data analytics based upon the kind of research question and hypothesis being asked and tested, and the nature of data, variables, and measures being used in your research. To help you make this decision, we offer a general discussion of several statistical approaches found within NPF research. We acknowledge that there are many more analytical tools—both qualitative and quantitative—that are not covered here and yet are also appropriate for NPF research. For those interested in in-depth discussions of applying both qualitative and quantitative analytic techniques, we recommend consulting additional texts dedicated to such discussions (e.g., Agresti, 2018; Miles et al., 2018).

Descriptive Statistics

What are the primary considerations when you conduct univariate analysis with descriptive statistics in your NPF research? First, it is important to be transparent about your data source and related data collection protocols (e.g.,

Shanahan et al., 2008, p. 123). Second, examining descriptive statistics of your data is a critical first step of your overall data analysis, as it will help you understand distributional characteristics of the key variables. This understanding can in turn guide you to the appropriate follow-up statistical analysis and the proper interpretation of related analytical results (e.g., Shanahan, Adams et al., 2014, p. 77). Finally, visualizing univariate distributions of your data and variables, using graphical tools, can be extremely effective, not only in terms of communicating your data but also in terms of capturing important patterns or trends that might be missed when data are presented in typical tabular forms (e.g., Shanahan et al., 2013, pp. 464–465 and p. 470).

Relational Statistics, Differences of Means, and Predictive Tools

Much meso-level NPF research uses content analysis-based coded narrative data, mainly consisting of nominal and ordinal scale variables. When examining hypothesized relationships between variables of this sort, where policy narrative is a primary concern, NPF researchers typically employ non-parametric statistics, such as chi-square (e.g., Crow and Berggren, 2014) or chi-square with odds ratios (e.g., McBeth et al., 2007). When using interval scale variables, such as narrativity (e.g., McBeth et al., 2012), raw count of narrative elements (e.g., Shanahan, Adams et al., 2014), measures of policy beliefs derived from content coded narrative elements (e.g., McBeth et al., 2005), or from survey data such as a narrative index measure computed from multiple survey items, parametric statistical analysis can be utilized (e.g., McBeth et al., 2012).

Group mean difference test (two sample *t*-test, ANOVA, etc.), for instance, is often used when examining the similarities or differences in interval scale narrative component measures within and between advocacy coalitions in a given policy subsystem (Shanahan et al., 2013) or when analyzing the interval scale changes to outcome variables as a result of discrete narrative intervention or stimulus (Shanahan, Adams et al., 2014). Of course, bivariate correlation analysis is useful in investigating the relationship between two interval scale NPF variables as well.

A growing number of NPF researchers utilize various regression-based techniques while modeling narrative components and strategies as independent variables in their regression equations. For instance, ordinary least squares (OLS) regression estimation is implemented while public opinion or risk perception is regressed on narrative elements (Shanahan, Adams et al., 2014) or strategies (Jones, 2014a). Probit regression analysis has also been conducted to investigate why and how various stakeholders select a certain type of policy narrative or a science statement as a preferred method of their public communications (Lybecker, McBeth, and Stoutenborough, 2016).

More recently, causal mediation analysis has been utilized to untangle more complex multilayered structural relationships between policy narratives, narrative elements, and policy outcomes in micro-level narrative cognition research (Guenther and Shanahan, 2020; Zanocco et al., 2018). Bayesian

posterior simulations have also been employed to maximize the utility and interpretability of OLS regression-based causal inferences regarding triadic relationships among important narrative and policy outcome variables while more explicitly addressing uncertainties surrounding related estimations derived from micro-level experimental narrative cognition data (Jorgensen et al., 2018).

Network Analysis

Networks are a way of conceptualizing, visualizing, and understanding various relationships and interactions among nodes within a system (Borgatti, Everett, and Johnson, 2018). Nodes have definable characteristics or attributes and are directly or indirectly linked to other nodes through ties, also variously specified. What these nodes are in NPF studies can be anything from actors with coalitions or groups to elements of policy narratives (e.g., Heikkila et al., 2014). Studying these nodes and their linkages has proven useful for NPF researchers. A fairly recent NPF network application analyzed meso-level narrative networks for similarity within coalitions and dissimilarity between coalitions, in terms of the use of narrative elements (Smith-Walter et al., 2019). Similarly, Weible et al. (2016) applied network analyses to explore ties between narrative characters and actions. Relatedly, Jones and Song (2014) apply hierarchical cluster analysis to experimental micro-level NPF data to examine how narrative elements are cognitively organized by individuals after they were exposed to narrative experimental treatments. Though it is currently in its infancy within the NPF research community, NPF researchers are increasingly finding these kinds of network techniques useful in illustrating and comparing policy narrative uses and distributions across different actors.

Big Data Analytics?

With the recent advancement in information and communication technologies, including digital devices, network infrastructure, and online platform developments, "big data" has become a crucial part of our daily lives, and big data analytics serves as an important analytical tool for data-driven policy decision (McNeely and Hahm, 2014). Some policy researchers utilize big data analytics to extract value-added information from a large volume of unstructured data of various types (e.g., number, image, audio, text) in an attempt to support better decision-making (e.g., Guo and Vargo, 2015). However, its big data's application to policy research in general and to NPF research in particular is still nascent and developing. We offer this commentary here because we believe that big data analytics are an important part of the NPF's future. Though not yet fully realized, we believe that machine learning-based large-scale text analytics, holds promise. Considering the ubiquitous nature of text-as-data in policy process, NPF researchers can build massive text data from various "digital archives" (e.g., *New York Times*, Facebook, Twitter) using application user interfaces (APIs), preprocess the unstructured text

data to generate document-term matrix (DTM) and document-feature matrix (DFM) where narrative elements are important terms or features, and perform an automated content analysis-based descriptive data analysis or structural topic modeling-based explanatory data analysis depending on the research question or hypothesis. We expect that such innovations would play a substantial role in moving the framework toward prediction.

Future NPF Research

In this chapter, we have provided a "how to" guide for the NPF that takes the reader from theory to method. We realize that much of the guide is disproportionately about methods. Given the methodological emphasis of this chapter, we think it is necessary to point out that methodologies do not define NPF research nor in any way can they independently signal an NPF study. Rather, it is the understanding and acceptance of NPF's theoretical assumptions (i.e., social constructions, bounded relativity, generalizable structural elements, levels of analysis, *Homo narrans* model of the individual), how those assumptions facilitate and allow for concept construction and validity, and how various methodologies attach to those concepts and the relationships between them, that define an NPF study. This, we think, is the most essential takeaway from this chapter.

As to future NPF research, we expect that research ventures following this guide will no doubt continue the tradition of scientific methodologies that are transparent, focused on replicability, and "clear enough to be wrong" (Sabatier, 2000 p. 137). And while there is room enough for additional NPF hypotheses, we would like to nudge readers in the direction of testing and refining the framework's existing hypotheses, as such testing and retesting is paramount to the development of scientific knowledge, especially if that attention manifests in comparative applications across diverse contexts. We also hope this chapter has helped show the continued need for the NPF to be methodologically pluralistic, as no one methodology has presented itself as singularly best at describing policy narratives nor at explaining their role in the policy process, across all contexts. Finally, in the process of presenting this chapter, we identified many puzzles and challenges for future NPF research: e.g., concept operationalizations, at different levels of analysis, big and small data, data formats, and so on. We hope you find our observations both interesting and useful, perhaps even to the extent that they help motivate your own research. We expect, however, that inquisitive and curious scholars will find their own lacunas in the literature, gaps in theory, and opportunities for methodical and operational innovations. Indeed, it is our sincerest hope that you will do just that.

Notes

1 A content analysis was conducted by multiple coders who read the articles and then answered questions about the article's use of NPF concepts, levels of analysis, hypotheses tested, methodological approach, and statistical analyses. We would

like to thank Rachael Moyer and Briana Huett for their invaluable assistance in collecting and evaluating these articles.

2 Importantly, this assumption is built upon ten postulates that aggregate many scientific studies and theories from multiple academic fields. It is most often necessary to make this assumption at meso and macro levels, but aspects of the individual postulates are regularly tested and expanded upon in micro-level NPF studies.

3 For integrating alternate epistemologies with the NPF, see Gray and Jones (2016) and Jones and Radaelli (2015).

4 We do not intend in any way to demean the work of those who have examined the macro level. Rather, we appreciate these studies as they are both laudable and provide critical insights into what the next steps ought to be.

References

Agresti, Alan. 2018. *Statistical Methods for the Social Sciences*, 5th edition. Boston, MA: Pearson.

Arieli, Sharon, and Lilach Sagiv. 2018. Culture and problem-solving: Congruency between the cultural mindset of individualism versus collectivism and problem type. *Journal of Experimental Psychology: General* 147(6): 789.

Arnold, Annika. 2018. How to Understand the Role of Narratives in Environmental Communication: Cultural Narrative Analysis. In: Climate Change and Storytelling. Palgrave Studies in Environmental Sociology and Policy. Palgrave Macmillan. DOI: 10.1007/978-3-319-69383-5_3

Beck, Marisa. (2018). Telling stories with models and making policy with stories: An exploration. *Climate Policy* 18(7): 928–941. doi:10.1080/14693062.2017.1404439

Borgatti, Stephen P., Martin G. Everett, and Jeffrey C. Johnson. 2013. *Analyzing Social Networks*. New York, NY: Sage.

Boscarino, Jessica E. 2018. From Three Mile Island to Fukushima: The impact of analogy on attitudes toward nuclear power. *Policy Sciences* 52(1): 21–42.

Boscarino, Jessica E. 2020. Constructing visual policy narratives in new media: The case of the Dakota access pipeline. *Information, Communication & Society* 1–17. www.tandfonline.com/doi/abs/10.1080/136i9118X.2020.1787483

Brewer, Adam M. 2019. A bridge in flux: Narratives and the policy process in the pacific northwest. *Review of Policy Research*. 36(4): 497–522. doi:10.1111/ropr.12343

Brewer, Adam M. 2020. Strategic policy narratives: A narrative policy study of the Columbia river crossing. *Public Policy and Administration* 36(4):496–517. doi:10.1177/0952076720904434

Büthe, Tim. 2002. Taking temporality seriously: Modeling history and the use of narratives as evidence. *American Political Science Review* 96(3): 481–493. doi:10.1017/s0003055402000278

Chang, Katherine T., and Koebele, Elizabeth A. 2020. What drives coalitions' narrative strategy? Exploring narratives around school choice. *Politics & Policy* 48(4): 618–657.

Choiruzzad, Shofwan Al Bannan. 2019. Save Palm Oil, Save the Nation: Palm Oil Companies and the Shaping of Indonesia's National Interest. *Asian Politics & Policy* 11(1): 8–26. DOI: 10.1111/aspp.12431

Clarke, Kris. 2015. The case of a needle exchange policy debate in Fresno, California. *Critical Social Policy* 36(2): 289–306. doi:10.1177/0261018315608726

Clemons, Randy S., Mark K. McBeth, and Elizabeth Kusko. 2012. Understanding the role of policy narratives and the public policy arena: Obesity as a lesson in public policy development. *World Medical & Health Policy* 4(2): 1–26.

Clemons, Randy S., Mark K. McBeth, Rolfe D. Peterson, and Carol L. Palmer. 2019. The narrative policy framework and sticky beliefs: An experiment studying islamophobia. *International Journal on Minority and Group Rights* 27: 1–29.

Cline, Kurt. 2015. Targeting agriculture: Air quality policy in California's San Joaquin Valley. *California Journal of Politics and Policy* 7(4): 1–15. doi:10.5070/p2cjpp7428923

Colville, K. 2019. "Please Consider My Situation": Policy narratives and epistemological authority. Presented at the 76th Midwest Political Science Association Annual Conference, Chicago, Illinois. April 4–7.

Cristoforetti, Brice, and Querton, Lara. 2019. Migration policies debates in the European Parliament: Does the mainstream right stand its ground? In Costa O. (eds) *The European Parliament in times of EU Crisis: Dynamics and Transformations.* European Administrative Governance. pp. 299–320. Palgrave Macmillan. doi:10.1007/978-3-319-97391-3_14

Crow, Deserai, and John Berggren. 2014. The science of stories: Applications of the narrative policy framework in public policy analysis. In Jones, Michael D., Shanahan Elizabeth A. and McBeth Mark K. (eds.) *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis* (pp. 131–156). New York, NY: Palgrave Macmillan.

Crow, Deserai, John Berggren, Lydia A. Lawhon, Elizabeth A. Koebele, Adrianne Kroepsch, and Juhi Huda. 2016. Local media coverage of wildfire disasters: An analysis of problems and solutions in policy narratives. *Environment and Planning C: Politics and Space* 35(5): 849–871. doi:10.1177/0263774x16667302

Crow, Deserai, Lydia A. Lawhon, John Berggren, Juhi Huda, Elizabeth A. Koebele, and Adrianne Kroepsch. 2017. A narrative policy FRAMEWORK analysis of wildfire policy discussions in Two Colorado communities. *Politics & Policy* 45(4), 626–656. doi:10.1111/polp.12207

Crow, Deserai, and Andrea Lawlor. 2016. Media in the policy process: Using framing and narratives to understand policy influences. *Review of Policy Research* 33(5): 472–491. doi:10.1111/ropr.12187

Crow, Deserai and Laura Wolton. 2020. Talking policy in congressional campaigns: Construction of policy narratives in electoral politics. *Politics & Policy* 48(4): 658–699.

Danforth, Scot. 2016. Social Justice and Technocracy: Tracing the Narratives of Inclusive Education in the USA. *Discourse: Studies in the Cultural Politics of Education*. 37(4): 582-599. Doi: 10.1080/01596306.2015.1073022

Dupuis, Nicole. 2018. Stories of the sharing economy: Policy narratives surrounding the entry of transportation network companies into four mid-sized American cities. *Critical Policy Studies* 13(3): 306–327. doi:10.1080/19460171.2018.1437459

Eide Kjærård, Helene M. and Gunn Elisabeth Søreide. 2014. A Norwegian policy perspective on the relation between school leadership and pupils' learning outcomes. *Policy Studies* 35(6): 576–591. Dx.doi.org/10.1080/01442872.2014.971729

Ertas, Nevbahar. 2015. Policy narratives and public opinion concerning charter schools: Policy narratives and public opinion. *Politics & Policy* 43(3): 426–451. doi:10.1111/polp.12120

Ertas, Nevbahar, and Andrew N. McKnight. 2020. Clarifying and reframing the neoliberal critique of educational policy using policy process theories. Retrieved February 15, 2021, from www.tandfonline.com/doi/abs/10.1080/01596306.2019.1569881?journalCode=cdis20

Freelon, Deen. 2013. ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science* 8(1): 10–16.

Gottlieb, Madeline, Ernst Bertone Oehninger, and Gwen Arnold. 2018. "No fracking Way" VS. "drill BABY DRILL": A restructuring of who is pitted against whom in the narrative policy framework. *Policy Studies Journal* 46(4): 798–827. doi:10.1111/psj.12291

Gray, Garry, and Michael D. Jones. 2016. A qualitative narrative policy framework? Examining the POLICY narratives of US campaign finance regulatory reform. *Public Policy and Administration* 31(3): 193–220. doi:10.1177/0952076715623356

Guenther, Sarah K., and Elizabeth A. Shanahan. 2021. Communicating risk in human–wildlife interactions: How stories and images move minds. *PLoS ONE* 15(12): e0244440. https://doi.org/10.1371/journal.pone.0244440

Guo, Lei, and Chris Vargo. 2015. The power of message networks: A big-data analysis of the network agenda setting model and issue ownership. *Mass Communication and Society* 18(5): 557–576. doi:10.1080/15205436.2015.1045300

Guenther, Sara K. and Shanahan, Elizabeth A. 2020. Communicating Risk in Human-Wildlife: How stories and images move minds. PLoS ONE 15(12): e0244440. https://doi.org/10.1371/journal.pone.0244440

Gupta, Kuhika, Joseph T. Ripberger, and Savannah Collins. 2014. *The science of stories: Applications of the narrative policy framework in public policy analysis*. New York, NY: Palgrave Macmillan.

Gupta, Kuhika, Joseph T. Ripberger, and Wesley Wehde. 2018. Advocacy group messaging on social media: Using the narrative policy framework to study Twitter messages about nuclear energy policy in the United States. *Policy Studies Journal* 46(1): 119–136.

Healy, Kieran. 2017. Fuck nuance. *Sociological Theory* 35(2): 118–127.

Heikkila, Tanya, Christopher M. Weible, and Jonathan J. Pierce. 2014. *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis*. New York, NY: Palgrave Macmillan.

Herman, David. 2009. *Basic Elements of Narrative*. Malden, MA: Wiley-Blackwell.

Honeck, Thomas. 2018. A touch of post-truth: the roles of narratives in urban policy mobilities. *Geographica Helvetica* 73: 133-145. Doi.org/10.5194/gh-73-133-2018.

Husmann, Maria A. 2015. Social constructions of obesity target population: An empirical look at obesity policy narratives. *Policy Sciences* 48(4): 415–442.

Huda, Juhi. 2018. An Examination of Policy Narratives in Agricultural Biotechnology Policy in India. *World Affairs* 181(1): 2018. 10.1177/0043820018783046

——— 2019. Policy Narratives across Two Languages: A Comparative Study using the Narrative Policy Framework. *Review of Policy Research* 36(4): 523–546.

Irvin, Renee A. 2019. How slogans curate public opinion: Hard lessons from Lakoff and the linguists. *Public Integrity* 21(4): 433–447. doi:10.1080/10999922.2018.1544022

Jones, Michael D. 2014a. Cultural characters and climate change: How heroes shape our perception of climate science. *Social Science Quarterly* 95(1): 1–39. doi:10.1111/ssqu.12043

Jones, Michael D. 2014b. Communicating climate change: Are stories better than "just the facts"? *Policy Studies Journal* 42(4): 644–673. doi:10.1111/psj.12072

Jones, Michael D. 2018. Advancing the narrative policy framework? The musings of a potentially unreliable narrator. *Policy Studies Journal* 46(4): 724–746. doi:10.1111/psj.12296

Jones, Michael D., and Deserai Crow. 2017. How can we use the "science of stories" to produce persuasive scientific stories? *Palgrave Communications* 3(1): 1–9. doi:10.1057/s41599-017-0047-7

Jones, Michael D., Kjersti Fløttum, and Øyvind Gjerstad. 2017. Stories about climate change. In Kjersti Fløttum (ed) *The Role of Language in the Climate Change Debate*. London: Routledge (pp. 49–68). doi:10.4324/9781315456935-4

Jones, Michael D., and Mark K. McBeth. 2010. A narrative policy framework: Clear enough to be wrong? *Studies Journal* 38(2): 329–353.

——— 2020. Narrative in the Time of Trump: Is the Narrative Policy Framework good enough to be relevant? *Administrative Theory & Praxis* 42(2): 91–110.

Jones, Michael D., and Holly Peterson. 2017. Narrative persuasion and storytelling as climate communication strategies. *Oxford Research Encyclopedia of Climate Science*: 1-22 doi:10.1093/acrefore/9780190228620.013.384

Jones, Michael D., and Claudio Radaelli. 2015. The narrative policy framework: Child or monster? *Critical Policy Studies* 9(3): 1–17. doi:10.1080/19460171.2015.1053959

Jones, Michael D., Elizabeth A. Shanahan, and Mark K. McBeth (Eds). 2014. *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis*. Palgrave Macmillan.

Jones, Michael D., and Geoboo Song. 2014. Making sense of climate change: How story frames shape cognition. *Political Psychology* 35(4): 447–476.

Jorgensen, Paul D., Geoboo Song, and Michael D. Jones. 2018. Public support for campaign finance reform: The role of policy narratives, cultural predispositions, and political knowledge in collective policy preference formation. *Social Science Quarterly* 99(1): 216–230. doi:10.1111/ssqu.12357

Kahan, Dan M. 2012. Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision Making* 8(4): 407–424.

Kear, Andrew R., and Dominic D. Wells. 2014. Coalitions are people: Policy narratives and the defeat of Ohio Senate Bill 5. In Michael D. Jones, Elizabeth Shanahan, and Mark McBeth (eds) *The Science of Stories*. London: Palgrave MacMillan (pp. 157–184). doi:10.1057/9781137485861_8

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton, NJ: Princeton University Press. doi:10.1515/9781400821211

Kirkpatrick, Kellee J. 2017. The "not yet pregnant": The impact of narratives on infertility identity and reproductive policy. In Brian Attebery, John Gribas, Mark K. McBeth, Paul Sivitz, and Kandi Turley-Ames (eds.) *Narrative, Identity, and Academic Community in Higher Education*. New York, NY: Routledge (pp. 111–128). doi:10.4324/9781315627090-8

Knackmuhs, Eric, James Farmer, and Doug Knapp. 2018. The relationship between narratives, wildlife value orientations, attitudes, and policy preferences. *Society & Natural Resources* 32(3): 303–321. doi:10.1080/08941920.2018.1517916

———. 2020. The interaction of policy narratives, moral politics, and criminal justice policy beliefs. *Politics & Policy* 48(2): 288–313. doi:10.1111/polp.12343

Knox, Claire C. 2013. Distorted communication in the Florida everglades: A critical theory analysis of "Everglades Restoration." *Journal of Environmental Policy and Planning* 15(2): 269–284.

Laufer, A. E., and Michael D. Jones. 2021. Who pays for marine conservation? Processes and narratives that influence marine funding. *Ocean & Coastal Management* 203(1): 105504. doi:10.1016/j.ocecoaman.2020.105504

Lawlor, Andrea, and Deserai Crow. 2018. Risk-based policy narratives. *Policy Studies Journal* 46(4): 843–867. doi:10.1111/psj.12270

Lawton, Rickey N., and Murray A. Rudd. 2014. A narrative policy approach to environmental conservation. *AMBIO* 43(7): 849–857. doi:10.1007/s13280-014-0497-8

———. 2016. Scientific evidence, expert entrepreneurship, and ecosystem narratives in the UK natural environment white paper. *Environmental Science & Policy* 61: 24–32. doi:10.1016/j.envsci.2016.03.015

Lebel, Louis, and Boripat Lebel. 2018. Nexus narratives and resource insecurities in the Mekong region. *Environmental Science & Policy* 90: 164–172. doi:10.1016/j.envsci.2017.08.015

Lombard, Matthew, Jennifer Snyder-Duch, and Cheryl Campanella Bracken. 2002. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research* 28(4): 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Lybecker, Donna L., Mark K. McBeth, Adam M. Brewer, and Carine De Sy. 2018. The social construction of a border: The US–Canada border. *Journal of Borderland Studies* 33(4): 529–547.

Lybecker, Donna L., Mark K. McBeth, Maria A. Husmann, and Nicholas Pelikan. 2015. Do new media support new policy narratives? The social construction of the US–Mexico border on YouTube. *Policy & Internet* 7(4): 497–525.

Lybecker, Donna L., Mark K. McBeth, and Elizabeth Kusko. 2013. Trash or treasure: Recycling narratives and reducing political polarisation. *Environmental Politics* 22(2): 312–332. doi:10.1080/09644016.2012.692935

Lybecker, Donna L., Mark K. McBeth, and James W. Stoutenborough. 2016. Do we understand what the public hears? Stakeholders' preferred communication choices for discussing river issues with the public. *Review of Policy Research* 33(4): 376–392. doi:10.1111/ropr.12182

Lyotard, Jean Francois 1984. *The Postmodern Condition: A Report on Knowledge*. Minnesota, MN: University of Minnesota Press.

McBeth, Mark K. 2014. Preface: The Portneuf School of Narrative. In: *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis*. pp. *xiv–xviii* New York, NY: Palgrave Macmillan.

McBeth, Mark K., Randy S. Clemons, Maria A. Husmann, Elizabeth Kusko, and Alethea Gaarden. 2013. The social construction of a crisis: Policy narratives and contemporary U.S. obesity policy. *Risk, Hazards & Crisis in Public Policy* 4(3): 135–163. doi:10.1002/rhc3.12042sdaaa

McBeth, Mark K., and Donna L. Lybecker. 2018. The narrative policy framework, agendas, and sanctuary cities: The construction of a public problem. *Policy Studies Journal* 46(4): 868–893.

McBeth, Mark K., Donna L. Lybecker, and Maria A. Husmann. 2014. The narrative policy framework and the practitioner: Communicating recycling policy. *The Science of Stories* Chapter 3 in *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis*. Edited by Elizabeth A. Shanahan, Michael

D. Jones, and Mark K. McBeth. New York, NY: Palgrave-McMillian. 45–68. doi:10.1057/9781137485861_3

McBeth, Mark K., Donna L. Lybecker, James W. Stoutenborough, Sarah N. Davis, and Katrina Running. 2017. Content matters: Stakeholder assessment of river stories or river science. *Puablic Policy and Administration* 32(3): 175–196. doi:10.1177/0952076716671034

McBeth, Mark K., Elizabeth A. Shanahan, Molly Anderson, and Barbara Rose. 2012. Policy story or gory story: Narrative policy framework analysis of buffalo field campaign's YouTube videos. *Policy & Internet* 4(3–4): 159–183.

McBeth, Mark K., Elizabeth A. Shanahan, Ruth J. Arnell, and Paul L. Hathaway. 2007. The intersection of narrative policy analysis and policy change theory. *Policy Studies Journal* 35(1): 87–108.

McBeth, Mark K., Elizabeth A. Shanahan, Paul L. Hathaway, Linda E. Tigert, and Lynette J. Sampson. 2010. Buffalo tales: Interest group policy stories in greater Yellowstone. *Policy Sciences* 43(4): 391–409.

McBeth, Mark K., Elizabeth A. Shanahan, and Michael D. Jones. 2005. The science of storytelling: Measuring policy beliefs in greater Yellowstone. *Society and Natural Resources* 18(5): 413–429. McGough, David J., Claudine Bedell, and Barri Tinkler. 2018. Building a dangerous outpost in the Green mountain state: A case study of educator preparation policymaking. *Education Policy Analysis Archives* 26: 37. doi:10.14507/epaa.26.2848

McHugh, Mary L. 2012. Interrater reliability: The kappa statistic. *Biochemia Medica* 22(3): 276–282.

McMorris, Claire, Chad Zanocco, and Michael D. Jones. 2018. Policy narratives and policy outcomes: An NPF examination of Oregon's ballot measure 97. *Policy Studies Journal* 46(4): 771–797. doi:10.1111/psj.12263

McNeely, Connie L., and Jong-on Hahm. 2014. The big (data) bang: Policy, prospects, and challenges. *Review of Policy Research* 31(4): 304–310. doi:10.1111/ropr.12082

Merry, Melissa. 2016. Constructing policy narratives in 140 characters or less: The case of gun policy organizations. *Policy Studies Journal* 44(4): 373–395.

———. 2018. Narrative strategies in the gun policy debate: Exploring proximity and social construction. *Policy Studies Journal* 46(4): 747–770.

———. 2019. Angels versus devils: The portrayal of characters in the gun policy debate. *Policy Studies Journal* 47(4): 882–904.

———. 2020. *Warped Narratives: Distortion in the Framing of Gun Policy*. Ann Arbor, MI: University of Michigan Press.

Miles, Matthew B., A. Michael Huberman, and Johnny Saldaña. 2018. *Qualitative Data Analysis: A Methods Sourcebook*. London: Sage.

Miller, Hugh T. 2015. Scientism versus social constructionism in critical policy studies. *Critical Policy Studies* 9(3): 356–360.

Mosley, Jennifer E., and Katherine Gibson. 2017. Strategic use of evidence in state-level policymaking: Matching evidence type to legislative stage. *Policy Sciences* 50(4): 697–719. doi:10.1007/s11077-017-9289-x

Ney, Steven. 2014. The governance of social innovation: Connecting meso and macro levels of analysis. In Jones, M.D., Shanahan, E.A. and McBeth M.K. (eds) *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis*. New York, NY: Palgrave Macmillan.

Nisbett, Nicholas. 2017. A narrative analysis of the political economy shaping policy on child undernutrition in India. *Development and Change* 48(2): 312–338. doi:10.1111/dech.12297

O'Donovan, Kristin Taylor. 2018. Does the narrative policy framework apply to local policy issues? *Politics & Policy* 46(4): 532–570. doi:10.1111/polp.12265

O'Leary, Renee, Ron Borland, Tim Stockwell, and Marjorie MacDonald. 2017. Claims in vapour device (e-cigarette) regulation: A narrative policy framework analysis. *International Journal of Drug Policy* 44: 31–40.

Olofsson, Kristin L., Christopher M. Weible, Tanya Heikkila, and J.C. Martel. 2017. Using nonprofit narratives and news media framing to depict air pollution in Delhi, India. *Environmental Communication* 12(7): 956–972. doi:10.1080/17524032.2017.1309442

Peltomaa, Juha. 2018. Drumming the Barrels of Hope? Bioeconomy Narratives in the Media. *Sustainability* 10: 4,278; DOI:10.3390/su10114278

Peltomaa, Juha, Mikael Hildén, and Suvi Huttunen. 2016. Translating institutional change—Forest journals as diverse policy actors. *Forest Policy and Economics* 70: 172–180. doi:10.1016/j.forpol.2016.06.029

Peterson, Holly L. 2018. Political information has bright colors: Narrative attention theory. *Policy Studies Journal* 46(4): 828–842. doi:10.1111/psj.12272

———. 2019. Macro stories: Policy process dynamics of presidential environmental ideas. Unpublished doctoral dissertation. Oregon State University.

Price, Monica Hatfield. 2019. Strategic stories: Analysis of prior learning assessment policy narratives. *The Review of Higher Education* 42(2): 511–535. doi:10.1353/rhe.2019.0005

Radaelli, Claudio M., Claire A. Dunlop, and Oliver Fritsch. 2013. Narrating impact assessments in the European Union. *European Political Science* 4(3): 31–54.

Sabatier, Paul A. 2000. Clear enough to be wrong. *Journal of European Public Policy* 7(1): 135–140.

Schattschneider, E. E. 1960. *The Semisoverign People: A Realist's View of Democracy in America*. New York, NY: Holt, Rinehard and Winston.

Schlaufer, Caroline. 2018. The narrative uses of evidence. *Policy Studies Journal* 46(1): 90–118. doi:10.1111/psj.12174

Schneider, Anne, and Helen Ingram. 1993. Social construction of target populations: Implications for politics and policy. *American Political Science Review* 87(2): 334–347.

Schwartz, Noah S. 2019. Called to arms: The NRA, the gun culture & women. *Critical Policy Studies* 15(1): 74–89. doi:10.1080/19460171.2019.1697892

Shanahan, Elizabeth A., Stephanie M. Adams, Michael D. Jones, and Mark K. McBeth. 2014. The blame game: Narrative persuasiveness of the intentional causal mechanism. In Michael D. Jones, Elizabeth A. Shanahan, and Mark K. McBeth (eds) *The Science of Stories: Applications of Narrative Policy Framework*. New York, NY: Palgrave Macmillan (pp. 69–88).

Shanahan, Elizabeth A., Michael D. Jones, and Mark K. McBeth. 2018. How to conduct a narrative policy framework study. *The Social Science Journal* 55(3): 332–345.

Shanahan, Elizabeth A., Michael D. Jones, Mark K. McBeth, and Ross R. Lane. 2013. An angel on the wind: How heroic policy narrative shape policy realities. *Policy Studies Journal* 41(3): 453–483.

Shanahan, Elizabeth A., Michael D. Jones, Mark K. McBeth, and Claudio M. Radaelli. 2018. The narrative policy framework. In Chris M. Weible and Paul A. Sabatier (eds) *Theories of the Policy Process*, 4th edition. Boulder, CO: Westview Press (Chapter 5, pp. 173–214).

Shanahan, Elizabeth A., Mark K. McBeth, and Paul L. Hathaway. 2011. Narrative policy framework: The influence of media policy narratives on public opinion. *Policy & Politics* 39(3): 373–400.

Shanahan, Elizabeth A., Mark K. McBeth, Paul L. Hathaway, and Ruth J. Arnell. 2008. Conduit or contributor? The role of media in policy change theory. *Policy Sciences* 41(2): 115–138.

Shanahan, Elizabeth A., Eric D. Raile, Kate A. French, and Jamie McEvoy. 2018. Bounded stories. *Policy Studies Journal* 46(4): 922–948.

Shanahan, Elizabeth A., Ann Marie Reinhold, Eric D. Raile, Geoffrey C. Poole, Richard C. Ready, Clemente Izurieta, Jamie McEvoy, Nicholas T. Bergmann, and Henry King. 2019. Characters matter: How narratives shape affective responses to risk communication. PLOS One 10.1371/journal.pone.0225968

Shenhav, Shaul R. 2015. *Analyzing Social Narratives*. New York, NY: Routledge.

Sievers, Tjorven, and Michael D. Jones. 2020. Can power be made an empirically viable concept in policy process theory? Exploring the power potential of the narrative policy framework. *International Review of Public Policy* 2(1): 90–114. doi:10.4000/irpp.942

Smith-Walter, Aaron. 2018. Victims of health-care reform: Hirschman's rhetoric of reaction in the shadow of federalism. *Policy Studies Journal* 46(4): 894–921.

Smith-Walter, Aaron, Michael D. Jones, Elizabeth A. Shanahan, and Holly L. Peterson. (2019, May 17). The stories groups tell: Campaign finance reform and the narrative networks of cultural cognition. Retrieved February 14, 2021, from https://link.springer.com/article/10.1007%2Fs11135-019-00884-8

Smith-Walter, Aaron, Holly L. Peterson, Michael D. Jones, and Ashley Nicole Reynolds Marshall. 2016. Gun stories: How evidence shapes firearm policy in the United States. *Politics & Policy* 44(6): 1053–1088. Stephan, Hannes R. 2020. Shaping the scope of conflict in Scotland's fracking debate: Conflict management and the narrative policy framework. *Review of Policy Research* 37(1): 64–91. doi:10.1111/ropr.12365

Stone, Deborah A. 2012. *Policy Paradox: The Art of Political Decision Making*. New York, NY: W.W. Norton & Co.

Veselková, Marcela. 2014. Science, stories and the anti-vaccination movement. *Human Affairs* 24(3): 287–298. doi:10.2478/s13374-014-0227-8

Weible, Christopher M., Kristin L. Olofsson, Daniel P. Costie, Juniper M. Katz, and Tanya Heikkila. 2016. Enhancing precision and clarity in the study of policy narratives: An analysis of climate and air issues in Delhi, India. *Review of Policy Research* 33(4): 420–441.

Weiss, Jens. 2018. The evolution of reform narratives: A narrative policy framework analysis of German NPM reforms. *Critical Policy Studies* 14(1): 106–123. doi:10.1080/19460171.2018.1530605

Yin, Robert K. 2003. *Case Study Research Design and Methods*. Thousand Oaks, CA: Sage Publications.

———. 2014. *Case Study Research: Design and Methods*, 5th edition. Thousand Oaks, CA: Sage Publications.

Zanocco, Chad M., Geoboo Song, and Michael D. Jones. 2018. Fracking bad guys: The role of narrative character affect in shaping hydraulic fracturing policy preferences. *Policy Studies Journal* 46(4): 978–999. doi:10.1111/psj.12278

Form-Oriented Concepts (Meso-Level Example)
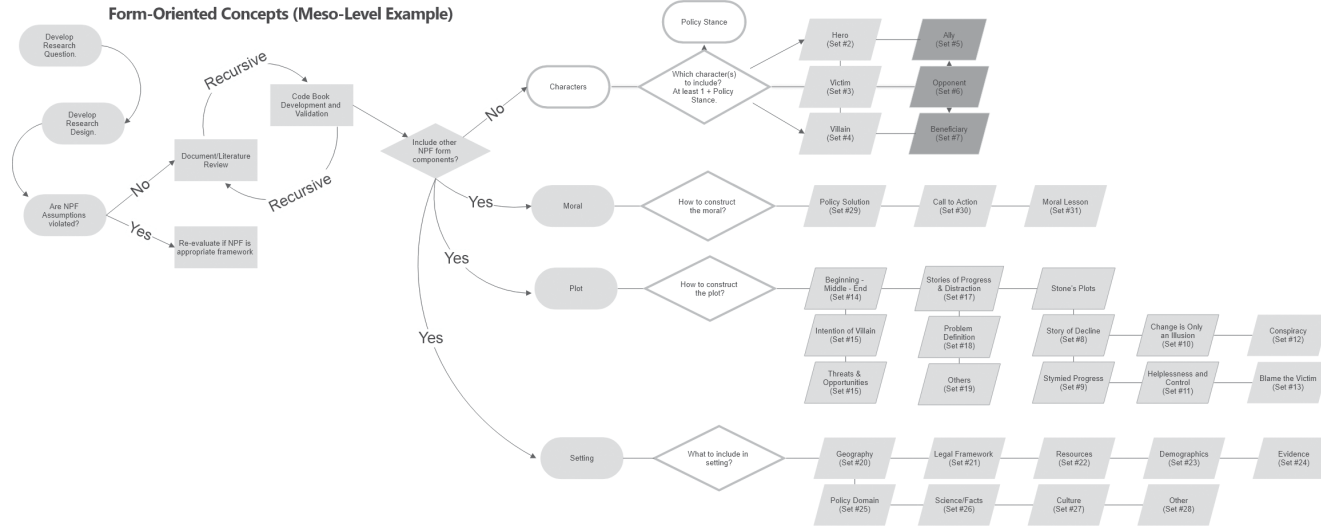
Figure 6A.1 Mapping an NPF research design

Appendix B

Table 6A.1 NPF concepts and corresponding literature

ID (See Flow Chart)	NPF Concept	Literature
Set #1	Policy Narrative	Shanahan, Jones, McBeth, and Lane 2013, p. 457 (but, see Weible , Olofsson, Costie, Katz, and Heikkila, 2016 for an alternative formulation)
	Narrativity	Boscarino, 2018; Brewer, 2020; Crow and Berggren, 2014; Crow and Lawlor, 2016; Crow and Wolton, 2020; Fløttum and Gjerstad, 2017; Huda, 2018; Merry, 2016; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, Raile, French, and McEvoy, 2018

Narrative Form

Characters

Set #2	Hero	Beck, 2018; Blair and McCormack, 2016; Boscarino, 2018, 2020; Bragg and Soler, 2017; Brewer, 2019, 2020; Ceccoli, 2019; Clemons, McBeth, and Kusko, 2012; Crow, Berggren, Lawhon, Koebele, Kroepsch, and Huda, 2016; Crow and Berggren, 2014; Crow, Lawhon, Berggren, Huda, Koebele, and Kroepsch, 2016; Crow and Lawlor, 2016; Crow and Wolton, 2020; Ertas, 2015; Fløttum and Gjerstad, 2017; Gray and Jones, 2016; Guenther and Shanahan, 2020; Heikkila, Weible, and Pierce, 2014; Honeck, 2018; Huda, 2018, 2019; Husmann, 2015; Jones, 2014a, 2014b, 2018; Jones Fløttum, and Gjerstad, 2017; Jones and McBeth, 2010; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones and Song 2014; Jones and Crow, 2017; Kear and Wells, 2014; Kirkpatrick, 2017; Knackmuhs, Farmer, and Knapp, 2018, 2020; Laufer and Jones, 2021; Lawton and Rudd, 2014; Lebel and Lebel, 2018; Lybecker, McBeth, Husmann, and Pelikan, 2015; Lybecker, McBeth, Brewer, and De Sy, 2018; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth and Lybecker, 2018; McBeth, Lybecker, and Husmann, 2014; McBeth, Lybecker, and Stoutenborough, 2016; McBeth, Shanahan, Anderson, and Rose 2012;McBeth, Shanahan, Anderson, and Rose 2012; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; McMorris, Zanocco, and Jones, 2018; Merry, 2016, 2018; Merry, 2020; Mosley and Gibson, 2017; Ney, 2014; Nisbett, 2017; O'Leary, Borland, Stockwell, and MacDonald, 2017; O'Donovan, 2018; Olofsson, Weible, Heikkila, and Martel, 2018; Peterson, 2018; Peterson, 2019; Price, 2019; Radaelli, Dunlop, and Fritsch, 2013; Raile, Shanahan, Ready, McEvoy, Izurieta, Reinhold, Poole, Bergmann, and King 2022; Schlaufer, 2018; Shanahan, Adams, Jones,

ID (See Flow Chart)	NPF Concept	Literature
		and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, McBeth, and Hathaway, 2011; Shanahan, McBeth, Hathaway, and Arnell, 2008; Shanahan, Raile, French, and McEvoy, 2018; Smith–Walter, 2018; Smith–Walter, Jones, Shanahan, and Peterson 2020; Smith–Walter, Peterson, Jones, and Reynolds Marshall, 2016; Stephan, 2020; Veselkova, 2014; Weiss, 2018; Zanocco, Song, and Jones, 2018
Set #3	Victim	Beck, 2018; Blair and McCormack, 2016; Boscarino, 2018, 2020; Bragg and Soler, 2017; Brewer, 2019, 2020; Ceccoli, 2019; Clemons, McBeth, and Kusko, 2012; Cline, 2015; Crow, Berggren, Lawhon, Koebele, Kroepsch, and Huda, 2016; Crow and Berggren, 2014; Crow, Lawhon, Berggren, Huda, Koebele, and Kroepsch, 2017; Crow and Lawlor, 2016; Crow and Wolton, 2020; Ertas, 2015; Fløttum, and Gjerstad, 2017; Gray and Jones, 2016; Guenther and Shanahan, 2020; Heikkila, Weible, and Pierce, 2014; Honeck, 2018; Huda, 2018, 2019; Husmann, 2015; Jones, 2014b, 2018; Jones, Fløttum, and Gjerstad, 2017; Jones and McBeth, 2010; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones and Song, 2014; Jones and Crow, 2017; Kear and Wells, 2014; Kirkpatrick, 2017; Knackmuhs, Farmer, and Knapp, 2018, 2020; Laufer and Jones, 2021; Lawton and Rudd, 2014; Lebel and Lebel, 2018; Lybecker, McBeth, Husmann, and Pelikan, 2015; Lybecker, McBeth, and Kusko, 2013; Lybecker, McBeth, Brewer, and De Sy, 2018; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth and Lybecker, 2018; McBeth, Lybecker, and Husmann, 2014; McBeth, Lybecker, and Stoutenborough, 2016; McBeth, Shanahan, Anderson, and Rose 2012; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; McMorris, Zanocco, and Jones, 2018; Merry, 2016, 2018; Merry, 2020; Mosley and Gibson, 2017; Ney, 2014; Nisbett, 2017; O'Leary, Borland, Stockwell, and MacDonald, 2017; O'Donovan, 2018; Olofsson, Weible, Heikkila, and Martel, 2018; Peterson, 2018; Peterson, 2019; Price, 2019; Schlaufer, 2018; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, McBeth, and Hathaway, 2011; Shanahan, McBeth, Hathaway, and Arnell, 2008; Shanahan, Raile, French, and McEvoy, 2018; Smith–Walter, 2018; Smith–Walter, Jones, Shanahan, and Peterson 2020; Smith–Walter, Peterson, Jones, and Reynolds Marshall, 2016; Stephan, 2020; Veselkova, 2014; Weiss, 2018

ID (See Flow Chart)	NPF Concept	Literature
Set #4	Villain	Beck, 2018; Blair and McCormack, 2016; Boscarino, 2018, 2020; Bragg and Soler, 2017; Brewer, 2019, 2020; Ceccoli, 2019; Clemons, McBeth, and Kusko, 2012; Cline, 2015; Crow, Berggren, Lawhon, Koebele, Kroepsch, and Huda, 2016; Crow and Berggren, 2014; Crow, Lawhon, Berggren, Huda, Koebele, and Kroepsch, 2017; Crow and Lawlor, 2016; Crow and Wolton, 2020; Ertas, 2015; Fløttum and Gjerstad, 2017; Gray and Jones, 2016; Guenther and Shanahan, 2020; Heikkila, Weible, and Pierce, 2014; Honeck, 2018; Huda, 2018, 2019; Husmann, 2015; Jones, 2014a, 2014b, 2018; Jones, Fløttum, and Gjerstad, 2017; Jones and McBeth, 2010; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones & Song, 2014; Jones and Crow, 2017; Jorgensen, Song, and Jones, 2018; Kear and Wells, 2014; Kirkpatrick, 2017; Knackmuhs, Farmer, and Knapp, 2018, 2020; Laufer and Jones, 2021; Lawton and Rudd, 2014; Lebel and Lebel, 2018; Lybecker, McBeth, Husmann, and Pelikan, 2015; Lybecker, McBeth, Brewer, and De Sy, 2018; Lybecker, McBeth, and Kusko, 2013; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth and Lybecker, 2018; McBeth, Lybecker, and Husmann, 2014; McBeth, Lybecker, and Stoutenborough, 2016; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; McBeth, Shanahan, Anderson, and Rose 2012; McMorris, Zanocco, and Jones, 2018; Merry, 2016, 2018; Merry, 2020; Mosley and Gibson, 2017; Ney, 2014; Nisbett, 2017; O'Leary, Borland, Stockwell, and MacDonald, 2017; O'Donovan, 2018; Olofsson, Weible, Heikkila, and Martel, 2018; Peterson, 2018; Peterson, 2019; Price, 2019; Radaelli, Dunlop, and Fritsch, 2013; Schlaufer, 2018; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, McBeth, and Hathaway, 2011; Shanahan, McBeth, Hathaway, and Arnell, 2008; Shanahan, Raile, French, and McEvoy, 2018; Smith–Walter, 2018; Smith–Walter, Jones, Shanahan, and Peterson 2020; Smith–Walter, Peterson, Jones, and Reynolds, Marshall, 2016; Stephan, 2020; Veselkova, 2014; Weiss, 2018; Zanocco, Song, and Jones, 2018
Set #5	Ally	Boscarino, 2020; Merry, 2016
Set #6	Opponent	Merry, 2016
Set #7	Beneficiary	Huda, 2018, 2019; Laufer and Jones, 2021; O'Donovan, 2018; Weible, Olofsson, Costie, Katz, and Heikkila, 2016

ID (See Flow Chart)	NPF Concept	Literature

Plot

Set #8	Story of Decline	Cline, 2015; Dupuis, 2018; Gray and Jones, 2016; Lawton and Rudd, 2014; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth, Lybecker, and Stoutenborough, 2016; McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Olofsson, Weible, Heikkila, and Martel, 2018; Peltomaa, Hildén, and Huttunen, 2016; Schlaufer, 2018; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, McBeth, and Hathaway, 2011; Veselková, 2014; Weiss, 2018
Set #9	Stymied Progress	Gray and Jones, 2016; Lawton and Rudd, 2014; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Peltomaa, Hildén, and Huttunen, 2016; Shanahan, McBeth, and Hathaway, 2011; Weiss, 2018; Jones, 2014a; Jones, 2014b; Jones and Song, 2014
Set #10	Change is only an Illusion	McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Weiss, 2018; Schlaufer, 2018; Lybecker, McBeth, and Kusko, 2013
Set #11	Helplessness and Control	O'Donovan, 2018; Weiss, 2018; Schlaufer, 2018; Gray and Jones, 2016; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth, Shanahan, Anderson, and Rose 2012; Shanahan, Jones, McBeth, and Lane, 2013; Mosley, and Gibson, 2017
Set #12	Conspiracy	McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Weiss, 2018
Set #13	Blame the Victim	McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Weiss, 2018; Knackmuhs, Farmer, and Knapp, 2020; Lebel and Lebel, 2018
Set #14	Beginning/ Middle/End	Boscarino, 2018; Jones, Fløttum, and Gjerstad, 2017; Jones and Song, 2014; Honeck, 2018; Nisbet, 2017; Shanahan, Adams, Jones, and McBeth, 2014
Set #15	Villain Intention	Crow, Lawhon, Berggren, Huda, Koebele, and Kroepsch, 2017
Set #16	Threats and Opportunities	O'Leary, Borland, Stockwell, and MacDonald 2017
Set #17	Stories of Disorientation	Dupuis, 2018
Set #18	Problem Definition	Crow, Lawhon, Berggren, Huda, Koebele, and Kroepsch, 2017; Jones and Crow, 2017; Kear and Wells, 2014, and Price, 2019
Set #19	Other	Beck, 2018; Boscarino, 2020; Crow and Lawlor, 2016; Jones, 2018; Jones and Peterson, 2017; Jones and Radaelli, 2015; Lawlor and Crow, 2018; McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Radaelli, Dunlop, and Fritsch, 2013; Shanahan, Adams, Jones, and McBeth, 2014; Kirkpatrick, 2017

ID (See Flow Chart)	NPF Concept	Literature

Setting

Set #20	Geography	Boscarino, 2020;Brewer, 2019; Dupuis, 2018; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones and Song, 2014; Jones, 2014a; Jones, 2018; Jones, Fløttum, and Gjerstad, 2017; Knackmuhs, Farmer, and Knapp, 2018, 2020; Lawton and Rudd, 2014; McBeth and Lybecker, 2018; McMorris, Zanocco, and Jones, 2018; O'Leary, Borland, Stockwell, and MacDonald, 2017; Olofsson, Weible, Heikkila, and Martel, 2018; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Raile, French, and McEvoy, 2018; Smith-Walter, Peterson, Jones, and Reynolds Marshall, 2016
Set #21	Legal/ Constitutional Frameworks	Boscarino, 2020; Brewer, 2019; Dupuis, 2018; Gray and Jones, 2016; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones, 2014a; Jones, 2018; Jones, Fløttum, and Gjerstad, 2017; McMorris, Zanocco, and Jones, 2018; O'Leary, Borland, Stockwell, and MacDonald, 2017; Smith-Walter, Peterson, Jones, and Reynolds Marshall, 2016
Set #22	Resources	Boscarino, 2020; Brewer, 2019; Knackmuhs, Farmer, and Knapp, 2018; Laufer and Jones, 2021; Mosley, and Gibson, 2017; Husmann, 2015
Set #23	Demographics	Boscarino, 2020; Brewer, 2019; Husmann, 2015; Dupuis, 2018; Smith-Walter, Peterson, Jones, and Reynolds Marshall, 2016; Jones, Fløttum, and Gjerstad, 2017; McMorris, Zanocco, and Jones, 2018; Kirkpatrick, 2017
Set #24	Evidence	Crow, Berggren, Lawhon, Koebele, Kroepsch, and Huda, 2016; Eide and Søreide, 2014; Jones and Crow, 2017; Jones and Radaelli, 2015; Jones and Song, 2014; Jones, 2014a; Jones, 2014b; Jones, 2018; Lawlor and Crow, 2018; Lawton and Rudd, 2016; McBeth and Lybecker, 2018; Merry, 2016; Mosley, and Gibson, 2017; O'Donovan, 2018; Radaelli, Dunlop, and Fritsch, 2013; Schlaufer, 2018; Shanahan, Jones, McBeth, and Lane, 2013; Smith-Walter, Peterson, Jones, and Reynolds Marshall, 2016
Set #25	Policy Domain	Beck, 2018; Ceccoli, 2019; Ertas and McKnight, 2020; Gray and Jones, 2016; Jones, 2018; Kirkpatrick, 2017; Lawton and Rudd, 2014; Lebel and Lebel, 2018; McBeth and Lybecker, 2018; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; Merry, 2018; O'Leary, Borland, Stockwell, and MacDonald, 2017; O'Donovan, 2018; Olofsson, Weible, Heikkila, and Martel, 2018
Set #26	Science/Facts	Beck, 2018; Clemons, McBeth, Peterson, and Palmer, 2019; Jones, 2014b; Jones and Radaelli, 2015; Shanahan, McBeth, and Hathaway 2011
Set #27	Culture	Dupuis, 2018; Huda, 2019; Ney, 2014
Set #28	Others	Gottlieb, Oehninger, and Arnold 2018; Honeck, 2018; Peltomaa, 2018; Shanahan, Jones, McBeth, and Radaelli, 2018; Weiss, 2018

ID (See Flow Chart)	NPF Concept	Literature

Moral of the Story

Set #29	Policy Solution	Crow, Lawhon, Berggren, Huda, Koebele, Kroepsch, 2017; Dupuis, 2018; Ertas and McKnight, 2020; Gray and Jones, 2016; Heikkila, Weible, and Pierce, 2014; Honeck, 2018; Husmann, 2015; Jones and Crow, 2017; Jones and McBeth, 2010; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones and Song, 2014; Jones, 2014a; Jones, 2014b; Jones, 2018; Jones, Fløttum, and Gjerstad, 2017; Jorgensen, Song, and Jones, 2018; Kirkpatrick, 2017; Knackmuhs, Farmer, and Knapp, 2018; Lawlor and Crow, 2018; Lawton and Rudd, 2014; Lebel and Lebel, 2018; Lybecker, McBeth, and Kusko, 2013; McBeth and Lybecker, 2018; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth, Lybecker, and Stoutenborough, 2016; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; McGough, Bedell, and Tinkler, 2018; McMorris, Zanocco, and Jones, 2018; Merry, 2016; Mosley, and Gibson, 2017; Ney, 2014; O'Leary, Borland, Stockwell, and MacDonald, 2017; McBeth, Shanahan, Anderson, and Rose 2012; O'Donovan, 2018; Schlaufer, 2018; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, McBeth, and Hathaway, 2011; Smith-Walter, Jones, Shanahan, and Peterson 2020; Smith-Walter, Peterson, Jones, and Reynolds Marshall, 2016; Veselková, 2014; Zanocco, Song, and Jones, 2018; Lawhon, Koebele, Kroepsch, and Huda, 2017
Set #30	Call to Action	Beck, 2018; Brewer, 2020; Jones and McBeth, 2010; Jones, 2018
Set #31	Moral Lesson	Jones and Radaelli, 2015; Jones, 2014a; Kirkpatrick, 2017; Schwartz, 2019

Narrative Content

Narrative Strategies

| Causal Mechanisms | | Beck, 2018; Bragg and Soler, 2017; Ceccoli, 2019; Clemons, McBeth, Kusko, 2012; Crow and Berggren, 2014; Crow, Berggren, Lawhon, Koebele, Kroepsch, and Huda, 2016; Dupuis, 2018; Ertas and McKnight, 2020; Gray and Jones, 2016; Husmann, 2015; Jones, 2014a; Jones, 2014a; Jones, 2018; Jones, Fløttum, and Gjerstad, 2017; Jorgensen, Song, and Jones, 2018; Lawton and Rudd, 2014; Lawton and Rudd, 2016; Lybecker, McBeth, Husmann, and Pelikan, 2015; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth, Shanahan, Anderson, and Rose 2012; |

ID (See Flow Chart)	NPF Concept	Literature
		Merry, 2018; Merry, 2019; Ney, 2014; O'Donovan, 2018; Olofsson, Weible, Heikkila, and Martel, 2018; Peterson, 2018; Price, 2019; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Smith–Walter, Jones, Shanahan, and Peterson 2020; Smith–Walter, Peterson, Jones, and Reynolds Marshall, 2016; Veselková, 2014
	Cost–Benefit	Blair and McCormack, 2016; Brewer, 2019; Cline, 2015; Crow, Lawhon, Berggren, Huda, Koebele, Kroepsch, 2017; Heikkila, Weible, and Pierce, 2014; Huda, 2018; Jones and McBeth, 2010; Kear and Wells, 2014; McBeth, Shanahan, Arnell, and Hathaway, 2007; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; Mosley, and Gibson, 2017; Peltomaa, Hildén, and Huttunen, 2016; Schlaufer, 2018; Stephan, 2020; Weiss, 2018
	Devil Shift	Brewer, 2019; Brewer, 2020; Chang and Koebele 2020; Crow and Berggren, 2014; Crow, Lawhon, Berggren, Huda, Koebele, Kroepsch, 2017; Gottlieb, Oehninger, and Arnold, 2018; Heikkila, Weible, and Pierce, 2014; Jones and McBeth, 2010; Jones and Peterson, 2017; Jones, 2018; Lebel and Lebel, 2018; McBeth and Lybecker, 2018; Merry, 2016; Merry, 2019; Merry, 2020; Schlaufer, 2018; Smith–Walter, Peterson, Jones, and Reynolds Marshall, 2016; Stephan, 2020
	Angel Shift	Brewer 2020; Chang and Koebele 2020; Crow and Berggren, 2014; Crow, Lawhon, Berggren, Huda, Koebele, Kroepsch, 2017; Gottlieb, Oehninger, and Arnold, 2018; Heikkila, Weible, and Pierce, 2014; Jones and Radaelli, 2015; Jones, 2018; Lebel and Lebel, 2018; Merry, 2016; Merry, 2019; Merry, 2020; Mosley, and Gibson, 2017; Ney, 2014; Schlaufer, 2018; Shanahan, Jones, McBeth, and Lane, 2013; Smith–Walter, Peterson, Jones, and Reynolds Marshall, 2016; Stephan, 2020
	Winning/Losing	Boscarino, 2020; Brewer, 2019; Brewer, 2020; Crow and Berggren, 2014; Chang and Koebele 2020; Crow, Lawhon, Berggren, Huda, Koebele, Kroepsch, 2017; Gottlieb, Oehninger, and Arnold, 2018; Gupta, Ripberger and Collins, 2014; Gupta, Ripberger, and Wehde, 2018; Heikkila, Weible, and Pierce, 2014; Huda, 2018; Jones and McBeth, 2010; Jones and Radaelli, 2015; Jones, 2018; Kear and Wells, 2014; Lebel and Lebel, 2018; Lybecker, McBeth, Husmann, and Pelikan, 2015; McBeth, Shanahan, Arnell, and Hathaway, 2007; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; Merry, 2016; Merry, 2019; Mosley, and Gibson, 2017; Ney, 2014; Price, 2019; Schlaufer, 2018;

ID (See Flow Chart)	NPF Concept	Literature
		Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Shanahan, McBeth, and Hathaway, 2011; Shanahan, McBeth, Hathaway, and Arnell, 2008; Stephan, 2020; Veselková, 2014
	Scope of Conflict	Boscarino, 2018; Boscarino, 2020; Brewer, 2019; Brewer, 2020; Crow and Berggren, 2014; Crow, Lawhon, Berggren, Huda, Koebele, Kroepsch, 2017; Gupta, Ripberger and Collins, 2014; Jones and McBeth, 2010; Jones and Radaelli, 2015; Jones, 2018; Kear and Wells, 2014; McBeth, Lybecker, Stoutenborough, Davis, and Running, 2017; McBeth, Shanahan, Arnell, and Hathaway, 2007; McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; Merry, 2016; Merry, 2018; Merry, 2019; Ney, 2014; Radaelli, Dunlop, and Fritsch, 2013; Shanahan, Adams, Jones, and McBeth, 2014; Shanahan, Jones, and McBeth, 2011; Shanahan, Jones, McBeth, and Lane, 2013; Smith–Walter, Peterson, Jones, and Reynolds Marshall, 2016; Stephan, 2020; Veselková, 2014; Weiss, 2018
	Other	Beck, 2018; Boscarino, 2020; Brewer, 2020; Colville, 2019; Crow and Lawlor, 2016; Honeck, 2018; Huda, 2018; Irvin, 2019; Kear and Wells, 2014; Knackmuhs, Farmer, and Knapp, 2018; Lebel and Lebel, 2018; McBeth and Lybecker, 2018; McBeth, Clemons, Husmann, Kusko, and Gaarden, 2013; McBeth, Shanahan, Arnell, and Hathaway, 2007; McGough, Bedell, and Tinkler, 2018; Sievers and Jones, 2020; Smith–Walter, 2018; Veselková, 2014

Belief Systems

	ACF Policy Beliefs	Dupuis, 2018; Ertas, 2015; Gupta, Ripberger, and Collins, 2014; Kear and Wells, 2014; Lawton and Murray, 2014; Mosley and Gibson, 2017; O'Donovan, 2018; O'Leary, Borland, Stockwell, and MacDonald, 2017
	Ideology	Clemons, McBeth, Peterson, and Palmer, 2019; Cristoforetti and Querton, 2018; Crow and Lawlor, 2016; Jones and Crow, 2017; Jones and McBeth, 2010; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones, 2018; Peterson, 2018; Schlaufer, 2018; Shanahan, Jones, and McBeth, 2011
	Cultural Theory	Jones and McBeth, 2010; Jones and Peterson, 2017; Jones and Radaelli, 2015; Jones and Song, 2014; Jones, 2014a; Jones, 2014b; Jones, 2018; Jones, Fløttum, and Gjerstad, 2017; Jorgensen, Song, and Jones, 2018; Lebel and Lebel, 2018; McMorris, Zanocco, and Jones, 2018; Ney, 2014; Shanahan, Jones, and McBeth, 2011; Zanocco, Song, and Jones, 2018

ID (See Flow Chart)	NPF Concept	Literature
	Moral Politics	Clemons, McBeth, and Kusko, 2012; Knackmuhs, Farmer, and Knapp, 2020
	Old & New West	Jones, 2018; Shanahan, McBeth, and Hathaway, 2011; Shanahan, McBeth, Hathaway, and Arnell, 2008
	Federalism	McBeth, Shanahan, Hathaway, Tigert, and Sampson, 2010; McBeth, Shanahan, and Jones 2005; Smith-Walter, Peterson, Jones, and Reynolds Marshall, 2016; Smith-Walter, 2018
	Other	Beck, 2018; Crow and Wolton, 2020; Knackmuhs, Farmer, and Knapp, 2018; Lebel and Lebel, 2018; Lybecker, McBeth, and Stoutenborough, 2016; McBeth, Lybecker, and Husmann, 2014; McBeth, Shanahan, and Jones 2005; Shanahan, Jones, McBeth, and Lane, 2013

Appendices References

Beck, M. (2018). Telling stories with models and making policy with stories: An exploration. *Climate Policy, 18*(7), 928–941. doi:10.1080/14693062.2017.1404439

Blair, B. D., and McCormack, L. (2016). Applying the narrative policy framework to the issues surrounding hydraulic fracturing within the news media: A research note. *Research & Politics, 3(1)*: 2053168016628334. 10.1177/2053168016628334

Boscarino, J. E. (2018). From Three Mile Island to Fukushima: the impact of analogy on attitudes toward nuclear power. *Policy Sciences, 52*, 21–42.

Boscarino, J. E. (2020). Constructing Visual Policy Narratives in New Media: The Case of the Dakota Access Pipeline. Information, Communication & Society, 1–17. www.tandfonline.com/doi/abs/10.1080/136i9118X.2020.1787483.

Bragg, D. D., and Soler, M.C. (2017). Policy narratives on applied baccalaureate degrees: Implications for student access to and progression through college in the United States. *Journal of Vocational Education & Training, 69(1)*: 123–146. 10.1080/13636820.2017.1289550

Brewer, A. M. (2019). A bridge in flux: Narratives and the policy process in the pacific northwest. *Review of Policy Research*. doi:10.1111/ropr.12343

Brewer, A. M. (2020). Strategic policy Narratives: A narrative policy study of the Columbia river crossing. *Public Policy and Administration,* 095207672090443. doi:10.1177/0952076720904434

Ceccoli, S. (2019). "The Language We Use Matters": Streams, Narratives, and the Obama Administration Drone Strike Program in Yemen. *Presidential Studies Quarterly, 49(3)*: 498–526. 10.1111/psq.12526

Chang, K. T., and Koebele, E. A. (2020). What Drives Coalitions' Narrative Strategy? Exploring Narratives around School Choice. *Politics & Policy, 48*(4), 618–657.

Clemons, R. S., McBeth, M. K., and Kusko, E. (2012). Understanding the Role of Policy Narratives and the Public Policy Arena: Obesity as a Lesson in Public Policy Development. *World Medical & Health Policy, 4*(2), 1–26.

Clemons, R. S., McBeth, M. K., Peterson, R. D., and Palmer, C. L. (2019). The Narrative Policy Framework and Sticky Beliefs: An Experiment Studying Islamophobia. *International Journal on Minority and Group Rights, 27,* 1–29.

Cline, K. (2015). Targeting agriculture: Air Quality policy in California's San Joaquin Valley. *California Journal of Politics and Policy, 7*(4). doi:10.5070/p2cjpp7428923

Colville, K. (2019). "Please Consider My Situation:" Policy Narratives and Epistemological Authority. Presented at the 76th Midwest Political Science Association Annual Conference, Chicago, Illinois. April 4–7.

Cristoforetti, B., and Querton, L. (2018). Migration policies debates in the European Parliament: Does the mainstream Right stand Its Ground? *European Administrative Governance,* 299–320. doi:10.1007/978-3-319-97391-3_14

Crow, D., and Berggren, J. (2014). The science of stories: Applications of the narrative policy framework in public policy analysis. In *The science of stories: Applications of the narrative policy framework in public policy analysis* (pp. X–XX). New York, NY: Palgrave Macmillan.

Crow, D. A., Berggren, J., Lawhon, L. A., Koebele, E. A., Kroepsch, A., and Huda, J. (2016). Local media coverage of wildfire disasters: An analysis of problems and solutions in policy narratives. *Environment and Planning C: Politics and Space, 35*(5), 849–871. doi:10.1177/0263774x16667302

Crow, D. A., Lawhon, L. A., Berggren, J., Huda, J., Koebele, E., and Kroepsch, A. (2017). A narrative policy FRAMEWORK analysis of wildfire policy discussions in Two Colorado Communities. *Politics & Policy, 45*(4), 626–656. doi:10.1111/polp. 12207

Crow, D. A., and Lawlor, A. (2016). Media in the policy process: Using framing and narratives to understand policy influences. *Review of Policy Research, 33*(5), 472–491. doi:10.1111/ropr.12187

Crow, D., and Wolton, L. (2020). Talking Policy in Congressional Campaigns: Construction of Policy Narratives in Electoral Politics. *Politics & Policy*, *48*(4), 658–699.

Dupuis, N. (2018). Stories of the sharing economy: POLICY narratives surrounding the entry of transportation network companies into Four mid-sized American cities. *Critical Policy Studies, 13*(3), 306–327. doi:10.1080/19460171.2018.1437459

Ertas, N. (2015). Policy Narratives and Public Opinion concerning Charter Schools: Policy Narratives and Public Opinion. *Politics & Policy, 43*(3), 426–451. 10.1111/polp.12120

Ertas, N., and McKnight, A. (2020). Clarifying and reframing the neoliberal critique of educational policy using policy process theories. Retrieved February 15, 2021, from www.tandfonline.com/doi/abs/10.1080/01596306.2019.1569881?journalCode=cdis20

Fløttum, K., and Gjerstad, Ø. (2017). Narratives in climate change discourse. *Wiley Interdisciplinary Reviews: Climate Change, 8(1)*: e429. https://doi.org/10.1002/wcc.429

Gottlieb, M., Oehninger, E. B., and Arnold, G. (2018). "No fracking Way" VS. "drill BABY DRILL": A restructuring of who is pitted against whom in the Narrative policy framework. *Policy Studies Journal, 46*(4), 798–827. doi:10.1111/psj.12291

Gray, G., and Jones, M. D. (2016). A qualitative narrative Policy framework? Examining the POLICY narratives of US campaign finance regulatory reform. *Public Policy and Administration, 31*(3), 193–220. doi:10.1177/0952076715623356

Guenther, S., and Shanahan, E. A. (2020). Communicating Risk in Human-Wildlife: How stories and images move minds. *PLoS ONE* 15(12): e0244440. https://doi.org/10.1371/journal.pone.0244440

Gupta, K., Ripberger, J. T., and Collins, S. (2014). *The science of stories: Applications of the narrative policy framework in public policy analysis*. New York, NY: Palgrave Macmillan.

Gupta, K., Ripberger, J. T., and Wehde, W. (2018). Advocacy group messaging on social media: Using the narrative policy framework to study Twitter messages

about nuclear energy policy in the United States. *Policy Studies Journal,* 46(1), 119–136.

Heikkila, T., Weible, C. M., and Pierce, J. J. (2014). *The science of stories: Applications of the narrative policy framework in public policy analysis.* New York: Palgrave Macmillan.

Honeck, T. A touch of post-truth: The roles of narratives in urban policy mobilities. *Geographica Helvetica* 73(2): 133–145. 10.5194/gh-73-133-2018

Huda, Juhi. (2018). An Examination of Policy Narratives in Agricultural Biotechnology Policy in India. *World Affairs* 181(1): 2018. 10.1177/0043820018783046

Huda, Juhi. (2019). Policy Narratives across Two Languages: A Comparative Study using the Narrative Policy Framework. *Review of Policy Research* 36(4): 523–546.

Husmann, M. A. (2015). Social constructions of obesity target population: an empirical look at obesity policy narratives. *Policy Sciences*, 48(4), 415–442.

Irvin, R. A. (2019). How slogans curate public OPINION: Hard lessons from Lakoff and the linguists. *Public Integrity, 21*(4), 433–447. doi:10.1080/10999922.2018.1544022

Jones, M. D. (2014a). Cultural characters and climate change: How heroes shape our perception of climate science. *Social Science Quarterly, 95*(1), 1–39. doi:10.1111/ssqu.12043

Jones, M. D. (2014b). Communicating climate change: Are stories better than "just the facts"? *Policy Studies Journal, 42*(4), 644–673. doi:10.1111/psj.12072

Jones, M. D. (2018). Advancing the narrative policy framework? The musings of a potentially unreliable narrator. *Policy Studies Journal, 46*(4), 724–746. doi:10.1111/psj.12296

Jones, M. D., and Crow, D. (2017). How can we use the 'science of stories' to produce persuasive scientific stories? *Palgrave Communications, 3*(1). doi:10.1057/s41599-017-0047-7

Jones, M. D., Fløttum, K., and Gjerstad, Ø. (2017). Stories about climate change. *The Role of Language in the Climate Change Debate,* 49–68. doi:10.4324/9781315456935-4

Jones, M. D., and McBeth, M. K. (2010). A Narrative Policy Framework: Clear Enough to be Wrong? *Policy Studies Journal, 38*(2), 329–353.

Jones, M. D., and Peterson, H. (2017). Narrative persuasion and storytelling as climate communication strategies. *Oxford Research Encyclopedia of Climate Science.* doi:10.1093/acrefore/9780190228620.013.384

Jones, M. D., and Radaelli, C. (2015). The Narrative Policy Framework: Child or Monster? *Critical Policy Studies, 9*(3), 1–17. DOI:10.1080/19460171.2015.1053959

Jones, M. D., and Song, G. (2014). Making Sense of Climate Change: How Story Frames Shape Cognition. *Political Psychology, 35*(4), 447–476.

Jorgensen, P. D., Song, G., and Jones, M. D. (2018). Public support for campaign finance Reform: The role of Policy Narratives, CULTURAL predispositions, and political knowledge in collective policy Preference Formation. *Social Science Quarterly, 99*(1), 216–230. doi:10.1111/ssqu.12357

Kear, A. R., and Wells, D. D. (2014). Coalitions are people: POLICY narratives and the defeat of Ohio Senate bill 5. *The Science of Stories,* 157–184. doi:10.1057/9781137485861_8

Kirkpatrick, K. J. (2017). The "not yet pregnant". *Narrative, Identity, and Academic Community in Higher Education,* 111–128. doi:10.4324/9781315627090-8

Knackmuhs, E., Farmer, J., and Knapp, D. (2018). The relationship between narratives, wildlife value orientations, attitudes, and policy preferences. *Society & Natural Resources, 32*(3), 303–321. doi:10.1080/08941920.2018.1517916

Knackmuhs, E., Farmer, J., and Knapp, D. (2020). The interaction of Policy narratives, Moral politics, and criminal justice POLICY BELIEFS. *Politics & Policy, 48*(2), 288–313. doi:10.1111/polp.12343

Laufer, A. E., and Jones, M. D. (2021). Who pays for marine conservation? Processes and narratives that influence marine funding. *Ocean & Coastal Management, 203*, 105504. doi:10.1016/j.ocecoaman.2020.105504

Lawlor, A., and Crow, D. (2018). Risk–based policy narratives. *Policy Studies Journal, 46*(4), 843–867. doi:10.1111/psj.12270

Lawton, R. N., and Rudd, M. A. (2014). A narrative policy approach to environmental conservation. *AMBIO, 43*(7), 849–857. doi:10.1007/s13280-014-0497-8

Lawton, R. N., and Rudd, M. A. (2016). Scientific evidence, Expert entrepreneurship, and Ecosystem narratives in the UK natural Environment white paper. *Environmental Science & Policy, 61*, 24–32. doi:10.1016/j.envsci.2016.03.015

Lebel, L., and Lebel, B. (2018). Nexus narratives and Resource insecurities in the Mekong region. *Environmental Science & Policy, 90*, 164–172. doi:10.1016/j.envsci.2017.08.015

Lybecker, D. L., McBeth, M. K., Brewer, A. M. and De Sy, C. (2018). The Social Construction of a Border: The US Canada Border. *Journal of Borderland Studies*, *33*(4), 529–547.

Lybecker, D. L., McBeth, M. K., Husmann, M. A. and Pelikan, N. (2015). Do New Media Support New Policy Narratives? The Social Construction of the US–Mexico Border on YouTube. *Policy & Internet, 7*(4), 497–525.

Lybecker, D. L., McBeth, M. K., and Kusko, E. (2013). Trash or treasure: Recycling narratives and reducing political polarisation. *Environmental Politics, 22*(2), 312–332. doi:10.1080/09644016.2012.692935

Lybecker, D. L., McBeth, M. K., and Stoutenborough, J. W. (2016). Do we understand what the public hears? Stakeholders' preferred communication choices for discussing river issues with the public. *Review of Policy Research, 33*(4), 376–392. doi:10.1111/ropr.12182

McBeth, M. K., Clemons, R. S., Husmann, M. A., Kusko, E., and Gaarden, A. (2013). The social construction of a CRISIS: POLICY narratives and CONTEMPORARY U.S. obesity policy. *Risk, Hazards & Crisis in Public Policy, 4*(3), 135–163. doi:10.1002/rhc3.12042sdaaa

McBeth, M. K. and Lybecker, D. L. (2018). The Narrative Policy Framework, Agendas, and Sanctuary Cities: The Construction of a Public Problem. *Policy Studies Journal*, *46*(4), 868–893.

McBeth, M. K., Lybecker, D. L., and Husmann, M. A. (2014). The narrative policy framework and the Practitioner: Communicating recycling policy. *The Science of Stories,* 45–68. doi:10.1057/9781137485861_3

McBeth, M. K., Lybecker, D. L., Stoutenborough, J. W., (2016). Do stakeholders analyze their audience? The communication switch and stakeholder personal versus public communication choices. *Policy Sciences, 49(4)*: 421–444. 10.1007/s11077-016-9252-2

McBeth, M. K., Lybecker, D. L., Stoutenborough, J. W., Davis, S. N., and Running, K. (2016). Content matters: Stakeholder assessment of river stories or river science. *Public Policy and Administration, 32*(3), 175–196. doi:10.1177/0952076716671034

McBeth, M. K., Shanahan, E. A., Anderson, M., and Rose, B. (2012). Policy Story or Gory Story: Narrative Policy Framework Analysis of Buffalo Field Campaign's YouTube Videos. *Policy & Internet, 4*(3–4), 159–183.

McBeth, M. K., Shanahan, E. A., Arnell, R. J., and Hathaway, P. L. (2007). The Intersection of Narrative Policy Analysis and Policy Change Theory. *Policy Studies Journal, 35* (1), 87–108.

McBeth, M. K., Shanahan, E. A., and Jones, M. D. (2005). The Science of Storytelling: Measuring Policy Beliefs in Greater Yellowstone. *Society and Natural Resources, 18*(5), 413–429.

McBeth, M. K., Shanahan, E. A., Hathaway, P. L., Tigert, L. E., and Sampson, L. J. (2010). Buffalo Tales: Interest Group Policy Stories in Greater Yellowstone. *Policy Sciences, 43*(4), 391–409.

McGough, D. J., Bedell, C., and Tinkler, B. (2018). Building a dangerous outpost in the Green mountain state: A case study of educator preparation policymaking. *Education Policy Analysis Archives, 26*, 37. doi:10.14507/epaa.26.2848

McMorris, C., Zanocco, C., and Jones, M. (2018). Policy narratives and POLICY OUTCOMES: An NPF examination of Oregon's ballot MEASURE 97. *Policy Studies Journal, 46*(4), 771–797. doi:10.1111/psj.12263

Merry, M. (2016). Constructing Policy Narratives in 140 Characters or Less: The Case of Gun Policy Organizations. *Policy Studies Journal, 44*(4), 373–395.

Merry, M. (2018). Narrative Strategies in the Gun Policy Debate: Exploring Proximity and Social Construction. *Policy Studies Journal, 46*(4), 747–70.

Merry, M. (2019). Angels versus Devils: The Portrayal of Characters in the Gun Policy Debate. *Policy Studies Journal 47*(4): 882–904.

Merry, M. (2020). *Warped Narratives: Distortion in the Framing of Gun Policy.* University of Michigan Press: Ann Arbor.

Mosley, J. E., and Gibson, K. (2017). Strategic use of evidence in state-level policy-making: Matching evidence type to legislative stage. *Policy Sciences, 50*(4), 697–719. doi:10.1007/s11077-017-9289-x

Ney, S. (2014). The Governance of Social Innovation: Connecting Meso and Macro Levels of Analysis. In *The science of stories: Applications of the narrative policy framework in public policy analysis*. New York, NY: Palgrave Macmillan.

Nisbett, N. (2017). A narrative analysis of the political Economy shaping policy on Child Undernutrition in India. *Development and Change, 48*(2), 312–338. doi:10.1111/dech.12297

O'Donovan, K. T. (2018). Does the Narrative policy framework apply to local Policy Issues? *Politics & Policy, 46*(4), 532–570. doi:10.1111/polp.12265

Olofsson, K. L., Weible, C. M., Heikkila, T., and Martel, J. C. (2018). Using non-profit narratives and news media framing to Depict air pollution in Delhi, India. *Environmental Communication, 12*(7), 956–972. doi:10.1080/17524032.2017.1309442

O'Leary, R., Borland, R., Stockwell, T., and MacDonald, M. (2017). Claims in vapour device (e-cigarette) regulation: a narrative policy framework analysis. *International Journal of Drug Policy, 44*, 31–40.

Peltomaa, J. (2018). Drumming the barrels of hope? Bioeconomy narratives in the media. *Sustainability 10(11)*: 4278. 10.3390/su10114278

Peltomaa, J., Hildén, M., and Huttunen, S. (2016). Translating institutional change - forest journals as diverse policy actors. *Forest Policy and Economics, 70*, 172–180. doi:10.1016/j.forpol.2016.06.029

Peterson, H. L. (2018). Political information has bright colors: Narrative attention theory. *Policy Studies Journal, 46*(4), 828–842. doi:10.1111/psj.12272

Peterson, H. L. (2019). Macro Stories: Policy Process Dynamics of Presidential Environmental Ideas. Unpublished doctoral dissertation. Oregon State University.

Price, M. H. (2019). Strategic stories: Analysis of prior learning assessment policy narratives. *The Review of Higher Education, 42*(2), 511–535. doi:10.1353/rhe. 2019.0005

Radaelli, C. M., Dunlop, C. A., and Fritsch, O. (2013). Narrating Impact Assessments in the European Union. *European Political Science, 4*(3), 31–54.

Raile, E. D., Shanahan, E. A., Ready, R. C., McEvoy, J., Izurieta C., Reinhold, A. M., Poole, G. C., Bergmann, N. T., and King, H. (2022). Narrative Risk Communication as a Lingua Franca for Environmental Hazard Preparation. Environmental *Communication 16(1)*: 108–124. DOI: 10.1080/17524032.2021.1966818

Schlaufer, C. (2018). The narrative uses of evidence. *Policy Studies Journal, 46*(1), 90–118. doi:10.1111/psj.12174

Schwartz, N. S. (2019). Called to arms: The NRA, the gun culture & women. *Critical Policy Studies,* 1–16. doi:10.1080/19460171.2019.1697892

Shanahan, E. A., Adams, S. M., Jones, M. D., and McBeth, M. K. (2014). "The Blame Game: Narrative Persuasiveness of the Intentional Causal Mechanism." Michael D. Jones, Elizabeth A. Shanahan, and Mark K. McBeth, eds. *The Science of Stories: Applications of Narrative Policy Framework* (pp. 69–88). New York, New York: Palgrave Macmillan.

Shanahan, E. A., Jones, M. D., and McBeth. M. K. (2018). How to Conduct a Narrative Policy Framework Study. *The Social Science Journal*, *55*(3), 332–345.

Shanahan, E. A., Jones, M. D., McBeth, M. K., and Lane, R. (2013). An Angel on the Wind: How Heroic Policy Narrative Shape Policy Realities. *Policy Studies Journal, 41* (3), 453–483.

Shanahan, E. A., Jones, M. D., McBeth. M. K., and Radaelli, C. M. (2018). "The Narrative Policy Framework." Chapter 5 (173–214) in *Theories of the Policy Process*, 4th edition. Edited by Chris M. Weible and Paul A. Sabatier. Boulder: Westview Press.

Shanahan, E. A., McBeth, M. K., and Hathaway, P. L. (2011). Narrative Policy Framework: The Influence of Media Policy Narratives on Public Opinion. *Policy & Politics, 39* (3), 373–400.

Shanahan, E. A., McBeth, M. K., Hathaway, P. L., and Arnell, R. J. (2008). Conduit or Contributor? The Role of Media in Policy Change Theory. *Policy Sciences, 41*(2), 115–138.

Shanahan, E. A., Raile, E. D., French, K. A., and McEvoy, J. (2018). Bounded stories. *Policy Studies Journal, 46*(4), 922–948.

Sievers, T., and Jones, M. D. (2020). Can power be made an empirically viable concept in policy process theory? Exploring the power potential of the narrative policy framework. *International Review of Public Policy, 2*(1), 90–114. doi:10.4000/irpp.942

Smith-Walter, A. (2018). Victims of Health-Care Reform: Hirschman's Rhetoric of Reaction in the Shadow of Federalism. *Policy Studies Journal, 46*(4), 894–921.

Smith-Walter, A., Jones, M. D., Shanahan, E. A., and Peterson, H. L. (2019, May 17). The stories groups tell: Campaign finance reform and the narrative networks of cultural cognition. Retrieved February 14, 2021, from https://link.springer.com/article/10.1007%2Fs11135-019-00884-8

Smith-Walter, A., Peterson, H. L., Jones, M. D., and Reynolds Marshall, A. N. (2016). Gun stories: How evidence shapes firearm policy in the United States. *Politics & Policy, 44*(6), 1053–1088.

Stephan, H. R. (2020). Shaping the scope of conflict in Scotland's fracking debate: Conflict management and the narrative policy framework. *Review of Policy Research, 37*(1), 64–91. doi:10.1111/ropr.12365

Veselková, M. (2014). Science, stories and the anti-vaccination movement. *Human Affairs, 24*(3). doi:10.2478/s13374-014-0227-8

Weible, C. M., Olofsson, K. L., Costie, D. P., Katz, J. M., and Heikkila, T. (2016). Enhancing Precision and Clarity in the Study of Policy Narratives: An Analysis of Climate and Air Issues in Delhi, India. *Review of Policy Research, 33*(4), 420–441.

Weiss, J. (2018). The evolution of reform Narratives: A narrative policy framework analysis of GERMAN NPM reforms. *Critical Policy Studies, 14*(1), 106–123. doi:10.1080/19460171.2018.1530605

Zanocco, C., Song, G., and Jones, M. (2018). Fracking bad guys: The role of narrative character affect in shaping hydraulic fracturing policy preferences. *Policy Studies Journal, 46*(4), 978–999. doi:10.1111/psj.12278

7 Innovation and Diffusion

Connecting Theory and Method

Andrew Karch

Introduction

The expansive scholarly literature on innovation and diffusion reaches across numerous disciplines and virtually every subfield of political science (Rogers 1995; Graham, Shipan, and Volden 2013). It is sometimes linked to Galton's problem, which explains that the widespread adoption of a specific form, institution, or policy cannot automatically be attributed to functional exigencies. Instead, this outcome might result from *interdependencies* among the adopters who influence, and are influenced by, one another. Scholars of international relations, comparative politics, and American politics use several distinct terms to refer to interjurisdictional interactions, such as policy transfer, harmonization, and convergence. These differences in terminology mask a shared interest in similar research questions, mechanisms, and concepts. The diverse strands of the diffusion literature differ in important ways,[1] but this chapter will treat them as complementary and emphasize areas of overlap.

The possibility of diffusion poses a fundamental challenge to correlational analysis, which requires that each case be independent of every other case. In contrast, diffusion scholars posit that interdependencies among units are part and parcel of the policy process. Can change be attributed to the impact of forces within the jurisdiction, to the influence of other units within the political or social system, or to some combination of the two? Teasing out the respective impact of internal and external forces is both a conceptual and an empirical challenge (Volden, Ting, and Carpenter 2008).

As scholars gravitated toward the "unified model of government innovation" (Berry and Berry 2018), they sought a quantitative approach that could acknowledge and identify a range of potential influences on the adoption decision. The introduction of event history analysis (EHA), which can accommodate internal and external factors in a single model (Berry and Berry 1990, 1992), sparked remarkable growth in published diffusion research. EHA and related approaches, such as pooled EHA (PEHA) and dyadic EHA, continue to be used widely. The mechanics of these approaches have been described in detail elsewhere (Box-Steffensmeier and Jones 1997, 2004; Volden 2006; Boehmke 2009a, 2009b; Kreitzer and Boehmke 2016; Buckley and Westerland 2004; Mooney 2020; LaCombe and Boehmke

2020). The first section of this chapter will provide a brief overview of these quantitative methods and discuss their respective strengths and limitations; it will also describe inferred network analysis, a relatively new approach to studying the diffusion process (Desmarais, Harden, and Boehmke 2015).

The second section of this chapter will focus on the methodological implications of recent conceptual innovations in the study of policy diffusion. Most recent research does not simply ask whether diffusion occurred. It also asks *why* these interdependencies exist, focusing on diffusion "mechanisms" like learning, imitation, and competition (Shipan and Volden 2008; Maggetti and Gilardi 2016). This constructive turn has produced a better understanding of the conditions under which developments in one jurisdiction might be affected by what has happened elsewhere, but it poses significant methodological challenges. Identifying the presence of a mechanism "requires careful thinking about how theoretical concepts map onto measures and which methods provide the most appropriate features for estimating them" (LaCombe and Boehmke 2020, 324). Scholars disagree on the best way to measure specific mechanisms. Even when they use the same measures, they sometimes disagree on how to interpret them. Resolving these disagreements will be essential to further conceptual development. In the words of Berry and Berry (2018, 279–280), "Developing methods for detecting the presence of specific mechanisms—thereby allowing our tools for empirical analysis to catch up with advances in theory—is the most important task on the agenda of scholars of policy innovation and diffusion." Multimethod research designs that incorporate quantitative and qualitative evidence offer one promising way to assess whether and how specific mechanisms influence the diffusion process.

The third section of this chapter examines forces that operate in multiple jurisdictions and transport innovations from one site to another. These "innovation vectors" may be responsible for the interdependencies that represent a diffusion episode. They include policy entrepreneurs, professional associations and interest groups, and mass media (Skocpol et al. 1993; Hays and Glick 1997; Mintrom 1997; Balla 2001; Boushey 2010; Finger 2018; Yi, Berry, and Chen 2018; Bromley-Trujillo and Karch 2021; Yu, Jennings, and Butler 2020). Canonical theories of the policy process—policy feedback theory, punctuated equilibrium theory, the advocacy coalition framework, and others—highlight the potential influence of these political forces, even if they do not refer to them as vectors. Identifying their impact poses analytical challenges; responding to those challenges will enable diffusion scholars to develop a more comprehensive understanding of the policy process and of interdependencies among jurisdictions.

The fourth section of this chapter highlights a long-standing limitation of policy diffusion research, namely, its focus on the adoption decision. While this emphasis has contributed to many conceptual and methodological advancements, it also leaves several important questions unasked. The result is a scholarly literature that mostly emphasizes widely adopted innovations at the expense of diffusion "failures" (but see Myers 2018) fails to consider

how the existence of a policy in one jurisdiction might make its adoption elsewhere less likely and does not acknowledge how innovations vary across the jurisdictions in which they exist. These limitations are significant because they represent additional ways that external forces can foster interdependencies. Thus, they are also missed opportunities for theory development and the advancement of hypotheses about the conditions under which specific mechanisms and vectors are likely to be impactful.

In sum, there is a tight relationship between theory and method in the study of innovation and diffusion. The interdependencies at the heart of a diffusion episode might even be described as a methodological challenge to be resolved. Over the years, political scientists have developed several ways to address this challenge while simultaneously embracing conceptual advancements that explain why developments in one jurisdiction might affect what happens elsewhere. These conceptual advancements, particularly the heightened focus on diffusion mechanisms, raise their own set of measurement and other challenges. This chapter will conclude by identifying several avenues for future research that will help this expansive scholarly literature maintain its forward momentum.

Common Quantitative Approaches

In his seminal article on the diffusion of policy innovations among the American states, Jack Walker (1969, 881–882) tells the story of the California fair trade law of 1931. Ten of the 20 states that enacted their own versions of the law copied "two serious typographical errors" that appeared in the California statute. These errors provide incontrovertible evidence of a diffusion episode in which legislation was affected by developments in other jurisdictions. This sort of evidence is exceedingly rare, however. It is far more common for diffusion scholars to simply have a list of jurisdictions that adopted the same innovation. Their analytical challenge is to ascertain whether this overlap results from common needs, interjurisdictional interactions, or some combination of the two. Standard correlational models are not well suited to this task because they require that each case be independent of every other case. In contrast, the concept of diffusion highlights the interactions and prioritizes the interdependencies that conventional quantitative methods largely assume away (Mooney 2020).

Timing and sequence are central to theories of diffusion, which poses another analytical challenge. The fact that a jurisdiction adopted an innovation that already existed elsewhere does not definitively demonstrate that diffusion occurred. The second adoption could be motivated by internal factors rather than external pressures. Parametric survival models are ideal for analyses that have a temporal component because they account for time, rather than controlling for it as a nuisance (Box-Steffensmeier and Jones 1997, 2004). These models take various forms, including discrete time models that employ logistic regression and Cox proportional hazard models. They employ exponential distributions and Weibull distributions. In addition,

diffusion scholars use different strategies to account for duration dependence and the passage of time, including annual dummy variables, nonlinear trend variables, and cubic splines (Beck, Katz, and Tucker 1998; Mooney 2001; Buckley and Westerland 2004).

Since the early 1990s, EHA has been the dominant quantitative approach in diffusion research. It responds reasonably effectively to the above concerns, thereby enabling scholars to examine the impact of external forces—such as the existence of an innovation in a neighboring jurisdiction—while controlling for the jurisdiction's political and economic environment and accommodating longitudinal variation in these factors. Early applications of EHA to diffusion research sparked a flurry of studies relying on the technique, especially in the context of the American states (Pavalko 1989; Berry and Berry 1990, 1992).

Initially, most EHA-based diffusion studies examined a single innovation and evaluated whether external factors were influential in that specific case. The single-innovation approach suffers from two methodological drawbacks that Christopher Z. Mooney (2020, 52) labels "sample bias and heterogeneous processes." Sample bias implies that the trajectory of a single innovation may not be representative of all innovations or even of a broader class of innovations. Drawing inferences or making broad generalizations based on a single case is therefore inappropriate, especially since diffusion is a heterogeneous process that occurs for multiple reasons. The recent turn to diffusion mechanisms, described in the next section, highlights this heterogeneity, but it has long been part of the scholarly literature (Gray 1973).

PEHA, which allows scholars to assess many innovations simultaneously, emerged partly in response to concerns about sample bias. It "allows researchers to study the effects of variables across multiple [innovations] by stacking the data and estimating the parameters in a single model" (Boehmke and Kreitzer 2016, 121). Scholars of American politics have applied PEHA to numerous policy domains (Shipan and Volden 2006, 2008; Makse and Volden 2011; Kreitzer 2015; Boushey 2016), and it also has been used to examine specific policy tools like interstate compacts (Karch et al. 2016). PEHA begins to address concerns about sample bias even though it cannot eliminate it; there is no guarantee that any set of innovations, no matter how large, is representative of the entire universe. The downside of the data-stacking process, however, is that it effectively assumes that the impact of a given variable is constant across innovations. This assumption violates the notion that diffusion is a heterogeneous process that is associated with different mechanisms and vectors.[2]

Another common response to the issue of sample bias has been to collect adoption data for ever larger samples of innovations. In the context of the American states, the State Policy Innovation and Diffusion (SPID) database is the most ambitious of these efforts (Boehmke et al. 2020). It contains adoption data for 728 policy innovations that are coded by topic area. Similarly, the policy diffusion results (PDR) database attempts to facilitate the drawing of generalizations by compiling the results from every EHA

model of policy diffusion among the American states that was published from 1990 to 2018 (Mallinson 2020). It is well suited to meta-analyses of specific factors that appear frequently in existing research, both internal correlates of adoption and proxies for interdependencies.

In addition to drawing on larger databases of innovations, diffusion scholars increasingly embrace another quantitative approach—dyadic EHA—that responds directly to the conceptual issues at the heart of the scholarly literature. Conventional EHA models are monadic, examining adoption patterns in an individual jurisdiction in the context of what all the other jurisdictions in the system have done. However, diffusion is a relational process in which officials are drawn to specific models. By using the dyad-year as the unit of analysis (Volden 2006; Nicholson-Crotty and Carley 2016; Yu, Jennings, and Butler 2020) and limiting observations to instances in which there is a possible interdependency (Boehmke 2009b), dyadic EHA can assess the impact of specific mechanisms and vectors. However, it imposes a trade-off because it requires precise theorizing and measurement of interjurisdictional interactions. As a result, dyadic EHA often appears in studies of a single innovation or policy domain, raising questions about sample bias and external validity.

Network inference provides another way to model the transmission of information and ideas in the diffusion process (Desmarais, Boehmke, and Harden 2015). It identifies patterns based on the timing of adoption across many innovations, assessing whether jurisdictions repeatedly follow specific peers and are less likely to adopt if those peers have not previously done so. Network inference invokes conventional portrayals of the diffusion process, presuming that there are "persistent policy pathways" that link jurisdictions to one another. It is reminiscent of the leader-laggard diffusion model, which posits that certain jurisdictions tend to be pioneers and influence others' adoption decisions. The main weakness of network inference, like the leader-laggard model, is its failure to "identify a priori (1) the jurisdictions (or even types of jurisdictions) that are expected to be pioneers, and (2) the predicted order of adoption of the governments expected to follow" (Berry and Berry 2018, 265). It is an intriguing approach that has highlighted non-geographical diffusion patterns, but it will only generate major conceptual breakthroughs if it is combined with theorizing about the types of networks that are likely to facilitate specific diffusion mechanisms (LaCombe and Boehmke 2020; Mooney 2020).

Before turning to those mechanisms, it is important to acknowledge a shared limitation of all the quantitative methods described thus far. Ultimately, the study of diffusion is about individual behavior. It is an effort to identify instances when external developments influence the decisions made by policymakers. EHA, PEHA, dyadic EHA, and network inference differ in important ways, yet they rely on jurisdiction-level data. Individual decision-makers rarely appear in these models, which implicitly rely on the unitary actor assumption. The turn toward mechanisms is important, in part, because it returns these individuals to a central place in diffusion theorizing.

Assessing the Impact of Diffusion Mechanisms

Why does diffusion occur? This is a deceptively straightforward question. The most significant recent conceptual advance in diffusion research is a focus on several "mechanisms" that explain why officials take cues from what happened elsewhere. Institutional and policy innovations can diffuse for multiple reasons, and the turn toward mechanisms reflects a growing scholarly interest in distinguishing among these distinct external forces. The rest of this section describes different mechanisms that can drive a diffusion episode and the methodological challenge of identifying and operationalizing them.

As Christopher Z. Mooney (2020, 5) explains in a recent review essay, diffusion invokes the same underlying logic as incrementalism theory. Facing innumerable demands on their limited time, public officials cannot perform a comprehensive search for policy-relevant information. Instead, they gravitate toward timely and accessible sources, "muddling through" until they feel sufficiently well informed to make a specific decision (Lindblom 1959; see also Sabatier and Whiteman 1985; Mooney 1991; Karch 2007a). In other words, lawmakers rely on cognitive and informational shortcuts, one of which is turning to the models provided by other jurisdictions (Weyland 2006). These external models can provide programmatic information about whether an innovation achieved its goals and political information about how key constituencies responded.

There is no shortage of external models on which decision-makers can draw, but they are unlikely to be influenced equally by all of them. Thus, it is incumbent on scholars to explain why officials might gravitate toward certain exemplars rather than others and then to assess those possibilities empirically. Geographic proximity is a common theme in diffusion research. Many studies assume that developments in nearby jurisdictions provide models on which officials can draw and that these models make the adoption of an innovation more likely. A common statistical proxy for this effect is the percentage of a jurisdiction's neighbors in which an innovation has been enacted prior to the year of measurement (Mallinson 2020).

Geographic proximity can generate interdependencies in multiple ways. It can facilitate the development of communications networks through which information travels among decision-makers (Crain 1966; Foster 1978). Similarly, overlapping media markets can alert residents and public officials to the existence of innovations in nearby jurisdictions (Pacheco 2012; Mitchell 2016). Officials might also be drawn to models in nearby jurisdictions because they believe that their neighbors are culturally and demographically similar. Finally, geographic proximity can generate economic competition, in which the existence of an innovation in a nearby jurisdiction generates positive or negative externalities that lead officials to react accordingly (Boehmke and Witmer 2004; Baybeck, Berry, and Siegel 2011).

The standard proxy for geographic proximity cannot distinguish among these possibilities, showcasing yet again the importance of careful theorizing

about *why* developments in neighboring jurisdictions might be influential. Such theorizing can foster the development of nuanced measures that better reflect interjurisdictional interdependencies. Based on the international dynamics of human trafficking, for example, a recent study weighed policies in neighboring jurisdictions by the roads connecting them (Simmons, Lloyd, and Stewart 2021). This specificity is especially important because recent research finds that the impact of contiguous neighbors has declined over time, at least among the American states (Mallinson 2021).[3]

Given the theoretical and empirical shortcomings of the standard emphasis on geographic proximity, one of the most promising recent developments in the field is the growing attention paid to diffusion mechanisms like learning, competition, and imitation (Shipan and Volden 2008).[4] These mechanisms offer a precise explanation of the interdependencies that lie at the heart of a diffusion episode. However, incorporating them into empirical analyses poses a challenge.

The learning mechanism implies that officials emulate another jurisdiction because their counterparts adopted an innovation that achieved its goals. "Success" can take multiple forms. The existence of an innovation in another jurisdiction can give lawmakers more confidence in its "political palatability, social acceptability, and technical feasibility" (Pacheco-Vega 2021, 392). They might also be interested in its programmatic impact, evincing greater willingness to emulate innovations that are associated with desirable outcomes (Volden 2006; Yu, Jennings, and Butler 2020). Learning resonates with the notion that the American states function as "laboratories of democracy" in which policymakers can experiment with novel approaches to societal issues, but it can occur in virtually any setting. Importantly, the learning mechanism divorces the diffusion process from geographic proximity. Officials who want to address a specific problem presumably would be interested in innovations that seem to combat it successfully, regardless of where those external models are located.

Conducting an empirical assessment of the learning process is an extraordinarily daunting task. Innovations are inherently multidimensional, making it difficult to evaluate them on objective and widely agreed upon criteria (Nathan 2000; Marsh and Sharman 2009). Moreover, officials can change the criteria they use to judge success or adjust their aspirations with respect to these goals (Levitt and March 1988).[5] These general challenges are exacerbated in specific arenas such as environmental policy, where ecological phenomena are associated with an "extremely high degree of variability and uncertainty" (Pacheco-Vega 2021, 397). Many empirical assessments of the learning mechanism therefore employ process tracing and focus on a single innovation or a few closely related provisions that pursue the same objectives (Osorio Gonnet 2019). Quantitative analyses of learning tend to use dyadic EHA because it compares where jurisdictions stand vis-à-vis specific indicators (Nicholson-Crotty and Carley 2016; Yu, Jennings, and Butler 2020). Multimethod approaches can be especially revealing; in an excellent

study of the Children's Health Insurance Program, for example, Volden (2006) combines dyadic EHA with interviews of program administrators who were asked to identify their goals.

Other quantitative studies use a different proxy for the learning mechanism. Hypothesizing that the existence of an innovation in an increasing number of jurisdictions provides lawmakers with the "opportunity to learn," they use a count of the number of adoptions that occurred prior to the year of measurement (Shipan and Volden 2008; Makse and Volden 2011; Karch et al. 2016). Although this measure is simple to calculate and easy to apply to large databases of innovations, it cannot directly link learning to a specific goal or its successful pursuit. On its own, it cannot indicate whether officials adopted an innovation because it succeeded, because it was trendy, or because doing so conveyed legitimacy. Indeed, many sociologists employ a count of previous adopters to connote "normative pressures" (Maggetti and Gilardi 2016, 94). These conflicting interpretations suggest that attempts to identify learning through quantitative indicators will be more compelling if such "pattern-finding" efforts are combined with careful process tracing and case studies (Marsh and Sharman 2009).

Imitation—another diffusion mechanism—poses similar analytical challenges. During a diffusion episode driven by imitation, officials are motivated by the belief that their jurisdiction shares a policy-relevant attribute with another jurisdiction in which the innovation exists. The perceived connections can be based on political, demographic, or budgetary similarities. Political similarities seem especially important for scholars examining diffusion among the American states in an era of intense partisan polarization. Ideological cues can serve as an informational shortcut, enabling officials to "minimize the uncertainty about how issues fit in the liberal–conservative policy space" (Grossback, Nicholson-Crotty, and Peterson 2004, 521). In sum, shared ideological or partisan ties can generate interjurisdictional interdependencies.

Other types of perceived similarities can be responsible for an imitation-based diffusion episode. For example, Volden (2006, 310) finds that "similarities based on ideological leanings, per capita income, managed care structures, and budgetary considerations were all relevant to the diffusion process." The range of these factors is noteworthy. It suggests that specific features of a policy domain can affect the external examples toward which officials are drawn. Furthermore, imitation does not only affect the adoption decision. It can also influence decisions about program implementation and modification, perhaps because officials believe that certain environmental conditions affect the likelihood of programmatic success (Nicholson-Crotty and Carley 2016). The logic of the imitation mechanism is not novel; it resonates with the long-standing assumption that nearby jurisdictions share relevant attributes (Foster 1978).

Hypotheses about the impact of general jurisdictional characteristics, including ideological and budgetary similarities, are amenable to various research designs. They can be incorporated into quantitative studies of

individual innovations and pooled analyses, and they have appeared in studies relying on standard EHA techniques and the dyadic approach. Hypotheses about imitation that relate to specific features of an innovation or to its environment, in contrast, may render pooled analyses inappropriate. Regardless of which empirical approach they adopt, diffusion scholars with an interest in imitation must specify the conditions under which it is likely to occur and why.

The final diffusion mechanism featured in this section—competition— also raises several important conceptual and empirical issues. Competition drives diffusion when the existence of an innovation in one jurisdiction imposes positive or negative externalities. Positive externalities have received limited attention, but they give non-adopters a competitive advantage and an incentive to leave the status quo in place. In contrast, negative externalities can lead officials to believe that the failure to adopt the innovation will put their jurisdiction at a competitive disadvantage. Scholars hypothesize that this dynamic is especially common in economic development and tax policy, where jurisdictions compete for private investment because of its potential impact on job growth and tax revenue (Berry and Berry 1992; Arel-Bundock and Parinandi 2018). In other contexts, such as human trafficking, laws are designed to repel certain business activities and population movements that are viewed as undesirable (Simmons, Lloyd, and Stewart 2021). The mobility of firms and individuals can lead officials in specific jurisdictions to feel competitive pressures to act, and these pressures are an important form of interdependence.[6]

Interjurisdictional competition provides a compelling explanation of why diffusion occurs but identifying competition and its influence is an empirical challenge. The nature of that challenge depends on the political context. Scholars investigating diffusion among the American states often use measures that emphasize potential competition among neighboring jurisdictions (Boehmke and Witmer 2004; Arel-Bundock and Parinandi 2018). This geographic focus is justifiable in certain contexts. For example, the ease with which residents of one state could travel to purchase lottery tickets seemed to facilitate the spread of that policy innovation (Baybeck, Berry, and Siegel 2011). In other contexts, officials will compete with any jurisdiction that wants to attract private investment or become the recognized leader in a particular policy domain. As a result, diffusion scholars "must theorize carefully both about the conditions under which competition is likely to occur and about the type of competition that is likely to emerge" (Karch 2007b, 63).

In sum, competition-based accounts of diffusion face two empirical challenges. The first involves linking conceptual development to case selection. If competition is relevant in some contexts but not others and takes distinct forms, then empirical strategies based on massive databases may be less illuminating than studies that examine smaller sets of carefully chosen innovations. "Placebo tests" where competition is not expected to be influential can therefore be revealing (Gilardi 2015, 2016; Simmons, Lloyd, and Stewart 2021). Interjurisdictional competition will contribute to some

diffusion episodes but not others; a better understanding of its conditional impact will be a major conceptual and empirical advance.

Measurement poses a second empirical challenge for the analysis of competition-driven diffusion. What does competition look like if it is not based on geographic proximity? Consider the bidding war that ensued when Amazon requested proposals for a massive project that would effectively represent the company's "second" headquarters. It received bids from more than 230 locations across North America. This episode highlights how interjurisdictional competition can be "national" or even international in scope. Diffusion scholars, particularly those interested in economic development, might find it useful to incorporate insights from research in international political economy. Measures of trade flows or sectoral overlap might serve as proxies for non-geographic competition in various political contexts.

Why are jurisdictions interdependent? Why does diffusion occur? For too long, the most common empirical strategy for assessing the potential impact of external factors was to include a proxy for developments in nearby jurisdictions. The move away from geography and toward mechanisms such as learning, imitation, and competition is a very promising trend even though it generates the empirical challenges described in this section. In addition to acknowledging those challenges, it is important to acknowledge that mechanism-based accounts raise many conceptual questions. Diffusion mechanisms might interact, and their influence can vary across a diffusion episode or in different jurisdictions; these complexities often become apparent in detailed case studies based on fieldwork and interviews (Marsh and Sharman 2009; Osorio Gonnet 2019). How do learning and imitation occur? Where do officials turn for reliable information about whether an innovation has succeeded? How do lawmakers find out what their counterparts in ideologically or otherwise similar jurisdictions are doing? The next section of this chapter turns to several potential "vectors" that can facilitate the transfer of relevant information.

Assessing the Impact of Diffusion Vectors

Many political actors operate in multiple jurisdictions, and their geographic reach means that they can facilitate the interdependencies that are the core of a diffusion episode. These actors include policy entrepreneurs, professional associations and interest groups, and mass media; all of them are potential "vectors" that provide officials with programmatic and political information about external developments (Boushey 2010; Mooney 2020). These actors play a central role in several canonical theories of the policy process, but quantitative studies of the diffusion process often overlook their potential impact. This gap is partly due to the methodological challenge of assessing vectors' influence. It is not easy to design compelling measures that can be incorporated into empirical analyses treating the jurisdiction-year or jurisdiction-year-policy as the unit of analysis. The intellectual payoff of doing so, however, can be very high.

Consider, for example, the role of policy entrepreneurs in transmitting information across jurisdictional boundaries. These individuals include elected officials, bureaucrats, representatives of the business community, and activists (Keck and Sikkink 1998; Mintrom 2000). These "policy ambassadors" promote innovations at the local, national, and transnational levels, working in both governmental and nongovernmental settings (Porto de Oliveira 2020). They are often embedded in intellectual, professional, or advocacy communities that give them knowledge about the details and effects of innovations in other jurisdictions (Walker 1981; Kirst, Meister, and Rowley 1984). Moreover, they themselves can be professionally mobile "boundary spanners" who transfer external information and knowledge when they move to new agencies or different jurisdictions (Yi and Chen 2019). Interorganizational and interjurisdictional personnel flows are an impactful diffusion vector in settings as diverse as the American states and Chinese provinces (Teodoro 2009; Yi, Berry, and Chen 2018).

Acknowledging and assessing the potential impact of individual entrepreneurs responds to a long-standing methodological critique of diffusion scholarship. If the models provided by other jurisdictions represent a cognitive and informational shortcut for decision-makers (Weyland 2006; Mooney 2020), then theories of diffusion are an attempt to explain individual behavior. However, most quantitative studies use aggregated measures and units of analysis. They treat jurisdictions as unitary actors, examining the adoption decisions of countries, states, cities, or other geographic units. In doing so, they largely exclude the legislators, executive branch leaders, bureaucrats, and other decision-makers who must gather information and decide whether to follow the precedent set by other jurisdictions. In contrast, analyses of policy transfer in comparative politics tend to center the role of individual agents, typically employing process tracing based on field research and ethnography to focus on the "micro-dynamics of the diffusion process" (Porto de Oliveira 2019). These complementary literatures generally do not engage with one another; multimethod research designs would illuminate the interactive and iterative relationship between structure and agency (Marsh and Sharman 2009).

Constructing valid quantitative measures of entrepreneurial activity is inherently difficult. More than two decades after its publication, Michael Mintrom's (1997, 2000) study of policy entrepreneurs and school choice in the American states remains the gold standard. It used a national survey to construct its central indicator, and it generated new insights about what entrepreneurs do and the conditions under which their actions lead to policy change. Entrepreneurs nevertheless continue to play a peripheral role in quantitative diffusion research, partly because a survey-based approach is not feasible for scholars who rely on large databases of innovations. Even scholars who use EHA to study a single innovation find it challenging to construct measures that identify entrepreneurial activities at specific moments in time.

In contrast, recent methodological advances have propelled growing attention toward the potential influence of professional associations, interest

groups, international organizations, and the like. In addition to possessing the geographic reach necessary to facilitate diffusion, many of them view information transmission as a central goal. They publish reports and white papers, maintain websites that serve as information repositories, and host conferences and informal meetings that facilitate the development of professional networks. International organizations like the World Bank and the United Nations, large management consultancy firms, and philanthropic organizations have endorsed innovations and taken steps to encourage their transnational spread (Dolowitz and Marsh 2000; Osorio Gonnet 2019; Stone, Porto de Oliveira, and Pal 2020). The information they disseminate facilitates interjurisdictional interdependencies by making decision-makers aware of external developments. Similarly, in the United States, the diffusion of innovations has been linked to interest group campaigns (Skocpol et al. 1993; Haider-Markel 2001) and leadership roles in professional associations (Balla 2001; McNeal et al. 2003).

Model legislation is one of many tools that professional associations and interest groups have at their disposal, and it has recently drawn considerable scholarly attention for theoretical and methodological reasons. In the words of Dolowitz and Marsh (1996, 351), "The easiest way to prove that copying has occurred is to examine the wording of the legislative bill authorizing a program." Public officials may rely on examples of statutory language that are either developed or endorsed by organizations. Empirically, computational text analysis provides a way to assess the amount of overlap between two statutes or between a statute and model legislation.[7] It encompasses many different techniques, all of which produce a "similarity score" for two texts. It has therefore become an increasingly popular tool for diffusion scholars who want to evaluate the influence of interest groups in the American states (Garrett and Jansa 2015; Jansa, Hansen, and Gray 2018; Collingwood, El-Khatib, and Gonzalez O'Brien 2019; Hertel-Fernandez 2019).

Computational text analysis represents a promising tool for investigating diffusion. It illuminates the extent to which officials use the templates produced by professional associations and interest groups, think tanks, corporations, and other organizations. At this point, diffusion scholars have primarily used this method to identify cases where this vector has been influential. While understandable, this approach is incomplete in ways that highlight the connection between theory and method. Theoretically, this vector may be more influential under certain conditions. Decision-makers might be more inclined to rely on the information transmitted by professional associations when the innovation under consideration is complex or difficult to implement (McNeal et al. 2003). In contrast, organizations may exercise less influence when an innovation is noncomplex or when it is politically salient (Nicholson-Crotty 2009). In sum, computational text analysis must be married to careful theorizing about the conditions under which officials will be prone to rely on this informational shortcut.

Similarly, diffusion scholars must compare the influence of model legislation to officials' reliance on alternative sources of information. Just as

computational text analysis has been used to evaluate the impact of certain groups, it can also be used to assess the extent to which jurisdictions rely on each other during the policy process (Callaghan, Karch, and Kroeger 2020; Linder, Desmarais, Burgess, and Giraudy 2020). Before issuing definitive statements about the dominance of specific organizations, scholars must calculate a baseline against which to judge their impact. Comparing the influence of model legislation to that of other informational shortcuts is a promising avenue for future research, especially when combined with the sort of theorizing outlined in the previous paragraph. Interjurisdictional interdependencies are only possible if officials become aware of external developments. Various organizations view information dissemination as a key goal, and their potential to serve as diffusion vectors merits additional study.

The importance of information transmission highlights the potential role of mass media as a diffusion vector. By stimulating awareness of innovations and disseminating programmatic or political information about them, media coverage can spur interjurisdictional interdependencies. As has already been mentioned, overlapping media markets can inform residents and public officials about the existence of innovations in nearby jurisdictions (Mitchell 2016). Yet the impact of media coverage is not limited by geography. Influential channels of the international press, such as *The Economist*, facilitated the diffusion of Brazilian social policies like participatory budgeting and the family allowance program by extolling and endorsing them (Porto de Oliveira 2019).

While the potential impact of media coverage on the diffusion process has long been acknowledged, this vector is often invoked in a general and imprecise way. Many quantitative studies use article counts as a proxy for salience and an indicator of public concern. This approach has generated important insights. In the context of the American states, higher levels of national coverage in outlets like *The New York Times* have been linked both to a more rapid diffusion process and a higher likelihood that innovations will gain adoption (Lax and Phillips 2009, 2012; Nicholson-Crotty 2009; Winburn, Winburn, and Niemeyer 2014; Bromley-Trujillo and Poe 2020).

Conventional approaches to the study of media influence highlight two methodological challenges. Measures of national media coverage highlight the availability of information; in the words of a recent study, they are a proxy for the "broader informational environment" in which decision-makers operate (Bromley-Trujillo and Karch 2021, 1003). More precise measures, such as indicators of media coverage in specific jurisdictions, may be better. However, they raise issues of endogeneity. Differences in mass media coverage may help account for jurisdictions' divergent approaches to a specific innovation, but it is also plausible that they are a response to the political agenda (Shanahan et al. 2008; Lax and Phillips 2009). Is media coverage a cause or an effect? Scholars who rely on broad indicators must be circumspect in interpreting their results.

The second methodological challenge has appeared throughout this chapter. There is a fundamental disconnect between theories emphasizing the

cognitive and informational shortcuts used by individuals and the aggregated measures used in most studies. The standard quantitative approaches that treat the jurisdiction-year or jurisdiction-year-policy as the unit of analysis can help scholars build a circumstantial case that media coverage and other informational resources contributed to a diffusion episode. On their own, however, article counts do not demonstrate that decision-makers gathered information through specific media sources, let alone that the coverage significantly influenced their choices. Multimethod research designs can make this case more compelling. Supplementary content analysis can identify coverage that refers to developments in other jurisdictions, and interviews with key officials can facilitate a better understanding of whether and how they gather and utilize different informational resources (Karch 2007a). In a similar vein, experimental designs can illuminate the information gathering process by assessing whether ideological or other cues affect officials' willingness to consult specific sources (Butler et al. 2017; Zelizer 2018; Pereira 2021); they possess the advantage of explicitly examining the behavior of individuals who make the decisions in which scholars are interested.

Canonical theories of the policy process highlight the potential impact of all the diffusion vectors discussed in this section. Yet policy entrepreneurs, professional associations and interest groups, and mass media are less central to quantitative diffusion research than one might expect, possibly due to the methodological challenges outlined above. The innovative techniques that have been used to surmount these challenges—surveys, computational text analysis, and experimental designs—are often easier to execute in the context of a single innovation, which raises concerns about external validity. The possible intellectual payoff of giving these vectors a more prominent place in diffusion research nevertheless seems very high. All of them have the capacity to affect information generation and dissemination, which are central to a literature devoted to identifying the cognitive shortcuts that officials use in determining which innovations to consider and adopt.

Overlooked Forms of Interdependence

The interdependence at the heart of a diffusion episode can take a variety of forms. This is one of the key insights of the recent turn toward diffusion mechanisms like learning, imitation, and competition. Importantly, however, these interdependencies will not always lead to program adoption. An elected official might become aware of a policy that exists elsewhere, investigate it, and decide that it is not appropriate for their jurisdiction. In some cases, "a negative lesson may be drawn about how not to proceed" (Dolowitz and Marsh 1996, 344). Similarly, an innovative policy in one jurisdiction might generate positive externalities elsewhere, leading policymakers in those locations to prefer the status quo. Both possibilities meet the conventional standard that "policy diffusion occurs if the probability of adoption of a policy by one governmental jurisdiction is influenced by the policy choices of other governments in the system" (Berry and Berry 2018, 256).

However, they showcase how the existence of a policy in one jurisdiction might decrease its likelihood of being adopted elsewhere. Diffusion is not simply about an increasing number of adoptions. Instead, it is a "consequence of interdependence," which implies that scholars should be "interested more in the process than in the outcome" (Gilardi 2016, 9; see also Elkins and Simmons 2005; Karch 2007a). Understanding the diffusion process requires expanding the range of scholarly inquiry beyond questions of adoption and its timing.

Looking beyond the adoption decision offers considerable analytical potential. If a policy in another jurisdiction inspires a legislator to introduce a bill establishing the same policy, their decision showcases the impact of interdependence and increases the likelihood of adoption even if the bill is not enacted. Investigating whether new policies are given serious consideration can therefore illuminate whether and how external developments affect the policy process. Agenda setting has received less scholarly attention than the adoption decision, but existing research highlights the impact of various diffusion vectors—policy entrepreneurs, interest groups, and mass media—in moving items onto the governmental agenda (Mintrom 1997; Haider-Markel 2001; Bromley-Trujillo and Karch 2021). There is also a vertical component to these interdependences in the United States, with national developments shaping state political agendas (Roh and Haider-Markel 2003; Baumgartner, Gray, and Lowery 2009; Karch 2012).

Examinations of other stages of the policy process, such as implementation, can illuminate the influence (or lack thereof) of external forces (Gilardi 2016; Nicholson-Crotty and Carley 2016). A gap between adoption and implementation can highlight the impact of internal, as opposed to interjurisdictional imperatives. For scholars of transnational diffusion, it illustrates the necessity of thoroughly investigating the domestic context. In China, for example, information from international contacts was used very selectively as global regulatory standards against money laundering "were weakened or even neutralized through discretionary enforcement" (Heilmann and Schulte-Kulkmann 2011, 639). Economic interests facilitated the standards' adoption, whereas domestic political imperatives influenced their implementation. Administrative capacity can affect whether diffusion occurs; desirable innovations may not diffuse "if implementation is beyond a jurisdiction's technological abilities" (Dolowitz and Marsh 1996, 354). It can also affect how diffusion occurs, moderating the impact of certain mechanisms in specific jurisdictions (Osorio Gonnet 2019).

The prevailing focus on the adoption decision is understandable from a methodological perspective. Policy adoptions "represent easily observable, discrete events that are conducive to large-N quantitative analysis" (LaCombe and Boehmke 2020, 311). While contemporary scholars use a variety of research methods, their outcome of interest generally remains whether and when a jurisdiction adopted an innovation. The SPID database, with adoption-related information about 728 policies, illustrates this tendency (Boehmke et al. 2020). The large database approach has facilitated numerous

conceptual advances in the study of diffusion, but it comes with certain limitations. In addition to neglecting other stages of the policy process, it is not well suited to examine episodes of limited or "negative" diffusion and it cannot address differences in the "depth" or "extent" of innovation (Berry and Berry 2018). These limitations merit additional explanation.

Most diffusion studies gravitate toward policies for which there is sufficient variation in the "adopt or not" dependent variable to permit productive quantitative analysis. As a result, they focus almost exclusively on policies that were adopted by numerous jurisdictions. This tendency to examine widely adopted policies is sometimes called the "pro-innovation bias" (Karch et al. 2016).[8] For example, most efforts to construct state innovativeness scores include only policies that reach a specific adoption threshold, typically 20 states (Walker 1969; Savage 1978; Boushey 2010; Boehmke and Skinner 2012). Focusing on these policies can offer important insights into diffusion mechanisms, but it is less well suited to explain why some innovations are more popular than others. Nor can it explain instances of "negative" diffusion in which the existence of a policy in one jurisdiction makes it less likely that the same policy will be adopted elsewhere. For example, lawmakers might steer clear of an external model that proved technically infeasible or that sparked a political or social backlash.

Indeed, existing research rarely examines episodes of "failed" diffusion in which very few jurisdictions enact a novel innovation. During the early 1990s, for example, the state of Oregon implemented an innovative Medicaid reform program that effectively rationed the medical care services provided to program recipients. This experiment generated national interest, but not a single state adopted the Oregon model (Jacobs, Marmor, and Oberlander 1999). Studying failed or limited diffusion requires looking beyond the adoption decision. With limited adoptions to assess, a qualitative, process-oriented approach may be necessary. Case studies of jurisdictions in which the innovation was considered but not enacted can help identify the factors that stood in the way of its widespread adoption (Myers 2018; Shriver, Szabo, and Bray 2020). Quantitative studies can make a concerted effort to incorporate episodes of limited diffusion.[9] More inclusive databases can facilitate comparisons across innovations that achieve various thresholds of adoptions (Karch et al. 2016).

The "adopt or not" dependent variable used in most diffusion research cannot illuminate questions of program content. What, exactly, is being diffused from one jurisdiction to another? A policy template might gain widespread enactment but take diverse forms in the jurisdictions in which it is adopted (Grattet, Jenness, and Curry 1998; Mossberger 2000; Jones and Newburn 2002). In that sense, diffusion might be a "matter of degree" (Marsh and Sharman 2009, 278). Lawmakers can customize innovations for political, programmatic, or other reasons, and these initial differences can be extended by subsequent modifications, repeals, amendments, and reinstatements (Eyestone 1977; Karch and Cravens 2014).

Public policies typically have multiple components. Relying on dichotomous dependent variables that simply indicate whether a policy has been adopted therefore obscures differences in the expansiveness or comprehensiveness of existing programs (Hays 1996; Mooney and Lee 1995; Yu, Jennings, and Butler 2020). Interest groups, professional associations, policy entrepreneurs, and other political actors frequently view these details as critical. Thus, the conventional focus on the adoption decision overlooks central features of the policy process, thereby missing an opportunity to evaluate the potential impact of external actors. Using a process-tracing approach to investigate these "pathways of policy tinkering and adjustment" can help scholars identify which intermediaries are responsible for programmatic variation (Stone, Porto de Oliveira, and Pal 2020, 5). PEHA offers another potential path forward (Shipan and Volden 2006; Boehmke 2009a; Kreitzer and Boehmke 2016), not only in terms of empirical approach but also in terms of theory development. Scholars can stack data for multiple policy components rather than the adoption of multiple innovations.

In sum, this section has identified several lacunae in the scholarly literature on innovation and diffusion. As scholars address them, they should remain attuned to the conceptual advances described elsewhere in this chapter. Their focus should not be on establishing that diffusion took place or that a specific policy diffused. Instead, it should be on generating novel insights about diffusion mechanisms, diffusion vectors, and their interaction. Some interest groups contribute to "negative" diffusion by working to prevent policy change (Finger 2018). Interest groups can also affect the content of programs that gain enactment. Yu, Jennings, and Butler (2020) examine the link between lobbying by Mothers against Drunk Driving (MADD) and the comprehensiveness of drunk driving regulations, describing the group as a facilitator of policy learning. By examining specific policy provisions, this study highlights the connection between diffusion vectors and mechanisms. The more general point is that looking beyond adoption, examining episodes of "negative" or limited diffusion, and investigating program content all are ways to expand on recent theorizing about interdependencies among jurisdictions. This chapter will conclude by identifying additional constructive avenues for future research.

Moving Forward

The study of innovation and diffusion poses vexing analytical challenges, and scholars have responded to these challenges in myriad creative ways. This chapter has highlighted some of the most common approaches, focusing on their general strengths and limitations rather than their mechanics. It has also described various measurement issues that arise when scholars try to operationalize diffusion theories. Before concluding with suggested avenues for future research, it is important to acknowledge the existence of other research strategies.

Previous reviews of the diffusion literature have called for more qualitative and multimethod research designs (Gray 1994; Graham, Shipan, and Volden 2013). These designs are more common among scholars of comparative politics and policy and have produced additional insights about the impact of specific diffusion mechanisms and vectors.[10] Indeed, scholars can use several sources—administrative reports, legislative testimony and related documents, interviews, among others—to ascertain the specific information sources on which decision-makers rely (Jacob 1988; Mossberger 2000; Weyland 2006; Karch 2007a; Porto de Oliveira 2019; Pacheco-Vega 2021). Scholarship on diffusion among the American states would benefit greatly from a greater emphasis on process tracing and other qualitative techniques.

Although EHA remains the most common quantitative approach to the study of diffusion, several promising alternatives have emerged recently. For example, some diffusion studies use spatial econometrics to assess the impact of geographic proximity (Arel-Bundock and Parinandi 2018; Mitchell 2018). Other studies use a text-as-data approach, relying on computational analysis to quantify the overlap between statutes or between legislative proposals and the templates provided by various organizations (Collingwood, El-Khatib, and Gonzalez O'Brien 2019; Callaghan, Karch, and Kroeger 2020; Linder, Desmarais, Burgess, and Giraudy 2020). Finally, experimental designs investigate the factors that influence the sources to which decision-makers are drawn as they learn about and evaluate novel innovations (Butler et al. 2017; Zelizer 2018; Pereira 2021). These and other alternative approaches deserve serious examination, as they might address some of the limitations of EHA and its extensions. They might verify, extend, or challenge the ways in which diffusion scholars have interpreted their findings, potentially leading to novel quantitative approaches, measurement techniques, and conceptual developments.

In closing, it is critical to revisit the connection between conceptual advances in the study of innovation and diffusion and research methods. Amassing ever larger databases of innovations is a common trend in the scholarly literature (Boushey 2010; Boehmke and Skinner 2012; Boehmke et al. 2020). To be sure, these data repositories provide treasure troves of information and a valuable service for the community of scholars engaged in diffusion research. Yet it is equally important to recognize that they privilege a certain type of study, one focused nearly exclusively on adoption patterns. The adoption decision is only one component of the policy process, and this single-minded focus renders the literature incomplete. The other chapters of this book assess canonical theories that invoke policy-making stages from agenda setting to implementation. Developing a stronger connection between these theories and diffusion research must be a high priority. Diffusion is an important concept, but scholars must not lose sight of its capacity to offer general lessons about the policy process, information transmission, and decision-making.

One way to generate useful lessons for policy scholars is to build on recent conceptual developments in diffusion research. Earlier reviews of the

diffusion literature characterized it as overly mechanical (Karch 2007b), and the recent turn toward diffusion mechanisms and vectors represents a promising corrective. As scholars move beyond illustrating the potential impact of these mechanisms and vectors, they must strive to identify the conditions and contexts in which they are most likely to be influential. Again, this requires a close connection between theory and method. Sometimes studies of large numbers of innovations will prove constructive; sometimes a narrower inquiry will be more appropriate. Indeed, undertaking various approaches will likely be the best path forward.

The past offers important lessons. When scholars introduced EHA to diffusion research in the early 1990s, the novel statistical technique sparked a massive outburst of published work on the topic. Studies of individual innovations accumulated, but major conceptual advances were uncommon as scholars emphasized geographic proximity at the expense of other mechanisms of interjurisdictional interdependencies. Contemporary diffusion scholars do not only have the advantage of better developed statistical approaches. They also possess a more robust set of theories and concepts on which to base their work. In addition, the creation of the SPID and PDR databases demonstrate a broader interest in advancing scholars' collective knowledge by bringing together various strands of the disparate scholarly literature (Boehmke et al. 2020; Mallinson 2020). Most of these efforts are limited to analyses of American state politics; similar efforts that cross conventional subfield boundaries can help move the literature forward, since the balance between internal and external influences is a widespread concern (Graham, Shipan, and Volden 2013).

Regardless of the context in which they work, diffusion scholars must engage in careful theorizing about the conditions in which interjurisdictional interdependencies are expected to be influential. They must remain attuned to recent conceptual developments, considering whether and how specific mechanisms and vectors might influence the substantive phenomena in which they are interested. Their decisions about research design must be guided by the questions they ask and the hypotheses they seek to evaluate. Fortunately, they have a growing methodological tool kit on which to draw. Their collective willingness to embrace these varied tools and their complementary strengths will contribute to a better understanding of the dynamics of the diffusion process.

Notes

1 For example, several review essays assess the conceptual and methodological relationship between policy transfer and diffusion (Marsh and Sharman 2009; Benson and Jordan 2011; Dussauge-Laguna 2012).
2 Multilevel modeling is a way to accommodate this heterogeneity (Kreitzer and Boehmke 2016).
3 In addition, multiple studies suggest that conventional studies of the neighboring state effect tend to overstate its impact (Mooney 2001; Karch et al. 2016).

4 Shipan and Volden (2008) describe coercion as a fourth diffusion mechanism. Although coercion plays a central role in studies of policy transfer (Dolowitz and Marsh 2000; Benson and Jordan 2011; Dussauge-Laguna 2012), some scholars argue that it is not a mechanism because "diffusion implies that no central actors are coordinating the spread of a policy" (Maggetti and Gilardi 2016, 90). Coercion will not be discussed in this essay, but other studies examine the topic in more detail (Mooney 2020).

5 In addition, a policy innovation can gain widespread adoption even when multiple evaluations suggest that it has not achieved its primary objectives (Park and Berry 2014).

6 Participation in the global economy can drive similar convergence among firms, although the strength of this effect varies across industries (Malesky and Mosley 2018).

7 Although the focus here is on organizations and model legislation, computational text analysis can also be used to assess the impact of other jurisdictions' legislative templates (Callaghan, Karch, and Kroeger 2020).

8 In his classic synthesis of research on the diffusion of innovations, Everett Rogers (1995, 100) defines the pro-innovation bias as "the implication … that an innovation should be diffused and adopted by all members of a social system, that it should be diffused more rapidly, and that the innovation should be neither reinvented nor rejected." Since this chapter focuses on methodological issues, it offers a slightly different definition that focuses on the issue of case selection.

9 The SPID, which "relatively equally covers widely and narrowly diffusing policies" (Boehmke et al. 2020, 526), may help address this limitation.

10 Some comparative politics scholars have called for the more widespread use of statistical techniques, highlighting the complementary nature of quantitative and qualitative research designs (Marsh and Sharman 2009, 273).

References

Arel-Bundock, Vincent, and Srinivas Parinandi. 2018. Conditional Tax Competition in American States. *Journal of Public Policy* 38 (2): 191–220.

Balla, Steven J. 2001. Interstate Professional Associations and the Diffusion of Policy Innovations. *American Politics Research* 29 (3): 221–245.

Baumgartner, Frank R., Virginia Gray, and David Lowery. 2009. Federal Policy Activity and the Mobilization of State Lobbying Organizations. *Political Research Quarterly* 62 (3): 552–567.

Baybeck, Brady, William D. Berry, and David A. Siegel. 2011. A Strategic Theory of Policy Diffusion via Intergovernmental Competition. *Journal of Politics* 73 (1): 232–247.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable. *American Journal of Political Science* 42 (4): 1260–1288.

Benson, David, and Andrew Jordan. 2011. What Have We Learned from Policy Transfer Research? Dolowitz and Marsh Revisited. *Political Studies Review* 9 (3): 366–378.

Berry, Frances Stokes, and William D. Berry. 1990. State Lottery Adoptions as Policy Innovations: An Event History Analysis. *American Political Science Review* 84 (2): 395–415.

Berry, Frances Stokes, and William D. Berry. 1992. Tax Innovation in the States: Capitalizing on Political Opportunity. *American Journal of Political Science* 36 (3): 715–742.

Berry, Frances Stokes, and William D. Berry. 2018. "Innovation and Diffusion Models in Policy Research." In *Theories of the Policy Process*, edited by Christopher M. Weible and Paul A. Sabatier. Fourth Edition. New York, NY: Routledge (pp. 253–297).

Boehmke, Frederick J. 2009a. Approaches to Modeling the Adoption and Diffusion of Policies with Multiple Components. *State Politics & Policy Quarterly* 9 (2): 229–252.

Boehmke, Frederick J. 2009b. Policy Emulation or Policy Convergence? Potential Ambiguities in the Dyadic Event History Approach to State Policy Emulation. *Journal of Politics* 71 (3): 1125–1140.

Boehmke, Frederick J., Mark Brockway, Bruce A. Desmarais, Jeffrey J. Harden, Scott LaCombe, Fridolin Linder, and Hanna Wallach. 2020. SPID: A New Database for Inferring Public Policy Innovativeness and Diffusion Networks. *Policy Studies Journal* 48 (2): 517–545.

Boehmke, Frederick J., and Paul Skinner. 2012. State Policy Innovativeness Revisited. *State Politics & Policy Quarterly* 12 (3): 303–329.

Boehmke, Frederick J., and Richard Witmer. 2004. Disentangling Diffusion: The Effects of Social Learning and Economic Competition on State Policy Innovation and Expansion. *Political Research Quarterly* 57 (1): 39–51.

Boushey, Graeme. 2010. *Policy Diffusion Dynamics in America*. New York, NY: Cambridge University Press.

Boushey, Graeme. 2016. Targeted for Diffusion? How the Use and Acceptance of Stereotypes Shape the Diffusion of Criminal Justice Policy Innovations in the American States. *American Political Science Review* 110 (1): 198–214.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. Time Is of the Essence: Event History Models in Political Science. *American Journal of Political Science* 41 (4): 1414–1461.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Bromley-Trujillo, Rebecca, and Andrew Karch. 2021. Salience, Scientific Uncertainty, and the Agenda-Setting Power of Science. *Policy Studies Journal* 49 (4): 992–1018.

Bromley-Trujillo, Rebecca, and John Poe. 2020. The Importance of Salience: Public Opinion and State Policy Action on Climate Change. *Journal of Public Policy* 40 (2): 280–304.

Buckley, Jack, and Chad Westerland. 2004. Duration Dependence, Functional Form, and Corrected Standard Errors: Improving EHA Models of State Policy Diffusion. *State Politics & Policy Quarterly* 4 (1): 94–113.

Butler, Daniel M., Craig Volden, Adam M. Dynes, and Boris Shor. 2017. Ideology, Learning, and Policy Diffusion: Experimental Evidence. *American Journal of Political Science* 61 (1): 37–49.

Callaghan, Timothy, Andrew Karch, and Mary Kroeger. 2020. Model State Legislation and Intergovernmental Tensions over the Affordable Care Act, Common Core, and the Second Amendment. *Publius: The Journal of Federalism* 50 (3): 518–539.

Collingwood, Loren, Stephen Omar El-Khatib, and Benjamin Gonzalez O'Brien. 2019. Sustained Organizational Influence: American Legislative Exchange Council and the Diffusion of Anti-Sanctuary Policy. *Policy Studies Journal* 47 (3): 735–773.

Crain, Robert L. 1966. Fluoridation: The Diffusion of an Innovation among Cities. *Social Forces* 44 (4): 467–476.

Desmarais, Bruce A., Jeffrey J. Harden, and Frederick J. Boehmke. 2015. Persistent Policy Pathways: Inferring Diffusion Networks in the American States. *American Political Science Review* 109 (2): 392–406.

Dolowitz, David, and David Marsh. 1996. Who Learns What from Whom: A Review of the Policy Transfer Literature. *Political Studies* 44 (2): 343–357.

Dolowitz, David, and David Marsh. 2000. Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making. *Governance* 13 (1): 5–24.

Dussauge-Laguna, Mauricio I. 2012. On the Past and Future of Policy Transfer Research: Benson and Jordan Revisited. *Political Studies Review* 10 (3): 313–324.

Elkins, Zachary, and Beth Simmons. 2005. On Waves, Clusters, and Diffusion: A Conceptual Framework. *Annals of the American Academy of Political and Social Science* 598: 33–51.

Eyestone, Robert. 1977. Confusion, Diffusion, and Innovation. *American Political Science Review* 71 (2): 441–447.

Finger, Leslie K. 2018. Vested Interests and the Diffusion of Education Reform across the States. *Policy Studies Journal* 46 (2): 378–401.

Foster, John L. 1978. Regionalism and Innovation in the American States. *Journal of Politics* 40 (1): 179–187.

Garrett, Kristin N., and Joshua M. Jansa. 2015. Interest Group Influence in Policy Diffusion Networks. *State Politics & Policy Quarterly* 15 (3): 387–417.

Gilardi, Fabrizio. 2015. The Temporary Importance of Role Models for Women's Political Representation. *American Journal of Political Science* 59 (4): 957–970.

Gilardi, Fabrizio. 2016. Four Ways We Can Improve Policy Diffusion Research. *State Politics & Policy Quarterly* 16 (1): 8–21.

Graham, Erin R., Charles R. Shipan, and Craig Volden. 2013. The Diffusion of Policy Diffusion Research in Political Science. *British Journal of Political Science* 43 (3): 673–701.

Grattet, Ryken, Valerie Jenness, and Theodore R. Curry. 1998. The Homogenization and Differentiation of Hate Crime Law in the United States, 1978–1995: Innovation and Diffusion in the Criminalization of Bigotry. *American Sociological Review* 63 (2): 286–307.

Gray, Virginia. 1973. Innovation in the States: A Diffusion Study. *American Political Science Review* 67 (4): 1174–1185.

Gray, Virginia. 1994. "Competition, Emulation, and Policy Innovation." In *New Perspectives on American Politics*, edited by Lawrence C. Dodd and Calvin Jillson. Washington, DC: CQ Press (pp. 230–248).

Grossback, Lawrence J., Sean Nicholson-Crotty, and David A. M. Peterson. 2004. Ideology and Learning in Policy Diffusion. *American Politics Research* 32 (5): 521–545.

Haider-Markel, Donald P. 2001. Policy Diffusion as a Geographical Expansion of the Scope of Political Conflict. *State Politics & Policy Quarterly* 1 (1): 5–26.

Hays, Scott P. 1996. Influences on Reinvention during the Diffusion of Innovations. *Political Research Quarterly* 49 (3): 631–650.

Hays, Scott P., and Henry R. Glick. 1997. The Role of Agenda Setting in Policy Innovation: An Event History Analysis of Living-Will Laws. *American Politics Quarterly* 25 (4): 497–516.

Heilmann, Sebastian, and Nicole Schulte-Kulkmann. 2011. The Limits of Policy Diffusion: Introducing International Norms of Anti-Money Laundering into China's Legal System. *Governance* 24 (4): 639-664.

Hertel-Fernandez, Alexander. 2019. *State Capture: How Conservative Activists*, *Big Businesses*, *and Wealthy Donors Reshaped the American States—And the Nation*. New York, NY: Oxford University Press.

Jacob, Herbert. 1988. *Silent Revolution: The Transformation of Divorce Law in the United States*. Chicago, IL: University of Chicago Press.

Jacobs, Lawrence, Theodore Marmor, and Jonathan Oberlander. 1999. The Oregon Health Plan and the Political Paradox of Rationing: What Advocates and Critics Have Claimed and What Oregon Did. *Journal of Health Politics*, *Policy and Law* 24 (1): 161–180.

Jansa, Joshua M., Eric R. Hansen, and Virginia H. Gray. 2018. Copy and Paste Lawmaking: Legislative Professionalism and Policy Reinvention in the States. *American Politics Research* 47 (4): 739–767.

Jones, Trevor, and Tim Newburn. 2002. Learning from Uncle Sam? Exploring U.S. Influence on British Crime Control Policy. *Governance* 15 (1): 97–119.

Karch, Andrew. 2007a. *Democratic Laboratories: Policy Diffusion among the American States*. Ann Arbor, MI: University of Michigan Press.

Karch, Andrew. 2007b. Emerging Issues and Future Directions in State Policy Diffusion Research. *State Politics & Policy Quarterly* 7 (1): 54–80.

Karch, Andrew. 2012. Vertical Diffusion and the Policy-Making Process: The Politics of Embryonic Stem Cell Research. *Political Research Quarterly* 65 (1): 48–61.

Karch, Andrew, and Matthew Cravens. 2014. Rapid Diffusion and Policy Reform: The Adoption and Modification of Three Strikes Laws. *State Politics & Policy Quarterly* 14 (4): 461–491.

Karch, Andrew, Sean C. Nicholson-Crotty, Neal D. Woods, and Ann O'M. Bowman. 2016. Policy Diffusion and the Pro-Innovation Bias. *Political Research Quarterly* 69 (1): 83–95.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca, NY: Cornell University Press.

Kelley, Judith, and Beth A. Simmons. 2015. Politics by Number: Indicators as Social Pressure in International Relations. *American Journal of Political Science* 59 (1): 55–70.

Kirst, Michael W., Gail Meister, and Stephen R. Rowley. 1984. Policy Issue Networks: Their Influence on State Policymaking. *Policy Studies Journal* 13 (2): 247–263.

Kreitzer, Rebecca J. 2015. Politics and Morality in State Abortion Policy. *State Politics & Policy Quarterly* 15 (1): 41–66.

Kreitzer, Rebecca J., and Frederick J. Boehmke. 2016. Modeling Heterogeneity in Pooled Event History Analysis. *State Politics & Policy Quarterly* 16 (1): 121–141.

LaCombe, Scott J., and Frederick J. Boehmke. 2020. "Learning and Diffusion Models." In *The Sage Handbook of Research Methods in Political Science and International Relations*, edited by Luigi Curini and Robert Franzese. Newbury Park, CA: Sage Publications (pp. 311–328).

Lax, Jeffrey R., and Justin H. Phillips. 2009. How Should We Estimate Opinion in the States? *American Journal of Political Science* 53 (1): 107–121.

Lax, Jeffrey R., and Justin H. Phillips. 2012. The Democratic Deficit in the States. *American Journal of Political Science* 56 (1): 148–166.

Levitt, Barbara, and James G. March. 1988. Organizational Learning. *Annual Review of Sociology* 14: 319–340.

Lindblom, Charles E. 1959. The Science of "Muddling Through." *Public Administration Review* 19 (2): 79–88.

Linder, Fridolin, Bruce A. Desmarais, Matthew Burgess, and Eugenia Giraudy. 2020. Text as Policy: Measuring Policy Similarity through Bill Text Reuse. *Policy Studies Journal* 48 (2): 546–574.

Maggetti, Martino, and Fabrizio Gilardi. 2016. Problems (and Solutions) in the Measurement of Policy Diffusion Mechanisms. *Journal of Public Policy* 36 (1): 87–107.

Makse, Todd, and Craig Volden. 2011. The Role of Policy Attributes in the Diffusion of Innovations. *Journal of Politics* 73 (1): 108–124.

Malesky, Edmund J., and Layna Mosley. 2018. Chains of Love? Global Production and the Firm-Level Diffusion of Labor Standards. *American Journal of Political Science* 62 (3): 712–728.

Mallinson, Daniel J. 2020. Policy Diffusion Results: A New Database for Systematic Review and Meta-Analysis. *State and Local Government Review* 52 (1): 18–27.

Mallinson, Daniel J. 2021. Who Are Your Neighbors? The Role of Ideology and Decline of Geographic Proximity in the Diffusion of Policy Innovations. *Policy Studies Journal* 49 (1): 67–88.

Marsh, David, and J.C. Sharman. 2009. Policy Diffusion and Policy Transfer. *Policy Studies* 30 (3): 269–288.

McNeal, Ramona S., Caroline J. Tolbert, Karen Mossberger, and Lisa J. Dotterweich. 2003. Innovating in Digital Government in the American States. *Social Science Quarterly* 84 (1): 52–70.

Mintrom, Michael. 1997. Policy Entrepreneurs and the Diffusion of Innovation. *American Journal of Political Science* 41 (3): 738–770.

Mintrom, Michael. 2000. *Policy Entrepreneurs and School Choice*. Washington, DC: Georgetown University Press.

Mitchell, Joshua L. 2016. Examining Media Markets' Influence on the Diffusion of Antismoking Legislation in US Counties. *Local Government Studies* 42 (5): 665–688.

Mitchell, Joshua L. 2018. Does Policy Diffusion Need Space? Spatializing the Dynamics of Policy Diffusion. *Policy Studies Journal* 46 (2): 424–451.

Mooney, Christopher Z. 1991. Information Sources in Legislative Decision Making. *Legislative Studies Quarterly* 16 (3): 445–455.

Mooney, Christopher Z. 2001. Modeling Regional Effects on State Policy Diffusion. *Political Research Quarterly* 54 (1): 103–124.

Mooney, Christopher Z. 2020. *The Study of US State Policy Diffusion: What Hath Walker Wrought?* Cambridge: Cambridge University Press.

Mooney, Christopher Z., and Mei-Hsien Lee. 1995. Legislative Morality in the American States: The Case of Pre-*Roe* Abortion Regulation Reform. *American Journal of Political Science* 39 (3): 599–627.

Mossberger, Karen. 2000. *The Politics of Ideas and the Spread of Enterprise Zones*. Washington, DC: Georgetown University Press.

Myers, Adam S. 2018. The Failed Diffusion of the Unicameral State Legislature, 1934–1944. *Studies in American Political Development* 32 (2): 217–235.

Nathan, Richard P. 2000. *Social Science in Government: The Role of Policy Researchers*. Albany, NY: Rockefeller Institute Press.

Nicholson-Crotty, Sean. 2009. The Politics of Diffusion: Public Policy in the American States. *Journal of Politics* 71 (1): 192–205.

Nicholson-Crotty, Sean, and Sanya Carley. 2016. Effectiveness, Implementation, and Policy Diffusion: Or "Can We Make That Work for Us?" *State Politics & Policy Quarterly* 16 (1): 78–97.

Osorio Gonnet, Cecilia. 2019. A Comparative Analysis of the Adoption of Conditional Cash Transfers Programs in Latin America. *Journal of Comparative Policy Analysis: Research and Practice* 21 (4): 385–401.

Pacheco, Julianna. 2012. The Social Contagion Model: Exploring the Role of Public Opinion on the Diffusion of Antismoking Legislation across the American States. *Journal of Politics* 74 (1): 187–202.

Pacheco-Vega, Raul. 2021. "Policy Transfer of Environmental Policy: Where Are We Now and Where Are We Going? Examples from Water, Climate, Energy, and Waste Sectors." In *Handbook of Policy Transfer, Diffusion and Circulation*, edited by Osmany Porto de Oliveira. Cheltenham: Edward Elgar Publishing, Ltd. (pp. 384–403).

Park, Seejeen, and Frances Berry. 2014. The Successful Diffusion of a Failed Policy: The Case of Pay-for-Performance in the US Federal Government. *Public Management Review* 16 (6): 763–781.

Pavalko, Eliza K. 1989. State Timing of Policy Adoption: Workmen's Compensation in the United States, 1909–1929. *American Journal of Sociology* 95 (3): 592–615.

Pereira, Miguel M. 2021. How Do Public Officials Learn about Policy? A Field Experiment on Policy Diffusion. *British Journal of Political Science*. https://doi.org/10.1017/S0007123420000770

Porto de Oliveira, Osmany. 2019. Brazil Exporting Social Policies: From Local Innovation to a Global Model. *Journal of Politics in Latin America* 11 (3): 249–271.

Porto de Oliveira, Osmany. 2020. Policy Ambassadors: Human Agency in the Transnationalization of Brazilian Social Policies. *Policy and Society* 39 (1): 53–69.

Rogers, Everett M. 1995. *Diffusion of Innovations*. Fourth Edition. New York, NY: Free Press.

Roh, Jongho, and Donald P. Haider-Markel. 2003. All Politics Is Not Local: National Forces in State Abortion Initiatives. *Social Science Quarterly* 84 (1): 15–31.

Sabatier, Paul, and David Whiteman. 1985. Legislative Decision Making and Substantive Policy Information: Models of Information Flow. *Legislative Studies Quarterly* 10 (3): 395–421.

Savage, Robert L. 1978. Policy Innovativeness as a Trait of American States. *Journal of Politics* 40 (1): 179–187.

Shanahan, Elizabeth A., Mark K. McBeth, Paul L. Hathaway, and Ruth J. Arnell. 2008. Conduit or Contributor? The Role of Media in Policy Change Theory. *Policy Sciences* 41 (2): 115–138.

Shipan, Charles R., and Craig Volden. 2006. Bottom-Up Federalism: The Diffusion of Antismoking Policies from U.S. Cities to States. *American Journal of Political Science* 50 (4): 825–843.

Shipan, Charles R., and Craig Volden. 2008. The Mechanisms of Policy Diffusion. *American Journal of Political Science* 52 (4): 840–857.

Shriver, Thomas E., Adriana M. Szabo, and Laura A. Bray. 2020. Opportunity and Threat behind the Iron Curtain: The Failed Diffusion of Human Rights Activism in Romania. *International Sociology* 35 (6): 739–757.

Simmons, Beth A., Paulette Lloyd, and Brandon M. Stewart. 2021. The Global Diffusion of Law: Transnational Crime and the Case of Human Trafficking. *International Organization* 72 (2): 249–281.

Skocpol, Theda, Marjorie Abend-Wein, Christopher Howard, and Susan Goodrich Lehmann. 1993. Women's Associations and the Enactment of Mothers' Pensions in the United States. *American Political Science Review* 87 (3): 686–701.

Stone, Diane, Osmany Porto de Oliveira, and Leslie A. Pal. 2020. Transnational Policy Transfer: The Circulation of Ideas, Power and Development Models. *Policy and Society* 39 (1): 1–18.

Teodoro, Manuel P. 2009. Bureaucratic Job Mobility and the Diffusion of Innovations. *American Journal of Political Science* 53 (1): 175–189.

Volden, Craig. 2006. States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program. *American Journal of Political Science* 50 (2): 294–312.

Volden, Craig, Michael M. Ting, and Daniel P. Carpenter. 2008. A Formal Model of Learning and Diffusion. *American Political Science Review* 102 (3): 319–332.

Walker, Jack L., Jr. 1969. The Diffusion of Innovations among the American States. *American Political Science Review* 63 (3): 880–899.

Walker, Jack L., Jr. 1981. "The Diffusion of Knowledge, Policy Communities and Agenda Setting: The Relationship of Knowledge and Power." In *New Strategic Perspectives on Social Policy*, edited by John E. Tropman, Milan J. Dluhy, and Roger M. Lind. New York, NY: Pergamon (pp. 79–96).

Weyland, Kurt. 2006. *Bounded Rationality and Policy Diffusion: Social Sector Reform in Latin America*. Princeton, NJ: Princeton University Press.

Winburn, Jonathan, Amanda Winburn, and Ryan Niemeyer. 2014. Media Coverage and Issue Visibility: State Legislative Responses to School Bullying. *The Social Science Journal* 51 (4): 514–522.

Yi, Hongtao, Frances Stokes Berry, and Wenna Chen. 2018. Management Innovation and Policy Diffusion through Leadership Transfer Networks: An Agent Network Diffusion Model. *Journal of Public Administration Research and Theory* 28 (4): 457–474.

Yi, Hongtao, and Wenna Chen. 2019. Portable Innovation, Policy Wormholes, and Innovation Diffusion. *Public Administration Review* 76 (5): 737–748.

Yu, Jinhai, Edward T. Jennings, Jr., and J. S. Butler. 2020. Lobbying, Learning, and Policy Reinvention: An Examination of the American States' Drunk Driving Laws. *Journal of Public Policy* 40 (2): 259–279.

Zelizer, Adam. 2018. How Responsive Are Legislators to Policy Information? Evidence from a Field Experiment in a State Legislature. *Legislative Studies Quarterly* 43 (4): 595–618.

Online Data Repositories

The **Correlates of State Policy Project** includes more than 900 variables, with observations across the 50 American states and over time (from approximately 1900 to 2016). It includes policy outputs and various political, social, and economic factors that might be correlated with differences in those outputs.

http://ippsr.msu.edu/public-policy/correlates-state-policy

The **Global BRT Data** platform provides information about Bus Rapid Transit (BRT) systems, an innovation in sustainable urban transport that has received significant scholarly attention. The platform consolidates data from 177 cities across the globe. See Pacheco-Vega (2021) for more information.

http://brtdata.org

The **Policy Diffusion Results (PDR) Database** includes 6,641 variables from 507 models of policy diffusion among the American states. The models are drawn

from peer–reviewed articles published from 1990 to 2018. See Mallinson (2020) for more information.

   https://doi.org/10.7910/DVN/NASPUC

The **State Policy Innovation Database (SPID)** includes the year of adoption for more than 700 policy innovations enacted in the American states. It also includes information about the policies. See Boehmke et al. (2020) for more information.

   https://dataverse.harvard.edu/dataverse/spid#

8 Methods for Analyzing Social Dilemmas and Institutional Arrangements Within the Institutional Analysis and Development Framework

Edella Schlager, Saba Siddiki, and Michael Cox

Introduction

Elinor Ostrom, in her presidential address to the American Political Science Association stated, "the theory of collective action is the central subject of political science" (1998, p. 1).

Across the subfields of political science, whether international relations or comparative politics, American politics or political theory, rich veins of research have focused on the many puzzles presented by collective action and social dilemmas. For instance, why do citizens vote, which is a form of collective action, if a single vote is unlikely to decide an election? Or, why do nation-states invest in, adopt, and abide by treaties, even though the treaties lack enforcement mechanisms?

The Institutional Analysis and Development (IAD) framework, developed by E. Ostrom and colleagues over the years (Kiser and Ostrom 1982; Ostrom, Gardner, and Walker 1994; Ostrom 1999, 2009) provides social scientists with the means to systematically study the role of institutional arrangements in exacerbating or resolving social dilemmas, and how actors devise, implement, and revise such arrangements. Scholars working with the IAD framework employ multiple types of methods particularly appropriate for the study of social dilemmas, collective action, and the design and effects of institutional arrangements on behavior.

The IAD framework is well suited for studying these topics because at its core is the action situation, which represents *interdependent* actions among actors. An action situation is defined as any setting in which two or more actors "are faced with a set of potential actions that jointly produce outcomes" (Ostrom 2005:32). In other words, the action situation captures social dilemmas and efforts at engaging in collective action to resolve such dilemmas. In addition, the IAD framework incorporates categories of factors that shape action situations: rules-in-form and use, community attributes, and physical/material conditions. Institutions, the rules and norms that delineate what participants operating within action situations are required, allowed, or forbidden to do, are central to IAD framework studies. Institutional

arrangements (in conjunction with community characteristics and physical/material conditions) shape behavior, and actors create, revise, or eliminate institutional arrangements to achieve more desirable shared outcomes.

Studies of institutional arrangements and behavior in different contexts have largely focused on comparative analyses of institutional arrangements structuring operational level action situations; the individual, collective, and systemic outcomes of interactions governed by exogenously or endogenously generated institutions and how outcomes feed back into action situations to shape institutional designs. Scholars have studied these questions using case studies, laboratory and field experiments, agent-based modeling (ABM), and network analysis. Case studies allow for a comparative analysis of action situations, within or across types of setting, e.g., coastal fisheries, irrigation systems, or community gardens (Cox et al. 2010; Baggio et al. 2016). Laboratory and field experiments create carefully controlled action situations allowing analysts to examine how actors' behavior (and group outcomes) is affected by a specific treatment, typically a rule change, or a change in physical/material conditions. Agent-based models support assessments of interactions among actors in simulated action situations (Frantz et al. 2014). Finally, social network analysis allows for the examination of the structure and patterns of networks of rules constituting action situations. More recently, how networks of prescribed interactions (i.e., rules) and perceived interactions (i.e., behavior) jointly structure outcomes has been explored (Olivier 2019; Olivier and Schlager 2021).

What these methods have in common is that they allow analysts to identify, represent, and systematically analyze interactions among actors in analytically defined settings called action situations. An IAD study begins by identifying the action situations that will be studied by these methods. Action situations represent settings in which interdependent actors experience and/or have resolved social dilemmas by devising, implementing, and enforcing institutional arrangements that support coordination and cooperation. For instance, street gangs that have organized their control over city neighborhoods (Matsueda 2006), or harbor gangs that have organized access and use of lobster territories within and adjacent to harbors (Acheson 1988), or illicit business cartels that have defined and organized their regional territories and activities for extracting resources (Jaspers 2017). However, action situations need not be strictly grounded in specific geographies. For instance, action situations include officials elected to decision-making bodies who share similar interests and who must overcome collective action dilemmas in order to realize their shared goals in the larger decision-making body, such as party caucuses (Forgette 2004). The key is identifying social dilemmas and the actors involved.

Action situations are embedded in contexts, as explicated by the IAD framework. Variation in the types of collective action dilemmas that actors experience, and the types of institutional arrangements appropriate to their resolution are conditioned by context, i.e., the physical and material conditions, the community setting, the rules in use, and levels of action.

Identifying action situations begins with identifying collective action dilemmas, but explaining action situations requires understanding how they are conditioned by their context, as we explain below.

In the remainder of this chapter, we describe each method; major questions examined and how they relate to the IAD framework; and data collection and analysis. We conclude this chapter with promising lines of research on institutions.

Methods

In 2010, Poteete, Janssen, and Ostrom published *Working Together*, describing the different methods used in theory development and testing regarding common pool resource governance. They argued that the use of multiple methods is critical for theory development. Moving from cases to experiments, to ABM and simulations, and back to cases or experiments allows scholars to explore a panoply of questions and build a rich body of evidence. While we describe the methods in order, they often work together to allow scholars to pinpoint factors that are critical for explaining a specific social dilemma or identify a necessary condition for the emergence of cooperation. Case studies suggest "treatments" that may be explored in experiments; agent-based models allow for the "treatments" to be deployed across many diverse action situations, in order to explore patterns of individual and collective outcomes; and network analyses may reveal the logic of institutional design found in case studies. Or, experiments may highlight the importance of specific types of enforcement mechanisms for resolving collective action dilemmas that may then be identified and studied in cases, and so forth.

In this section, we describe and update each of the methods that appears in *Working Together*, as well as describe social network analysis and how it is used to analyze action situations and networks of rules.

Case Studies

A case study characterizes a time and place. An IAD case study does this by breaking a case down into one or more actor groups, their institutional arrangements, and their biophysical setting. It then attempts to explain outcomes based on these components. Like all case studies, the application of the IAD framework to a case depends on a variety of data sources. Several of these are tailored to the specific type of data being collected as directed by the categories of the framework. Institutional information often requires collecting documents (formal rules) and conducting in-person interviews (informal rules) with individuals who work and/or live within the system of interest. Participant observation is also frequently important, as institutions are expressed through patterns of behavior as much as they are through linguistic statements. Because of these features, many IAD case studies involve extensive fieldwork in the case study area. Additionally, because of

the geographically explicit nature of many IAD case studies, secondary spatial data are often collected and analyzed in geographic information system (GIS) environments. Increasingly, remote sensing analyses are incorporated as well.

A key theme in IAD case studies is the overlapping themes of comparison and multiple levels of analysis. IAD case studies are frequently "embedded" (Yin 2014), in that there are multiple levels or units of analysis, one embedded within the other. A common way in which this is done is by collecting data on a set of natural resource users in order to understand the community they constitute (Cox et al. 2018). Household-style interviews are commonly complemented in these studies by key informant interviews with other types of actors (Nigussie et al. 2018). As such, IAD case studies are also frequently "mixed method," combining quantitative analysis of household data with qualitative analysis of other data sources and of the case itself.

Equally common is the comparison of multiple cases, particularly if the case is regional to local in scale. What is less common is the use of a hierarchical statistical model (fixed or random effects) that describes both within- and across-case variation. Even without multiple cases, as has been discussed before with respect to case studies in general (George and Bennett 2005), single IAD case studies are frequently a comparative method as well: the inferences generated from them come from the comparison of features nested within them. A central way in which the IAD framework does this is through the concept of the action situation network (ASN) (Sendzimir et al. 2010; McGinnis 2011; Cox 2014; Möck et al. 2019).

The ASN is at once a major distinctive feature of the IAD framework, and also what ties it closely to other approaches, such as the "ecology of games framework" (Lubell 2013). It represents the observation that decisions made by one set of actors in one arena affect the incentives and actions of other actors in other arenas, often reciprocally. If we want to understand the incentives and actions of one set of actors, say a group of fishers, we can interview those fishers. But if we want to understand the full causal chain that led to those actions in this "focal action situation" (Ostrom 2009; Hinkel et al. 2015), we often will need to ask the same question to "upstream" actors whose decisions affect the incentives of these fishers. This reinforces the previous observation that IAD case studies frequently combine household interviews with more in-depth key informant interviews. By starting with a target behavior of a target actor group and going "upstream" in this way, IAD scholars can unpack the behavioral dynamics of a case study. Why we would want to can be expressed as a generic version of the problem of omitted variable bias from statistics: if we leave out a factor that relates to the outcomes we care about and the factors we do include, this will bias the interpretations we have of the factors we do include. Two systems might look quite similar with respect to their focal action situations, but if we expand out and go upstream, important differences can reveal themselves to explain divergent outcomes.

In applications of the IAD, the most common example of an ASN is based on the difference between operational and collective-choice (and sometimes constitutional) rules (Hardy and Koontz 2009). Operational rules are those that dictate everyday behavior and are similar to informal norms. Collective-choice rules are more formal and dictate how decisions are made regarding the content of operational rules. This spectrum of rule types correlates with increasing formality, increasing institutional "stickiness" (North 1990) and subsequent potential for institutional lock-in.

An ASN represents causal impact over space and time and organizational context. IAD case studies have addressed geographic and social space better than they have addressed the dynamics of time (see Johnson 2004). There is some important literature on studying institutional change (see again North 1990), and Ostrom commented on it increasingly throughout her career (Ostrom and Basurto 2011). But the idea of change is not firmly entrenched in the IAD framework itself, other than the arrow pointing back from the "Outcomes" box to the conditioning factors.

IAD case studies frequently make inferences with respect to the functionality of the institutions that are present. Doing so cannot only depend on traditional statistical or quantitative methods (Ostrom and Basurto 2011). Instead of, or at least in addition to this, appeals to function and adaptation are implicitly made, in line with similar analysis conducted by evolutionary biologists. We can call this the "adaptationist" perspective, as embodied in Ostrom (1990), who inferred the adaptability of institutional design principles based in large part on their presence in long-lasting community-based resource management systems. The adaptationist perspective often infers adaptability from longevity, assuming an evolutionary logic and the presence of a historical process that selected for traits that would help resolve adaptive challenges: in the case of IAD studies, these most often take the form of collective-action problems oriented around the provision of public goods. Common property arrangements themselves have been seen to be an adaptation to resource scarcity, explaining their disproportionate presence in resource-scarce environments. Local knowledge and customs are also often interpreted through an adaptationist lens (Boyd 2017).

While this perspective is a powerful framework for institutional theorists (Hodgson 2007), there are many reasons why a trait might persist, other than being adaptive (Gould and Lewontin 1979). We see institutions all around us that have persisted for some time, but which promote inefficient and inequitable outcomes, and many have argued persuasively that institutions are mostly a result of power relations and conflict (Knight 1992). This relates to a critique often applied to the IAD framework: that it is under-politicized (Clement 2010).

A further complication here is the possibility for mismatches between adaptations to previous circumstances and current ones. The IAD literature discusses different types of mismatches that are argued to affect outcomes. Most commonly discussed are mismatches between institutional arrangements and features of an environmental commons, such as its spatial

scale or temporal dynamics. We can add evolutionary mismatches to the common typology of mismatches that are discussed. Schlaepfer et al. (2002) describe this situation as an "evolutionary trap," in which "an environment that has been altered suddenly by human activities, an organism makes a maladaptive behavior or life-history choice based on formerly reliable environmental cues, despite the availability of higher quality options." In an institutional context, an evolutionary mismatch would indicate that institutions that have developed in response to historic challenges involved in, say, managing small-scale commons and the modern environment in which many communities find themselves (see, e.g., Cox 2014).

This discussion implies that IAD case studies should ideally contain a narrative history of a case, to unpack what selective pressure but also power dynamics and historical contingencies may have led to the current institutions that are observed (Johnson 2004; Blyth et al. 2011; Epstein et al. 2020). Without an awareness of what institutions were adapting to, it is difficult to make inferences about their functionality.

Ostrom and Basurto (2011) present a simple method for documenting changes in institutional arrangements over time. Additionally, several papers have combined a historical framing with the ASN approach, including those associated with the Social Ecological Systems Meta-Analysis Database (SESMAD) project (Fleischman et al. 2014). Epstein et al. (2020) take a similar approach to ASNs as applied to forest policy in Senegal. Finally, in her case study of the Great Barrier Reef governance system, Morrison (2017) describes sequential "regime structures," which roughly align with the ASN concept. Presenting a case as a series of snapshots of an ASN is a promising approach to be built on more in the future.

All of this discussion bears on how we ought to use the framework itself in a case study. Rules include informal norms, which are hard to measure. Power and politics and historical trajectories are also hard to measure quantitatively. This leads the case study approach to be particularly appropriate for the application of the IAD framework, in that it avails itself of multiple data sources in order to uncover frequently illegible factors in a system. Building on the concept of an embedded case study from above, we can present a basic recipe for IAD case studies:

1   A set of research questions associated with social and/or ecological outcomes
2   A focal action situation and an institutional analysis of this situation
3   An analysis of the ASN that contains this focal action situation
4   A historical analysis of the focal situation and the larger ASN, potentially as a series of "snapshots" to trace dynamics over time
5   Data collection via:
    a   Interviews with key stakeholders, which generally would include actors in a focal action situation and key informant interviews with other actors.

   b   Documents that describe formal rules and the case's historical tra-
jectory that led it to the current points.
   c   GIS data sources, which are increasingly publicly available, particu-
larly in wealthier states
6   Mixed-method analysis approaches

An example of an IAD case as a series of ASN snapshots using the steps
outlined above appears in a study on water governance in Kenya (McCord
et al. 2017). The research team examined the performance of the Kenyan
2002 water law reform, which created a form of polycentric water govern-
ance. Specifically, they studied the 25 community water projects (CWPs)
located in the Upper Ewaso Ng'iro basin, that are nestled within five water
user resource associations (WURAs). In total, the team gathered data on 30
action situations, the 25 CWPs, and the 5 WURAs. The CWP action situ-
ations each capture the interactions among individual water users; whereas
the WURAs focus on the interactions among clusters of CWPs. The research
team used surveys of the CWP's executive committees, household level water
flow on a weekly basis, and archival research. They examined whether CWPs
experimented with rule changes, water allocation outcomes, and forms of
coordination among CWPs within their respective WURAs.

Case studies were a foundational method for the development of the
common pool resource research program founded by Elinor Ostrom begin-
ning in the mid-1980s, and they continue to play a critical role in the devel-
opment of theories associated with the IAD framework. Case studies also
provide research questions to be systematically examined using lab and field
experiments as described next (Ostrom, Gardner, and Walker 1994).

Lab and Field Experiments

Laboratory and field experiments are a powerful method for instantiating
action situations and systematically examining the effects of altering a single
dimension of the situation on the outcomes individuals realize (Ostrom,
Gardner, and Walker 1994). Lab and field experiments associated with the
IAD framework examine social dilemmas that arise in providing for public
goods or governing common pool resources and "treatments" for resolving
the dilemmas. Experiments are best used for hypothesis testing and theory
development (Poteete, Janssen, and Ostrom 2010).

A succinct description of an experiment is provided by Poteete, Janssen,
and Ostrom (2010:141):

   the experimenter creates an environment where a number of human
   subjects make decisions in a controlled setting. Human subjects vol-
   untarily consent to participate in an experiment prior to participating.
   They receive instructions on the actions about which they can make
   decisions and outcomes that depend on the decisions of all in the experi-
   ment. Participants make each decision in private by writing it on a paper

form or entering it on a computer. Salient incentives are provided in terms of monetary returns depending on the decisions made, or other relevant rewards.

Beginning with a baseline experiment, rules in use, community characteristics, material and physical conditions, or individuals' decision-making/cognitive processes may be varied and outcomes compared with the baseline experiment. For instance, the baseline common pool resource (CPR) experiment involved eight subjects choosing to invest tokens in an activity that provided low, but stable returns that did not depend on what other subjects chose or invest in a common pool resource, the returns from which were a function of how many tokens all subjects as a group invested (see Ostrom, Gardner, and Walker 1994 for the models). Group returns were maximized if subjects invested just over half of their tokens in the CPR; however, the Nash equilibrium, which by definition means that an individual will make the choice that provides her with the greatest return given the choices of all subjects, predicts that individuals will invest most of their tokens in the CPR. Doing that lowers the group and individual returns. In other words, a social dilemma emerges. From this baseline experiment, many treatments are applied, as discussed below.

Given the high level of control that researchers exercise over the design and implementation of lab experiments, internal validity is quite high. It is further strengthened through replication. Often times, lab experiments are replicated multiple times using different participants. Or, different researchers execute the same experiments. Replication helps rule out unaccounted for factors that may affect the outcomes.

High internal validity is coupled with low external validity for laboratory experiments. One means of addressing this limitation, at least to a certain degree, is to use field experiments. In contrast to lab experiments, which typically take place in university labs with college students as subjects, field experiments are conducted in more naturalistic settings with participants who have experience with harvesting from CPRs or providing for public goods. Examining the choices made by people who, in pursuing their livelihoods, are engaged with social dilemmas provides a more realistic understanding of decision-making.

Field experiments are conducted similarly to lab experiments in that the experimenter creates a controlled decision-making setting. However, the experimenter must take extra precautions to ensure that any possible factors that may affect decisions are accounted for and addressed. This is most commonly accomplished by being knowledgeable of and sensitive to the setting, the culture, and norms and values of the subjects (see Cardenas, Janssen, and Bousquet 2013 for conducting specific field experiments).

The initial CPR experiments focused on communication treatments (Ostrom, Gardner, and Walker 1992, 1994; Ostrom 2005; Janssen 2010) and punishment treatments. Various forms of communication were examined, from a single opportunity for subjects to communicate to communication

prior to each decision round, and from face-to-face communication to texting and other computer-aided communication (Poteete, Janssen, and Ostrom 2010). The findings were consistent and significant, communication supports collective action and allows subjects to devise norms and strategies to limit overinvestment in the CPR. Likewise with punishment. Subjects were allowed to engage in costly punishment, in that they had to pay to impose a fine on another subject (Casari and Luini 2009). A punishment option also supported collective action. Group outcomes (or payoffs) improved, although the outcomes of individuals who chose to punish were negatively impacted (Janssen et al. 2010). When subjects were allowed to voluntarily design and adopt a punishment mechanism, cooperation in investing in the CPR was robust, punishment was used more judiciously, and individual and group outcomes were closer to optimal (Janssen et al. 2010).

A major development in CPR experiments occurred when Marco Janssen developed a spatially explicit, dynamic version of the CPR experiment in which subjects moved avatars across a landscape to harvest "tokens" from a CPR that was renewable and depletable (Janssen and Ostrom 2008; Janssen 2010). This allowed for the exploration of questions related to social–ecological systems, as features of ecological systems could be varied, along with features of the social system. (The software package for the experiment is available at https://gamesforsustainability.org/laboratory-experiments/#foraging_games, and can be downloaded for use).

Initially, treatments similar to the types used in the static CPR experiments were deployed in what are now called foraging games, along with varying the renewal rates of the resource. Findings regarding communication and punishment remained robust. Subjects allowed to communicate and punish, even as renewal rates of the resource were varied, realized and sustained cooperation in harvesting from a CPR (Janssen et al. 2010).

More recently, attention has turned to psychological processes of decision-making in common pool resource settings to explain why individuals cooperate and engage in rule following behavior (DeCaro et al. 2015; Coyle et al. 2018; Baggio et al. 2019; Freeman et al. 2020). DeCaro et al. (2015) used surveys of participants to measure the psychological effects of different experimental interventions. They found that participants who were allowed to adopt harvesting rules, rather than having them externally imposed, and were able to enforce rule compliance, expressed significantly higher levels of fundamental psychological needs of procedural justice, self-determination, and security.

Another line of work examines different forms of individual intelligence. One measures specific intelligence and represents the ability to understand the logics of systems, such as the complex dynamics of a common pool resource system. Another measures general intelligence and represents the ability to work with others. Scholars have found that in CPR experiments, groups that consist of combinations of individuals who score high on either of the cognitive traits perform better at sustaining the resource and they respond more effectively to negative disturbances. The two cognitive traits

work in tandem to support cooperation and sustainability (Coyle et al. 2018; Baggio et al. 2019; Freeman et al. 2020).

Field experiments have also been used to explore differences in decision-making as a function of the gender of the subjects. Cook, Grillos, and Andersson (2019) conducted field experiments among users of forests in Indonesia, Peru, and Tanzania. The experiment required subjects to decide how many trees to protect and how to allocate payments among participants from a payment for ecosystem services program. Groups consisting of at least 50% women were significantly more likely to protect more trees and allocate payments more equitably among participants.

Lab and field experiments have focused on actors' cognitive processes and ascriptive characteristics of subjects, as well as how actors' choices are affected by changes in rules or physical/material conditions that structure action situations. The effort is to contribute to a behavioral theory of collective action (E. Ostrom 1998) that can be used to develop more effective policy interventions.

ABM

ABM is a computational approach for simulating repeated interactions among autonomous and heterogeneous agents to understand systemic outcomes that result from these interactions. Within ABM, agents are endowed with characteristics that inform how they act within the context of prescribed choice sets and systemic parameters (e.g., natural resource constraints; number, characteristics and choices of other actors with whom they are interacting; events). ABM affords many analytical opportunities including the ability to: model intertemporal behavioral dynamics; explicitly model the impacts of cognitive constraints of agents; simultaneously account for variables situating at different analytical scales (i.e., individual/agent characteristics and systemic parameters and outcomes); and ascertain the effects of manipulations to individual or systemic variables on outcomes of interest.

ABM is well suited for analyzing the behavior of individuals within action situations, which themselves represent situations in which two or more actors are interacting within a set of constraints. Earlier work with ABMs focused on the emergence of norms that support cooperation and the resolution of social dilemmas (Poteete, Janssen, and Ostrom 2010). More recently, IAD scholars have used ABMs to study endogenous rule change, or evolution, emerging through collective action within the context of institutionally governed settings (Smajgl et al., 2008, Zellner et al., 2014; Frantz et al., 2015; Ghorbani and Bravo, 2016). In these cases, institutions are in essence part of the outcomes that are being evaluated and are thus necessarily reflecting institutions-in-use. Recently, however, scholars have highlighted how institutions-in-form can be used in the up-front parameterization of ABMs (Frantz and Siddiki, Forthcoming). Informing this approach is the recognition that institutions-in-form, through specific institutional directives, detail

choice sets available to agents. Siddiki and Frantz (2020) use an assessment of institutions-in-form – the US National Organic Program Rule – to support the parameterization of models that simulate interactions among different types of actors interacting within a simulated, regulated organic farming setting. From the actual regulation, they draw information about actors (e.g., farmers, regulatory inspectors, and regulatory certifiers), actions of these actors, constraints on their actions, and sanctions for noncompliance. These types of information map specifically to key ABM simulation parameters.

Finally, IAD scholars have also developed specific analytical platforms for conducting institutional analysis using ABM, or what they refer to as institutional modeling. One example of such an analytical platform is the Modeling Agent Systems based on Institutional Analysis (MAIA) framework developed by Ghorbani et al. (2013).

There are several basic steps to ABM (Ghorbani and Bravo, 2016; Abebe et al., Forthcoming; Nikolic and Ghorbani, 2011): system analysis, model conceptualization, detailed model design, implementation, and model evaluation. These steps are engaged once the modeler has identified the set of research questions that will inform the analytical exercise. In the system analysis step of ABM, the modeler seeks understanding of the system to be simulated, and its key features, including agents, agent attributes, and system features. For institutional analysis specifically, this is the step at which the modeler identifies the array of institutions that will be used to design a model (e.g., when the modeler is using institutions-in-form for model parameterization) or identifies the types of institutions the emergence of which the modeler is interested in evaluating. In the model conceptualization and design phases, the modeler identifies the specific values to assign to modeling features (e.g., types of agents, number of agents, and agent attributes; number of resource units upon which agents are acting). Again, in the context of institutional analysis, some of this information will be ascertained from institutional design or knowledge. In the model implementation phase, the modeler executes the model and, in the model evaluation phase, analyzes modeling results. In analyzing results, modelers typically evaluate a set of metrics that indicate model performance and another set of metrics that are specific to the modeling domain (the latter of which is discussed in more detail below). In the case of institutional analysis, where institutions are used in the ex ante parameterization of the model, the modeler is typically interested in evaluating outcomes associated with institutionally defined choice sets. In ex post assessment of institutions, the modeler is typically interested in describing emergent institutions, as indicated by the behavioral regularities that agents exhibit in the context of prescribed decision-making and systemic constraints.

A unique quality of ABMs relative to some of the other methods covered in this chapter is the level of control that the analyst exerts over the design of the analytical tool as well as the domain that is being studied. Compared to field studies, for example, in which the analyst is designing an instrument to be applied in an existing, real-world setting that is more or less understood

and hardly controlled by the analyst, the agent-based modeler is engaged in constructing the analytical exercise and the setting in which the exercise is to be conducted. Given the extent and breadth of control the analyst has, careful attention must be given to conceptualization and data sources relied on to inform model design. Conceptualization here, in complement to "model conceptualization" as described above, refers to identification of variables, relating to some aspects of a modeled system, for which the modeler will need to assign values.

Various data sources can be used to determine appropriate values for model variables. Zellner et al. (2014) highlight the following: surveys or questionnaires, field and lab experiments, companion modeling, GIS and other spatial data, and ethnographic methods. For values that may be difficult to derive from some type of observation (e.g., cognitive characteristics), values can be determined from theory (Poteete, Janssen, and Ostrom 2010).

Referenced above in association with the "model evaluation" step of ABM is the analysis of metrics that are specific to the modeling domain. These are the metrics that the modeler evaluates to respond to research questions that motivate his/her analytical exercise. Operationally, this evaluation is typically conducted using statistical methods. Descriptive methods are used to characterize behavioral or statistical dynamics that are realized through the modeling exercise. Diagnostic statistical methods are used to determine specific relationships between variables included in the models.

ABMs complement case studies and lab and field experiments. They can be used to examine and evaluate "how well alternative models of human behavior and social interactions account for empirical observations" identified in case studies and lab and field experiments (Poteete, Janssen, and Ostrom 2010:194). They can also be used to examine the robustness of institutional arrangements, as well as how institutions evolve.

Network Analysis

Networks, which are links or ties among actors and objects, can be thought of as individual or linked action situations. Thus, network analysis, like the methods discussed above, can be used to examine patterns of interaction among actors within and across action situations, allowing analysts to explore the patterns of interaction associated with emerging, existing, or resolved social dilemmas, as well as the design and structure of the configurations of rules shaping action situations.

The application of network analysis within an IAD framework study represents a relatively new line of research, although policy network analysis is well developed and extensively used within the subfield of policy studies (e.g., Schneider et al. 2003). IAD framework scholars incorporate forms of data and methods of analysis that are distinct from well-developed lines of research using policy network analysis (Berardo and Scholz 2010; Henry 2011; Ingold 2011; Yi and Scholz 2015; Herzog and Ingold 2019). Policy network analysis foregrounds perceived interactions among actors,

i.e., behavior; while setting aside the role of institutional arrangements in shaping actors interactions (Olivier and Schlager 2021). In contrast, the emerging work out of the IAD framework explores the structure and design of institutional arrangements and the interactions of actors and ecosystems in the context of social–ecological systems.

The analysis of institutional designs relies on data developed from the coding of institutional statements using the grammar of institutions (Crawford and Ostrom, 1995; Basurto et al. 2010; Siddiki et al. 2011; Olivier 2019). The focus has been on rules-in-form that are found in laws, regulations, treaties, and intergovernmental agreements (Brady 2020; Hanlon et al. 2019). Rules prescribe interactions among actors, creating ties between them. For instance, the institutional statement

> The project developer or the person commissioned by the project developer to carry out the work shall notify the competent authority without delay of any unintentional groundwater development pursuant to Article 49(2) of the Federal Water Act.

creates an information link between a project developer (called an attribute) and a competent authority (called an object, who is the receiver of the action). Configurations of institutional statements create networks of prescribed interactions. Action situations are configurations of institutional statements; thus, it is possible to explore the structure and patterns of the institutional scaffolding of action situations using social network analysis tools (Olivier 2019; Herzog et al. 2020; Olivier et al. 2020; Olivier and Schlager 2021).

Identifying action situations from laws, regulations, treaties, and intergovernmental agreements relies on identifying distinct programs and collective-choice venues. For instance, Schlager et al. (2021) identified action situations by heading and subheading titles from the intergovernmental agreement and program rules constituting the New York City Watersheds Governing System. Within each of the collective choice and public goods programs, monitoring, conflict resolution, and graduated sanctioning mechanisms unique to each program were identified as well. Thus, each program and venue consisted of linked action situations, and the programs and venues themselves were linked through the intergovernmental agreement, to create a regional scale, polycentric system of water quality governance.

Networks can be thought of as individual or linked action situations and, like action situations, can be treated as independent or dependent variables. Is the effort to explain the different interactions among actors and outcomes realized in a network/action situation? Or is the effort to explain the structure of the network/action situation? If the former, then network metrics such as centrality, density, and reciprocity may be used to describe the patterns of ties among nodes. If the latter, then different forms of modeling, such as exponential random graph models may be used to identify the endogenous and exogenous factors that play a statistically significant

role in reproducing the observed network (Robins et al. 2007; Cranmer and Desmarais 2011).

The questions addressed in this emerging line of research focus on empirically testing theories about institutional design. As E. Ostrom (1999, 2005, 2007, 2009) has repeatedly noted, the IAD framework is compatible with different theories that center around explaining the design and performance of institutional arrangements, such as transaction costs theory (Williamson 1985), local public economies (Oakerson 1999), common pool resource theory (E. Ostrom 1990), federalism theory (V. Ostrom 2007), among others. For instance, the local public contracting literature (Brown and Potoski 2003; Hefetz and Warner 2012) theorizes that the design of institutional arrangements for the provision of public goods will vary as a function of the type of good. And that the institutional arrangements for the provision of public goods will differ from the design of institutional arrangements for shared decision-making venues designed to support credible commitments and cooperation among actors (Miller 1992). Initial findings from research testing hypotheses from local public contracting literature suggest that the design of credible commitment arrangements is distinct from that of the provision of public goods (Oliver, Scott, and Schlager 2020; Olivier and Schlager 2021; Schlager, Bakkensen, Olivier, and Hanlon 2021). Moving forward attention is being given to incorporating networks of prescribed interactions and networks of perceived interactions or behavior into the same social network analysis model to explore the effects of institutional arrangements on the types of ties actors choose to create (Herzog et al. 2020; Olivier and Schlager 2021).

A second line of research uses network analysis to examine the ties among actors and ecosystems using multiplex network analysis (see Bodin et al. 2019 for best practices). Baggio et al. (2016), for instance, examine relations among households and the sharing of wild foods occurring in three indigenous Alaskan villages. The networks are constructed among households and types of wild foods. Households are linked to one another through sharing relations, and households are linked to the species they hunt. The scholars examined the robustness of the networks to different disturbances, such as the loss of access to a specie, or the out migration of households, by the removal of nodes in the networks. Baggio et al. (2016) found that the removal of social ties had greater negative impacts on communities than the removal of a specie, suggesting that the communities were more vulnerable to social disturbances compared to ecological disturbances. Multiplex or multilevel network analysis has also been used to explore the effects of scale mismatches between the boundaries of governing arrangements and the boundaries of ecological systems on the recovery of endangered species (Sayles and Baggio 2017). By incorporating organizations engaged in species recovery and ecological systems within a single network model, the patterns of interactions between actors and ecosystems can be measured and related to specific outcomes. This line of work has the potential to open up new lines of research on polycentric systems.

Network analysis provides another method of exploring the design of institutional arrangements and the patterns of perceived interactions that emerge as a result.

Methods Summary

The starting point for research grounded in the IAD framework is identifying relevant action situations, actors' interactions, and outcomes (see Table 8.1). With experiments and agent–based models, the researcher creates the action situation(s), observes choices and actions taken by actors, and the resulting outcomes. With case studies and network analyses, researchers identify action situations in the field and interrogate actors about their perceptions, values, choices, and understandings of the situations. At the same time, the physical/material conditions and the rules that structure the action situations are attended to. Oftentimes, it is the mismatch between physical/material conditions and institutional arrangements that trigger social dilemmas and efforts to engage in collective action to devise or adjust governing arrangements to resolve the dilemmas. Thus, research design centers on defining, identifying, and situating action situations, and observing and measuring interactions and outcomes among actors (Table 8.1, Columns 2 and 3).

Each method makes different demands on the researcher. Some methods make use of conventional analysis techniques, such as descriptive statistics, but require advanced modeling skills to establish the action situations, e.g., experiments, ABMs (see Table 8.1, Columns 4 and 5). Fortunately, scholars have begun to create repositories of models or coding packages to develop models that are freely available for others to use (see Table 8.1, Column 5). Other methods make use of conventional data gathering techniques, such as surveys and interviews, but require advanced data analysis skills (social network analysis). Most IAD framework scholars are expert in one or two of the methods, but not all. Poteete, Janssen, and Ostrom (2010) noted different teams of scholars with expertise encompassing these diverse methods and argued that progress in developing behavioral theories of collective action depends on teams of experts working together. In this way, it is possible for scholars to take advantage of the methods in different ways. Insights and data from case studies can be used to develop agent-based models, or the structure and patterns of interactions among actors revealed by network analyses may suggest questions best pursued through the use of case studies.

Moving Institutional Analysis Forward

Although the IAD framework is relatively complex with many interacting categories of concepts and variables operating at multiple levels of action, the core of the framework centers on institutions and how institutions condition behavior. That is not to dismiss physical/material conditions or community characteristics. Indeed, an IAD framework application would be incomplete without accounting for both. But institutions are the tools that actors have at

Table 8.1 Comparison across methods

Method	Conceptualization	Data Collection and Measurement	Analyzing Data	Major Data Sources
Cases	Systematic comparisons of action situations in the field	Field-based interviews combined with secondary data (legal documents)	Mixed-method, qualitative interpretation of interview data	SES Library, https://seslibrary.asu.edu/; Digital Library of the Commons, http://dlc.dlib.indiana.edu/dlc/; SES meta-analysis database, https://sesmad.dartmouth.edu/
Lab/field experiments	Create action situations in lab or field settings that represent different social dilemmas	Actors choices Group outcomes	Frequencies Descriptive statistics Regression models	https://gamesforsustainability.org/laboratory-experiments/
ABM	Patterns of interactions among actors in simulated action situations	Modeling of actor attributes and actor choice sets based on observed behavior, formal institutions, and theory	Descriptive and statistical analysis of behavioral trends	Network for Computational Modeling in Social–Ecological Sciences, www.comses.net/
Network analysis	Patterns of interactions within and across action situations	Surveys Grammar coding of legal documents	Quadratic Assignment Procedure Exponential Random Graph Models	None

their disposal to address social dilemmas by creating governing arrangements matched to a setting (Ostrom 2005).

Fundamentally, understanding social dilemmas, collective action, and related institutional dynamics requires approaches for reliably describing institutions. The IAD framework offers numerous approaches for classifying institutions which also make it especially well suited for the comparative study of institutions (Ostrom, 2005). For example, it offers guidance on differentiating between constitutional, collective choice, and operational "levels of action." Institutions are classified as linking to different levels of action based on their role in structuring day-to-day or operational activities; or structuring rule making, monitoring, and compliance activities (collective choice, constitutional choice). The framework also offers a "rule typology" that differentiates rules according to their functional properties. Yet another approach is the Institutional Grammar (IG) (Crawford and Ostrom, 1995). A core assumption of the IG is that institutional statements exhibit common structural elements, or a syntax, with individual syntactic elements corresponding to distinctive institutional information embedded in statements, and which combine to form a "grammar" of institutions. IG-based assessments of institutional design can yield insights on how constituent parts of institutions link together to govern behavior.

Moving forward, the further development of the IAD framework and associated theories will likely center on the different approaches for classifying institutions. In particular, complex, multilevel forms of governance, grounded in levels of action, present one promising line of research. These studies move beyond a focus on how single communities, organizations, or collections of resource users address social dilemmas, typically in a CPR context, to explore how social–ecological systems, or complex adaptive systems, perform (Dietz et al. 2003; Ostrom 2007; Anderies et al. 2016). Multiple linked action situations can be studied as dynamic systems or through static comparative analyses (Hanlon et al. 2019; Janssen et al. 2019). Studies of linked action situations as systems focus on feedback processes (Anderies et al. 2004; Anderies et al. 2016) and the stability properties and robustness of such systems. Case studies and ABM have been used to study system-level dynamics. For instance, Janssen et al (2019) examine the case study of the US highway system, examining feedback links between users; hard infrastructure; institutional arrangements, especially financing arrangements; and decision-makers using 30 plus years of data. By treating linked action situations as systems whose dynamics are a function of feedback loops, it is possible to identify whether feedback processes are aligned in such a way as to support system robustness (Muneepeerakul and Anderies 2017, 2020; Janssen et al. 2019). Agent-based models can also be used to explore system robustness. For instance, van Strien et al. (2019) develop an agent-based model simulating land-use change in an alpine mountain region in the Canton of Valais, Switzerland. They explore the behavior of the system in response to changing external stressors, such as changes in political, economic, and environmental conditions (i.e., external system stressors). Van Strien et al. (2019)

further explored the system effects of different policy interventions and demonstrated how those effects differ as a function of the initial conditions. As Poteete, Janssen, and Ostrom (2010) and van Strien et al. (2019) note, much work remains to further the systematic use of agent-based models in exploring social dilemmas, from developing agreed upon means for identifying and exploring the stability and instability boundaries of such models, to identifying best practices for using empirical data from cases studies and experiments to parameterize agent-based models, among other issues. As van Strien et al. (2019) point out, many agent-based models are one offs that contribute little to the accumulation of knowledge.

Another promising line of research on institutional arrangements centers on the institutional grammar that can be used to explore a variety of questions, from the legitimacy of institutional arrangements, to their design, to how institutional arrangements condition behavior in a multitude of different settings. A growing body of work using the institutional grammar for data collection and theory development is emerging. However, for this institutional classification tool to be more widely used, and to realize its full potential, will require both machine coding of institutional statements, and the development of theoretically grounded and policy-relevant approaches for analyzing the configurations of statements constituting policies, laws, and regulations. Machine coding will provide new data sets of institutions, grounded in the rule typology and grammar of institutions that can fuel work on institutional design and dynamics. Developing relevant applications of such data will allow for the development of theories of institutional change and of complex adaptive systems. Breakthroughs in each of these areas are likely through the concerted effort of networks of scholars, such as the institutional grammar research initiative (https://institutionalgrammar.org/).

Conclusion

The IAD framework was first published in 1982 (Kiser and Ostrom 1982). It provided the foundation for a research program on the governance of common pool resources initiated in 1985, that resulted in the awarding of a Nobel Prize in Economics to Elinor Ostrom in 2009. At the framework's core is the action situation, which allows analysts to systematically study interdependent behavior and collective action. From the beginning of that research program, multiple methods, appropriate for examining interdependent behavior, were used (Ostrom, Gardner, and Walker 1994). Initially, case studies and laboratory experiments were used for theory development and testing as well as sources of data to be used with other methods.

The case study is the most versatile tool in the IAD framework methods kit and is used to examine action situations, the components that structure action situations, and links among action situations, across a range of settings. In contrast, lab experiments have been used for theory testing and development by examining the effects of specific treatments, i.e., variations in the rules structuring action situations, on individual choice and group outcomes.

Addressing the limitations of lab experiments, scholars have developed dynamic experiments allowing for feedback effects. Just as important have been the development of field experiments, which allow resource users to participate in an attempt to resolve social dilemmas. Not only has this work begun to address external validity limitations of lab experiments, but it has also been used to allow resource users to experiment with different types of rules that they then may consider adopting in the governance of the common pool resources they depend on for their livelihoods (Meinzen-Dick et al. 2018).

Agent-based models and social network analysis are more recently adopted methods used for the exploration and analysis of institutional arrangements and behavior. Drawing on data from case studies, experiments, and institutional data collected through the application of the institutional grammar, ABMs provide means of exploring settings and interventions that may not have direct empirical analogs. They also allow analysts to identify and explore the robustness properties of dynamic systems. Social network analysis may be used to identify theoretical and policy-relevant patterns of institutional design and behavior that allow for theory testing as well as policy designs.

Although the IAD framework has been in use for almost four decades, it remains a vibrant framework, not just because it is compatible with multiple theories and continues to support theory development, but because it is compatible with multiple methods that allow policy analysts to explore and explain the diverse dimensions and factors affecting action situations and the possibilities of collective action.

References

Abebe, Yared Abayneh, Amineh Ghorbani, Igor Nikolic, Zoran Vojinovic, and Arlex Sanchez. 2019. "A Coupled Flood-Agent-Institution Modelling (CLAIM) Framework for Urban Flood Risk Management." Environmental Modelling and Software 111: 483-492.

Acheson, James M. 1988. *The Lobster Gangs of Maine*. Hanover: University Press of New England

Anderies, John M., Marco A. Janssen, and Elinor Ostrom. 2004. "A Framework to Analyze the Robustness of Social-Ecological Systems from an Institutional Perspective." *Ecology and Society* 9(1): 18.

Anderies, John M., Marco A. Janssen, and Edella Schlager. 2016. "Institutions and the Performance of Coupled Infrastructure Systems." *International Journal of the Commons* 10(2): 495–516.

Baggio, Jacopo, Allain Barnett, Irene Perez-Ibarra, Ute Brady, Elicia Ratajczyk, Nathan Rollins, Cathy Rubiños, Hoon C. Shin, David J. Yu, Rimjhim Aggarwal, John M. Anderies, and Marco A. Janssen. 2016. "Explaining Success and Failure in the Commons: The Configural Nature of Ostrom's Institutional Design Principles." *International Journal of the Commons* 10(2): 417–439.

Baggio, Jacopo A., Jacob Freeman, Thomas R. Coyle, Tam The Nguyen, Dale Hancock, Karrie E. Elpers, Samantha Nabity, HJ Francois Dengah II, and David Pillow. 2019. "The Importance of Cognitive Diversity for Sustaining the Commons." *Nature Communications* 10(875): 1–11

Basurto, Xavier, Gordon Kingsley, Kelly McQueen, Mshadoni Smith, and Christopher M. Weible. 2010. "A Systematic Approach to Institutional Analysis: Applying Crawford and Ostrom's Grammar." *Political Research Quarterly* 63(3): 523–553.

Berardo, Ramiro, and John Scholz. 2010. "Self-Organizing Policy Networks: Risk, Partner Selection, and Cooperation in Estuaries." *American Journal of Political Science* 54(3): 632–649. https://doi.org/10.1111/j.1540-5907.2010.00451.x.

Blyth, Mark, Geoffrey M. Hodgson, Orion Lewis, and Sven Steinmo. 2011. "Introduction to the Special Issue on the Evolution of Institutions." *Journal of Institutional Economics* 7(3): 299–315. https://doi.org/10.1017/S174413741 1000270.

Bodin, Örjan, Steven M. Alexander, Jacopo Baggio, Michele L. Barnes, Ramiro Berardo, Graeme S. Cumming, Laura E. Dee, Paige Fischer, Manuel Fischer, Maria Mancilla Garcia, Angela M. Guerrero, Karin Ingold, Jacob Hileman, Petr Matous, Tiffany Morrison, Jeremy Pittman, Garry Robins, and Jesse S. Sayles. 2019. "Improving Network Approaches to the Study of Complex Social-Ecological Interdependencies." *Nature Sustainability* 2(7): 551–559.

Boyd, Robert. 2017. *A Different Kind of Animal: How Culture Transformed Our Species*. Princeton, NJ: Princeton University Press.

Brady, Ute. 2020. Robust Conservation Anarchy: Comparing Treaty Institutional Design for Evidence of Ostrom's Design Principles, Fit, and Polycentricity. Doctoral Dissertation.

Brown, Trevor L., and Matthew Potoski. 2003. "Managing Contract Performance: A Transaction Costs Approach." *Journal of Policy Analysis and Management* 22(2): 275–297.

Cardenas, Juan-Camilo, Marco Janssen, and Francois Bousquet. 2013. "Dynamics of Rules and Resources: Three New Field Experiments on Water, Forests, and Fisheries." In John A. List and Michael K. Price (eds) *Handbook on Experimental Economics and the Environment*. Cheltenham: Edward Elgar Publishing. pp. 319–345.

Casari, Marco, and Luigi Luini. 2009. "Cooperation Under Alternative Punishment Institutions: An Experiment." *Journal of Economic Behavior and Organization* 71(2): 273–282.

Clement, Floriane. 2010. "Analysing Decentralised Natural Resource Governance: Proposition for a 'Politicised' Institutional Analysis and Development Framework." *Policy Sciences* 43(2): 129–156. https://doi.org/10.1007/s11077-009-9100-8.

Cook, Nathan J., Tara Grillos, and Krister P. Andersson. 2019. "Gender Quotas Increase the Equality and Effectiveness of Climate Policy Interventions." *Nature Climate Change* 9: 330–334.

Cox, Michael. 2014. "Applying a Social–Ecological System Framework to the Study of the Taos Valley Irrigation System." *Human Ecology* 42(2): 311–24. https://doi.org/10.1007/s10745-014-9651-y.

Cox, Michael, Gwen Arnold, and Sergio Villamayor Tomás. 2010. "A Review of Design Principles for Community-based Natural Resource Management." *Ecology & Society* 15(4): 38.

Cox, Michael, Frederick Payton, Tyler Pavlowich, Caroline Resor, Margaret Wilson, and Nina Dewees 2018. "The Dominican Fishery of Manzanillo: A Coastal System in Transition." *Ocean & Coastal Management*, Coastal Systems in Transition, 162 (August): 170–180. https://doi.org/10.1016/j.ocecoaman.2017.06.018.

Coyle, Thomas R., Karrie E. Elpers, Miguel C. Gonzalez, Jacob Freeman, and Jacopo A. Baggio. 2018. "General Intelligence(g), ACT Scores and Theory of Mind:

(ACT)g Predicts Limited Variance among Theory of Mind Tests." *Intelligence* 71: 85–91.

Cranmer, Skyler J., and Bruce A. Desmarais. 2011. "Inferential Network Analysis With Exponential Random Graph Models." *Political Analysis* 19(1): 66–86.

Crawford, Sue ES, and Elinor Ostrom. 1995. "A Grammar of Institutions." *American Political Science Review* 89(3): 582–600.

DeCaro, Daniel A., Marco A. Janssen, and Allen Lee. 2015. "Synergistic Effects of Voting and Enforcement on Internalized Motivation to Cooperate in a Resource Dilemma." *Judgement and Decision Making* 10(6): 511–537.

Dietz, Thomas, Elinor Ostrom, and Paul C. Stern. 2003. "The Struggle to Govern the Commons." *Science* 302: 1907–1912.

Epstein, Graham, Tiffany H. Morrison, Aaron Lien, Georgina G. Gurney, Daniel H. Cole, Martin Delaroche, Sergio Villamayor Tomas, Natalie Ban, and Michael Cox. 2020. "Advances in Understanding the Evolution of Institutions in Complex Social-Ecological Systems." *Current Opinion in Environmental Sustainability* 44(June): 58–66. https://doi.org/10.1016/j.cosust.2020.06.002.

Fleischman, Forrest, Natalie Ban, Louisa Evans, Graham Epstein, Gustavo Garcia-Lopez, and Sergio Villamayor Tomas. 2014. "Governing Large-Scale Social-Ecological Systems: Lessons from Five Cases." *International Journal of the Commons* 8(2): 428–456. www.thecommonsjournal.org/article/view/416/.

Frantz, Christopher K. and Saba Siddiki. Forthcoming. *Institutional Grammar: Foundations and Applications for Institutional Analysis*. London: Palgrave Macmillan.

Frantz, Christopher K, Martin K. Purvis, Bastin Tony Roy Savarimuthu, and Mariusz Nowostawski, M. 2015. "Modeling Dynamic Normative Understanding in Agent Societies." *Scalable Computing: Practice and Experience* 16(4): 355-380.

Frantz, Christopher, Martin K. Purvis, and Mariusz Nowostawski. 2014. "Agent Based Modeling in Information Transmission in Early Historic Trading." *Social Science Computer Review* 32(3): 393–416.

Freeman, Jacob, Jacopo A. Baggio, and Thomas R. Coyle. 2020. "The Benefits of Cognitive Diversity for Sustaining Natural Resources." *Proceedings of the National Academy of Sciences* 117(14): 7712–7718.

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

Ghorbani, Amineh, Pieter Bots, Virginia Dignum, and Gerard Dijkema. 2013. "MAIA: A Framework for Developing Agent-Based Social Simulations." *Journal of Artificial Societies and Social Simulation* 16(2): 9.

Ghorbani, Amineh, and Giangiacomo Bravo. 2016. "Managing the Commons: A Simple Model of the Emergence of Institutions through Collective Action." *International Journal of the Commons* 10(1): 200–2019.

Gould, Stephen Jay, and Richard C. Lewontin. 1979. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society* 205(1161): 581–598. https://doi.org/10.1098/rspb.1979.0086.

Hanlon, Jeffrey, Tomás Olivier, and Edella Schlager. 2019. "Suspicious Collaborators: How Governments in Polycentric Systems Monitor Behavior and Enforce Public Good Provision Rules Against One Another." *International Journal of the Commons* 13(2): 977–992.

Hardy, Scott D., and Tomas M. Koontz. 2009. "Rules for Collaboration: Institutional Analysis of Group Membership and Levels of Action in Watershed

Partnerships." *Policy Studies Journal* 37(3) 393–414. https://doi.org/10.1111/j.1541-0072.2009.00320.x.

Henry, Adam D., 2011. "Ideology, power, and the structure of policy networks." *Policy Studies Journal*, 39(3):361–383. doi:10.1111/j.1541-0072.2011.00413.x

Hefetz, Amir, and Mildred E. Warner. 2012. "Contracting or Public Delivery?" *Journal of Public Administration Research and Theory* 22: 289–317.

Henry, Adam D. 2011. "Ideology, Power, and the Structure of Policy Networks." *Policy Studies Journal* 39(3): 361–383.

Herzog, Laura M., and Karin Ingold. 2019. "Threats to Common-Pool Resources and the Importance of Forums: On the Emergence of Cooperation in CPR Problem Settings." *Policy Studies Journal* 47(1): 77–113.

Herzog, Laura, Karin Ingold, and Edella Schlager, E. 2021. "Prescribed by Law and Therefore Realized? Analyzing Rules and their Implied Actor Interactions as Networks. *Policy Studies Journal*. Early view. https://doi.org/10.1111/psj.12448

Hinkel, Jochen, Michael E. Cox, Maja Schlüter, Claudia R. Binder, and Thomas Falk. 2015. "A Diagnostic Procedure for Applying the Social-Ecological Systems Framework in Diverse Cases." *Ecology and Society* 20(1): 32. https://doi.org/10.5751/ES-07023-200132.

Hodgson, Geoffrey M. 2007. "Evolutionary and Institutional Economics as the New Mainstream?" *Evolutionary and Institutional Economics Review* 4(1): 7–25. https://doi.org/10.14441/eier.4.7.

Ingold, Karin. 2011. "Network Structures within Policy Processes: Coalitions, Power, and Brokerage in Swiss Climate Policy." *Policy Studies Journal* 39(3): 7–25.

Janssen, Marco. 2010. "Introducing Ecological Dynamics into Common-Pool Resource Experiments." *Ecology & Society* 15(2): 7.

Janssen, Marco A., John M. Anderies, Andres Baeza, Hanna L. Breetz, Tomasz Jasinski, Hoon C. Shin, and Sechindra Vallury. 2019. "Highways as Coupled Infrastructure Systems: An Integrated Approach to Address Sustainability Challenges." *Sustainable and Resilient Infrastructure* 4:1–12. doi: 10.1080/23789689.2019.1708181.

Janssen, Marco A., Robert Holahan, Allen Lee, and Elinor Ostrom. 2010. "Lab Experiments for the Study of Social-Ecological Systems." *Science* 328: 613–617.

Janssen, Marco A., and Elinor Ostrom. 2008. "TURFs in the Lab: Institutional Innovation in Dynamic Interactive Spatial Commons." *Rationality and Society* 20: 371–397.

Jaspers, Jelle D. 2017. "Managing Cartels: how Cartel Participants Create Stability in the Absence of Law." *European Journal on Criminal Policy and Research* 23: 319–335. https://doi.org/10.1007/s10610-016-9329-7.

Johnson, Craig. 2004. "Uncommon Ground: The 'Poverty of History' in Common Property Discourse." *Development and Change* 35(3): 407–434. https://doi.org/10.1111/j.1467-7660.2004.00359.x.

Kiser, Larry L., and Elinor Ostrom. 1982. "The Three Worlds of Action: A Metatheoretical Synthesis of Theoretical Approaches." In Elinor Ostrom (ed) *Strategies of Political Inquiry*. Beverly Hills, CA: Sage Publications. pp. 179–222.

Knight, Jack. 1992. *Institutions and Social Conflict*. Cambridge: Cambridge University Press.

Lubell, Mark. 2013. "Governing Institutional Complexity: The Ecology of Games Framework." *Policy Studies Journal* 41(3): 537–559. https://doi.org/10.1111/psj.12028.

Matsueda, Ross L. 2006. "Differential Social Organization, Collective Action, and Crime." *Crime, Law, and Social Change* 46(1–2): 3–33. doi: 10.1007/s10611-006-9045-1.

McCord, Paul, Jampel Dell'Angelo, Elizabeth Baldwin, and Tom Evans. 2017. "Polycentric Transformation in Kenyan Water Governance." *Policy Studies Journal* 45(4): 633–658.

McGinnis, Michael. D. 2011. "Networks of Adjacent Action Situations in Polycentric Governance." *Policy Studies Journal* 39(1): 51–78. https://doi.org/10.1111/j.1541-0072.2010.00396.x.

Meinzen-Dick, Ruth, Marco A. Janssen, Sandeep Kandikuppa, Rahul Chaturvedi, Kaushalendra Rao, and Sophie Theis. 2018. "Playing Games to Save Water: Collective Action Games for Groundwater Management in Andhra Pradesh, India." *World Development* 107: 40–53. https://doi.org/10.1016/j.worlddev.2018.02.006.

Miller, Gary. 1992. *Managerial Dilemmas: The Political Economy of Hierarchy*. Cambridge: Cambridge University Press.

Morrison, Tiffany H. 2017. "Evolving Polycentric Governance of the Great Barrier Reef." *Proceedings of the National Academy of Sciences of the United States of America* 114(15): E3013–E3021. https://doi.org/10.1073/pnas.1620830114.

Muneepeerakul, Rachata, and John M. Anderies. 2017. "Strategic Behaviors and Governance Challenges in Social-Ecological Systems." *Earth's Future* 5(8): 865–876.

Muneepeerakul, Rachata, and John M. Anderies. 2020. "The Emergence and Resilience of Self-Organized Governance and Coupled Infrastructure Systems." *Proceedings of the National Academy of Sciences* 117(9): 4617–4622. https://doi.org/10.1073/pnas.1916169117.

Nikolic, Igor and Amineh Ghorbani. 2011. "A Method for Developing Agent-Based Models of Socio-Technical Systems." International Conference on Networking, Sensing and Control. Delft, 44-49.

Nigussie, Zerihun, Atsushi Tsunekawa, Nigussie Haregeweyn, Enyew Adgo, Logan Cochrane, Anne Floquet, and Steffen Abele. 2018. "Applying Ostrom's Institutional Analysis and Development Framework to Soil and Water Conservation Activities in North-Western Ethiopia." *Land Use Policy* 71: 1–10. www.sciencedirect.com/science/article/pii/S0264837717312024.

North, Douglass. 1990. *Institutions, Institutional Change, and Economic Development*. Cambridge: Cambridge University Press.

Oakerson, Ronald, J. 1999. *Governing Local Public Economies*. San Francisco: ICS Press.

Olivier, Tomás. 2019. "How Do Institutions Address Collective-Action Problems? Bridging and Bonding in Institutional Design." *Political Research Quarterly* 72(1): 162–176.

Olivier, Tomás, and Edella Schlager. 2021. "Rules and the Ruled: Understanding Patterns of Institutional Design and Behavior in Complex Polycentric Systems." *Policy Studies Journal*. Early View. https://doi.org/10.1111/psj.12429

Olivier, Tomás, T. Scott, and E. Schlager. 2020. "Institutional Design and Complexity: Protocol Network Structure in Response to Different Collective-Action Dilemmas. In Manuel Fischer and Karin Ingold (eds) *Networks in Water Governance. Palgrave Studies in Water Governance: Policy and Practice*. London: Palgrave Macmillan.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Ostrom, Elinor. 1998. "A Behavioral Approach to the Rational Choice Theory of Collective Action." *American Political Science Review* 92(1): 1–22.

Ostrom, Elinor. 1999. "Coping with Tragedies of the Commons." *Annual Review of Political Science* 2: 493–535.

Ostrom, Elinor. 2005. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.

Ostrom, Elinor. 2009. "Institutional Rational Choice: An Assessment of the Institutional Analysis and Development Framework." In Paul A. Sabatier (ed) *Theories of the Policy Process.* Second edition. Boulder, CO: Westview Press. pp. 21–64.

Ostrom, Elinor, and Xavier Basurto. 2011. "Crafting Analytical Tools to Study Institutional Change." *Journal of Institutional Economics* 7(3): 317–43.

Ostrom, Elinor, Roy Gardner, and James Walker. 1994. *Rules, Games, and Common Pool Resources*. Ann Arbor, MI: University of Michigan Press.

Ostrom, Vincent. 2007. *The Political Theory of a Compound Republic*. Third edition. Lanham, MD: Lexington Books.

Poteete, Amy R., Marco A. Janssen, and Elinor Ostrom. 2010. *Working Together*. Princeton, NJ: Princeton University Press.

Robins, Garry, Tom Snijders, Peng Wang, Mark Handcock, and Philippa Pattison. 2007. "Recent Developments in Exponential Random Graph (p⋆) Models for Social Networks." *Social Networks* 29(2): 192–215. https://doi.org/10.1016/j.socnet.2006.08.003.

Sayles, Jesse S., and Jacopo A. Baggio. 2017. "Social-Ecological Network Analysis of Scale-Mismatches in Estuary Watershed Restoration." *Proceedings of the National Academy of Sciences* 114(10): E1776–E1785.

Schlaepfer, Martin A., Michael C. Runge, and Paul W. Sherman. 2002. "Ecological and Evolutionary Traps." *Trends in Ecology & Evolution* 17(10): 474–480.

Schlager, Edella, Laura Bakkensen, Tomás Olivier, and Jeffrey Hanlon. 2021. "Institutional design for a complex commons: Variations in the design of credible commitments and the provision of public goods." *Public Administration* 99:263-289.

Schneider, Mark, John Scholz, Mark Lubell, Denisa Mindruta, and Matthew Edwardsen. 2003. "Building Consensual Institutions: Networks and the National Estuary Program." *American Journal of Political Science* 47: 143–158. https://doi.org/10.1111/1540-5907.00010.

Sendzimir, Jan, Zsuzsana Flachner, Claudia Pahl-Wostl, and Christian Knieper. 2010. "Stalled Regime Transition in the Upper Tisza River Basin: The Dynamics of Linked Action Situations." *Environmental Science & Policy* 13(7): 604–619. https://doi.org/10.1016/j.envsci.2010.09.005.

Siddiki, Saba, and Christopher Frantz. 2020. "Understanding the Effects of Social Value Orientations in Shaping Regulatory Outcomes through Agent Based Modeling: An Application to Organic Farming." Working Paper.

Siddiki, Saba, Christopher M. Weible, Xavier Basurto, and John Calanni. 2011. "Dissecting Policy Designs: An Application of the Institutional Grammar Tool." *Policy Studies Journal* 39(1): 79–103.

Smajgl, Alex, Luis R. Izquierdo, and Marco Huigen. 2008. "Modeling Endogenous Rule Changes in an Institutional Context: The ADICO Sequence." *Advanced in Complex Systems* 11(2): 199–215.

Van Strien, Maarten J., Sibyl H. Huber, John M. Anderies, and Adrienne Grêt-Regamey. 2019. "Resilience in Social-Ecological Systems: Identifying Stable and Unstable Equilibria with Agent-Based Models." *Ecology and Society* 24(2): 8.

Williamson, Oliver, E. 1985. *The Economic Institutions of Capitalism*. New York, NY: Simon and Schuster.

Yi, Hongtao, and John Scholz. 2015. "Policy Networks in Complex Governance Subsystems: Observing and Comparing Hyperlink, Media, and Partnership Networks." *Policy Studies Journal* 44(3): 248–279. https://doi.org/10.1111/psj.12141.

Yin, Robert K. 2014. *Case Study Research: Design and Methods*. London: Sage.

Zellner, Moira, Cristy Watkins, Dean Massey, Lynne Westphal, Jeremy Brooks, and Kristen Ross. 2014. "Advancing Collective Decision-Making Theory with Integrated Agent-Based Modeling and Ethnographic Data Analysis: An Example in Ecological Restoration." *Journal of Artificial Societies and Social Simulation* 17(4): 11.

9 Methodological Approaches to the Ecology of Games Framework

Mark Lubell, Matthew Hamilton, Jack Mewhirter,
Francesca Vantaggiato, and Ramiro Berardo

Introduction

We begin with an assertion: complexity is a universal fact of life in governance systems. Policy researchers should only need a moment of introspection to recognize the many policy issues, policy actors, and policy forums that interact in a particular policy sector or region. The assertion that complexity is universal is supported by decades of research on polycentric governance systems – "a social system of many decision centers having limited and autonomous prerogatives and operating under an overarching set of rules" (Aligica and Tarko 2012: 237). Indeed, the concept of complexity has been an enduring thread in policy research (Cairney 2012) and more recent research on social–ecological systems (Ostrom 2009). The complexity of governance systems has a deep kinship with the overarching idea of self-organization in complex adaptive systems (Levin et al. 2013), which has long been an important topic in the natural sciences.

The ecology of games framework (EGF) offers a conceptual framework for developing theory that is applicable to complex, polycentric governance systems. The EGF builds on the metaphor introduced by sociologist Norton Long (1958) to conceptualize urban politics. The EGF argues that polycentric systems are composed of policy actors participating in policy forums where they deliberate about a set of interlinked policy issues or collective-action problems. Decision-making in policy forums is governed by what Ostrom would call "collective-choice rules" (Schlager and Ostrom 1992). The resulting constellation of interdependencies, which can be depicted as a network, constitutes a system of policy games in which actors make strategic decisions. As with Ostrom, the outputs of the system are sets of "operational rules" governing how people use resources or otherwise make on-the-ground decisions (Schlager and Ostrom 1992). However, the overall set of operational rules does not emerge from just one single collective-choice process but rather accumulates from decisions made in multiple policy forums.

From a methodological standpoint, one of the defining features of polycentric systems is interdependence – many components of the system are connected to each other in different ways. Decisions made in one

policy forum affect payoffs in other forums (Lubell 2013). Actors' strategies and experiences in one policy forum influence how they behave and their relationships with other actors throughout the system (Mewhirter and McLaughlin 2021). Decisions made about one issue or location may influence the outputs and outcomes for other issues or locations (Bodin, Alexander, et al. 2019). Researchers examining the components that comprise such systems must look beyond the multivariate linear regression models commonly used in policy research that assume that observations are independent and, instead, utilize empirical methods that take such interdependence into account.

This chapter is meant to help prospective students and researchers of the EGF plan their own EGF study. In the next section, we outline the key concepts underpinning the EGF. Then, we delve into the practicalities of research design, including common challenges and approaches to overcome them. In the fourth section, we outline techniques of data collection that are commonly used in studies adopting an EGF perspective. These include both qualitative and quantitative data. The fifth and longest section delves into methodological aspects of measurement and analysis: these range from analysis of interviews to multivariate regression to multilevel network models. Finally, we conclude by outlining emerging and promising avenues of empirical analysis and theoretical development of the EGF. These include causal inference, text-as-data, and agent-based models.

Key Concepts in the EGF

The EGF is based on six key concepts: policy issues, policy actors, policy forums, policy games, policy systems, and time. Policy actors are usually representatives of public or private organizations that are impacted by the decisions made in a specific system. Policy issues are social, economic, or environmental processes that are of interest to a certain number of policy actors and therefore shape the policy preferences of those actors. Policy issues often result from collective-action problems or distributional conflicts. Policy forums are the decision-making processes in which involved actors deliberate and make collective choices about the policy issues. Because collective decision-making is governed by formal and informal rules, policy forums may also be called policy institutions. Policy games are constituted by the interactions between policy actors, forums, and issues, along with the institutional rules governing decision-making. The set of policy games can be considered a polycentric system and features structural attributes that can be measured in research design. Each of these components also has individual-level attributes, such as the social values of the actor, the type of issue, or the geographic scope of the forum.

The interactions among policy issues, actors, and forums play out over time in the context of a policy system. Policy systems are "geographically defined territories" that encompass multiple issues, multiple forums, and multiple actors interacting over time (Lubell 2013: 542). The interactions

involve three key processes (also called "functions" in EGF parlance): (1) learning about causal drivers of policy issues and the features of the actors and policy forums; (2) cooperation in developing and implementing policy and using resources; and (3) bargaining over the distribution of the costs/ benefits of policies. These processes of system functions are often measured as dependent variables in EGF research. However, the dependent variable necessarily varies according to the research questions or hypotheses. For example, many analyses focus on two-mode networks linking actors to forums or the one-mode networks of actor-to-actor collaboration. In these studies, the network structure itself is the dependent variable.

EGF research to date has focused on several key hypotheses. The "risk hypothesis" (Berardo and Scholz 2010) holds that policy networks provide access to different types of social capital depending on the level of risk actors face in a governance system. Bonding social capital is useful for addressing cooperation problems: those collective-action problems where the risk of collaborators' defection is high. Bridging social capital is useful for addressing learning and coordination problems: collective-action problems where the risk of collaborators' defection is low and instead common knowledge is needed to orchestrate joint decisions (Berardo and Scholz 2010; Burt 2005).

The "institutional externalities" hypothesis posits that social and biophysical linkages among forums create externalities where interactions in one forum impose costs or benefits on interactions in other forums (Klasic and Lubell 2020). McLaughlin, Mewhirter, and Lubell (2021) demonstrate that actors exposed to high levels of conflict in one forum can carry negative experiences with them to other forums in which they participate, potentially disrupting forum processes. Mewhirter, McLaughlin, and Fischer (2019) find that greater diversity among forum participants increases the likelihood that actors identify how forum-specific issues are linked to other issues in the system, allowing them to develop complementary policy outputs.

The "transaction cost" hypothesis draws from neoinstitutional economics in arguing that policy forums will produce more cooperation if they reduce the transaction costs of searching for policy agreements, bargaining over the distribution of costs/benefits, and monitoring and enforcing the resulting policy agreements. Hamilton and Lubell (2018) provide empirical evidence that the transaction costs of cooperation increase at higher levels of geographic and institutional (i.e., operational versus collective choice) scales. Lubell et al. (2017) suggest that forums are more likely to survive over time if they reduce the transaction costs of cooperation and therefore attract enough political support from actors who enjoy the benefits of participating in that forum.

The "multi-functional hypothesis" is that polycentric systems must support learning, bargaining, and cooperation over time. Hence, different components of polycentric systems and network structures may specialize on developing bridging social capital for learning or bonding social capital for coordination (Levy, Lubell, and McRoberts 2018; R. R. J. McAllister et al. 2017). Like the human brain, effective polycentric systems are likely

to be functionally differentiated (R. McAllister et al. 2020). The relative importance of different processes may change over time, for example, in planning for a disaster or security threat, versus responding and recovering (Bodin, Nohrstedt, et al. 2019; Nohrstedt and Bodin 2014). There may also be tradeoffs between different functions, which requires balancing the costs and benefits of learning, cooperation, and bargaining across the system (Hamilton and Lubell 2019).

Key Considerations in Research Design

This section describes some of the basic research design considerations for studying polycentric systems from the EGF perspective.

Selecting the System

The EGF defines a policy system as a geographically defined territory encompassing multiple policy forums, actors, and issues connected via biophysical or other types of processes (Lubell 2013). Ostrom (2009) might focus on a "social-ecological system," and while the EGF has been most frequently applied to environmental policy, it is applicable to any policy sector in which collective-action problems are the source of policy conflict. For instance, Mclaughlin et al. (2020) apply the EGF to study the implementation of collaborative programs designed to address domestic violence.

Several different criteria can be used to select study systems. Given our argument that all systems are polycentric, the easiest way to get started with the EGF is to choose a close-by system that is logistically cheap to study and provides opportunities for sustained engagement and relationship building. Larger systems tend to have more complex institutional arrangements, but even very small systems (e.g., your academic department and its various committees!) are polycentric. Systems that are selected based on convenience will still be undergoing theoretically interesting changes such as exogenous shocks from natural disasters, political changes, the emergence of new institutions, or the discovery of a new collective-action problem, among others. Comparative analysis of two or more systems provides the possibility to consider how contextual variables might influence the evolution of polycentric systems.

However, it is preferable to choose systems based on some criteria that might help illuminate a particular theoretical process or hypothesis. For example, one might choose systems all facing the same exogenous threat such as a climate change impact, which represents a relatively new collective-action problem. Holding the collective-action problem constant reduces degrees of freedom and allows the researcher to examine other features of the EGF that might be important. Another interesting theoretical contrast is overall political culture and level of macro-political institutional capacity across regions or countries. Political systems with stronger institutions provide a more fertile environment to study the evolution of more stable

polycentric systems, while weaker institutionalized systems experience more transient and ephemeral changes in polycentric processes (Berardo, Olivier, and Lavers 2015; Olivier and Berardo 2021). The structure and evolution of a polycentric system may also vary depending on how the political culture, and associated political institutions, responds to top-down or centralized policies versus more bottom-up, decentralized approaches.

Identifying System Components and Boundaries

What are the geographic boundaries of the system under study, and how does one identify the actors, forums, and issues involved in the system? These questions are challenging because systems change over time, with the emergence of new actors, forums, and issues and potential expansion of geographic scope via spatial processes. In most cases, there is not a perfect solution to this problem – by the very nature of interconnected processes, one could imagine expanding any study to encompass the entire globe! Hence, boundary setting must be justified from the theoretical perspective of what makes the study useful for empirical testing and recognizable to involved stakeholders.

   Moreover, once the system is identified, the analyst must identify its issues, actors, and forums. The attributes of these components and the interactions among them are typically the subjects of empirical measurement and analysis. Triangulating qualitative and quantitative methods is helpful for identifying components: participant observation, informant interviews, internet searches, and surveys are all useful. For example, Vantaggiato and Lubell (2021) identified key actors via informant interviews and snowball sampling. These actors were then surveyed and asked about their collaborative relationships. Lubell and Robbins (2021) adopted the reverse approach: they conducted a thorough internet search aimed at identifying collaborative forums and then web-scraped these forums to identify involved actors and the issues considered by each forum.

Longitudinal Analysis

Much of the extant EGF literature is based on cross-sectional analysis that relies on quantitative or qualitative analysis of policy forums or different configurations of policy networks. The EGF literature is equally concerned with system evolution over time as forums are created, change, and disappear; as actors shift participation and strategies; and as new issues are discovered and addressed. Longitudinal analysis would shed light on many open questions on these processes of change: How do polycentric systems evolve given different contexts, actor compositions, and perturbations? What mechanisms drive change in the system? How do actors respond to change with different patterns of conflict and cooperation? How does the changing social system co-evolve with the environmental or other policy outcomes?

The main challenge to answering these compelling questions is the difficulty of collecting data over time. Even a single cross-sectional analysis requires substantial research effort. Researchers rarely have the resources to follow the evolution of the governance system through time in a sustained way. Nevertheless, there are contributions in the EGF tradition that involve longitudinal analyses. Research designs that relied on longitudinal data collection typically comprise survey designs replicated at least twice (Angst and Hirschi 2017; Berardo 2014b; Berardo and Scholz 2010), whose results are analyzed using temporal network models. The results of these studies show that networks appear to evolve from coordination to cooperation, i.e., collaboration deepens over time as actors learn about each other's priorities and preferences and form coalitions. However, these studies are scarce and typically cover a limited time span.

A data collection approach that allows for longitudinal data analyses at lower cost is collection of text-based empirical data linked to governance processes, e.g., meeting minutes, policy reports, and the like. For example, Fisher and Leifeld (2019) use the text of political speeches and intervention to study the change in discourse surrounding climate change in the US Congress using a longitudinal approach. By finding patterns in large collections of text, quantitative text analysis enables the researchers to observe the evolution of relational patterns between actors as well as between them and their discourse. There is also ample scope for more study of system evolution using agent-based models, which simulate the evolution of a system of autonomous agents interacting with each other based on certain pre-specified rules. We outline these methods in more detail in a later section on methodological approaches.

Data Collection and Measurement

The EGF adheres to a mixed-method philosophy for research design, which utilizes the complementary strengths of qualitative and quantitative methods.

Collecting Data for Qualitative Analysis

Qualitative research facilitates a deeper understanding of the social and political processes driving actor decision-making, the operation of forums, and the feedback between the governance system and the policy issues. Qualitative research can directly measure key aspects of the EGF and inform the design and interpretation of quantitative methods. The qualitative research conducted by EGF researchers is directly aligned with the epistemological principles articulated by the analytic narratives approach to institutional analysis, which seeks to trace the causal processes driving institutional change (Bates et al. 2000). Quantitative methods cannot fully capture the rich strategies that actors employ to select forums in which to participate, interact within forums, and seek to change forums. The dynamics of collective decision-making in forums is sometimes affected by variables that

are extremely difficult or impossible to observe with quantitative approaches, such as hallway conversations or personality conflicts. Qualitative research is also crucial for science communication and policy engagement because it provides a pathway for developing relationships with policy stakeholders, designing research that resonates with their mental models and policy parlance, and increasing the relevance of research through co-production. However, the complexity of polycentric systems is a challenge to qualitative research – it is difficult to keep track of all the system components, especially if one is trying to compare multiple systems or keep track of changes over time.

Participant Observation

Observing or participating in policy forums allows fine-grained observation of how decisions unfold over time, including the salience of different ideas and issues and how actors move around the system. Because there are many forums in a polycentric system, the researcher needs to decide which to observe. One clear necessity is to observe the largest or most central policy forums, where the most important policy actors are interacting. Most policy network theory suggests that the policy forums and important actors at the core of the network have a large influence on spreading cooperation and resolving conflict in the overall system (Robins and Alexander 2004; Robins, Bates, and Pattison 2011). Participant observation will usually uncover a network of "usual suspects" that show up across multiple forums and have repeated interactions that form the basis for cooperation and leadership. However, the researcher should not forget about the many smaller and more peripheral forums. These peripheral forums may be where excluded or dissatisfied actors will appear or where innovative ideas are developed that are outside the "normal" set accepted in the core of the system. Key informants may even nominate some peripheral forums as places where important events are happening. Hence, it is useful to at least sample a few of these more peripheral forums, selected based on how they might inform a particular theoretical question.

Key Informant Interviews

Key informant interviews are an important complement to participant observation. As mentioned earlier, polycentric systems usually have a core network of actors who participate across multiple forums and know each other fairly well. Many of these actors serve in leadership roles in a policy forum (e.g., chairpersons) or within their organizations (e.g., executive directors, program leaders). The researcher should interview as many of these people as possible using semi-structured interviews with a minimal number of questions (10 is a good target). The semi-structured interviews should facilitate theory-guided discussion rather than meticulously documenting answers to dozens of questions and sub-questions. The researcher needs to

understand how the respondents see the world and the language they use to describe their environment and decisions. This follows a more inductive or grounded theory approach that assumes the respondent is expressing the complex reality that the theory is supposed to describe and analyze.

Document and Archival Analysis

Polycentric systems leave archival traces especially in online documents, which can become important resources for understanding the history of decision-making and institutional change. Individual forums often produce vision documents, strategic plans, or annual reports that provide a good summary of the history and mission of the forum, the key issues under its jurisdiction, and the actors involved. If available, meeting minutes provide an opportunity for process tracing around key decisions and observing the behaviors and preferences of actors (Ulibarri 2015).

It is often not feasible to read the archival documents associated with all policy forums and actors, so it is important to focus first on the "big games" and "big actors" and their associated foundational policy documents. These can be analyzed via qualitative content analysis aimed at mapping actors to outcomes or to discursive frames that they use to further their agendas (Berardo, Heikkila, and Gerlak 2014). That said, there may be times the researcher is interested in comparing across all forums, for example, to analyze the extent to which climate change is integrated into planning processes. Archival documents may become good sources for the application of "big data" methods such as natural language processing, machine learning, network analysis, or other "computational social science" approaches.

Studies that rely on documents rather than fieldwork can sidestep major investments in time and effort required for eliciting data from respondents. This consideration is especially important in EGF research, which typically adopts a systems perspective that requires significant data inputs (e.g., from as complete a set of actors participating in related policy processes as possible). Likewise, the importance of social processes (e.g., cooperation) in EGF research highlights the value of longitudinal records, for which the accuracy of text-based data offers the potential to avoid well-documented challenges of eliciting information about social interaction using recall methods (e.g., Brewer 2000). Additionally, text-based records can reveal multiple phases of policy processes, including implementation of policy outputs, thereby offering a more complete depiction of how processes of self-organization play out among interdependent decision-making forums (Nohrstedt 2018). Surveys and interviews can likewise reveal such information but may require multiple waves of data collection.

Collecting Data for Quantitative Analysis

While qualitative research can produce valuable inputs to discern how polycentric systems are structured and evolve in specific cases, quantitative

analyses remain critical to produce results based on large-N empirical observations that can improve the external validity of findings. Furthermore, quantitative methods are often easier to replicate across cases and over time, in order to facilitate comparative analysis using common metrics.

Surveys

The bulk of EGF quantitative work to date has relied on online or telephone surveys of policy actors, where the sampling frame is assembled from a combination of internet searches, collating contact lists of participating actors from online documents, contact lists provided by policy actors or forum managers, and key informant nominations. While surveys typically target individual people, the main interest is usually in how their organizations participate in polycentric systems. Surveys typically collect four crucial types of information: actors' individual attributes, policy networks, forum participation, and issue attributes. Depending on the research question and analytical methods, the measured attributes of any of these elements of the EGF may serve as independent or dependent variables.

Actors' individual attributes are usually measured by asking survey respondents about the positions their organizations have on different topics; for instance, how organizations react to risky environments in which defection is possible (or even likely), their experience with various policy issues, their policy core values, and other theoretically interesting variables. These individual attributes are often considered important in shaping overall policy preferences and actors' capacity and opportunities to develop policy networks and participate in policy forums.

Second, surveys often collect basic data to measure the structure of policy networks, usually based on collaborative or information-sharing relationships. Angst and Hirschi (2017) collect longitudinal data by administering two different surveys in which they ask actors to indicate their collaboration relationships from a roster of all policy organizations involved with the establishment of a park in Switzerland. In the context of water governance in the lower valley of the Chubut River in Argentinian Patagonia, Olivier and Berardo (2021) use a "name-generator" question that asks respondents to "name up to five organizations with whom you/your organization collaborate/s periodically to address water management problems in the lower valley." Another example is in the work of Hamilton and Lubell (2018), who study climate change adaptation around Lake Victoria in East Africa and use an in-person survey that includes the following name-generator question: "Please list the organizations your organization has collaborated with in the context of climate change adaptation in the past year."

The third type of information collected through surveys pertains to the forums or venues in which policy stakeholders participate to advance their agendas. For instance, Lubell, Henry, and McCoy (2010) conducted a mixed survey (with both an online and a telephone component) that asked actors to indicate their participation in a collaborative land-use and transportation

forum, along with more traditional local planning processes. Fischer and Maag (2019) use a survey to examine the participation of over 150 actors in eight cross-sectoral policy forums in Switzerland that affect habitat and land-use governance. Berardo et al. (2015) studied water governance in South America's Parana River delta with three waves of surveys that included the following question:

> The issues of water and land use in the delta can be discussed in different forums, such as regional planning councils, advisory boards, workshops, etc. Could you mention the names of the forums in the delta in which you/your organization have/has participated in the last year?

The same survey was administered in the Tampa Bay region and northern California, producing results to assess the functioning of governance systems across different levels of institutional strength (Berardo and Lubell 2016; Mewhirter, Lubell, and Berardo 2018). Most of the time, once patterns of participation in a forum are observed, surveys often ask respondents their perceptions about various aspects of forum performance or their experiences while participating in the governance system. For example, Mancilla García and Bodin (2019) study the management of the Paraíba do Sul river in Brazil and use a survey to learn what are the forums in which actors participate; they find out that increased participation leads to the accrual of greater levels of influence throughout the governance system.

The final type of valuable information that surveys produce is how actors work on specific issues or problems that drive decision-making processes. Olivier and Berardo (2021) ask survey respondents to identify seven issues on which actors are interested, including water quality, water scarcity, and securing sufficient water flows for recreation activities such as fishing. Bodin and Nohrstedt (2016) also explored the topic of issue interdependency and individual-level action, showing that interdependency among fire management tasks in Sweden was correlated with collaborative relationships between them. Berardo and Lubell (2016) examined three ecologies of games at different stages of "institutional development" by first identifying water-related issues and stakeholders through an online media search and then used a combination of phone and online surveys to gauge both how much of the actors' work concentrated on specific issues and how they participate in relevant forums where the topics of water quality and land-use drive policy decisions. Angst's (2020) study of the Swiss water governance system administered an online survey to identify how 467 organizations connected to 26 issues related to water governance. The surveys thus can collect actor perceptions of various issues and build networks based on those actors who jointly work on issues.

Natural Language Processing: Text-as-Data

Natural language processing (NLP) methods provide automated, computational approaches for analyzing large amounts of text, which presents

significant opportunities to measure the social and behavioral processes involved with polycentric governance at a scale for which manual coding would not be practical. NLP methods can roughly be divided into approaches that emphasize information extraction (e.g., named entity recognition, geo-tagging, network relationships) versus content analysis (e.g., topic modeling, word embeddings, sentiment analysis). NLP is a rapidly changing field of "big data" science, with many tools available in R and Python, as well as stand-alone software such as Automap and Discourse Network Analysis. The specific NLP methods used in an EGF study will hinge not only on the research questions of interest but also the characteristics of the text that serves as the key source of data for analysis (e.g., the level of detailed discussion of decision-making processes may vary significantly between meeting minutes, transcriptions of speeches)

For example, Fisher and Leifeld (2019) analyze the content of US Congressional testimony on climate change to measure a polycentric discourse network (i.e., with linkages between policy actors and policy beliefs). Murphy and colleagues (2014) apply NLP and named entity recognition approaches to a large dataset of newspaper articles to map relationships among actors and characterize decentralized water governance networks in the Southwestern United States. Bell and Scott (2020) use named entity recognition and topic modeling to analyze the "paper trail" of meeting notes, scoping documents, and other materials, to demonstrate significant variation in the outputs of a set of regional water planning processes despite their common design.

Ulibarri and Scott (2017) combine qualitative and quantitative approaches in their analysis of three hydropower relicensing processes held by the Federal Energy Regulatory Commission (FERC) of the USA. They extract policy networks from meeting minutes and find that collaborative processes featured high reciprocity (indicating bonding social capital) in comparison to asymmetrical triadic relationships that indicate power differentials in non-collaborative processes. In their study of the process surrounding a case of dam relicensing by FERC, Scott, Ulibarri, and Scott (2020) relied on NLP to perform name entity extraction to identify actors and their affiliations, and part of speech tagging to identify the actions that actors performed within the collaborative governance forums. They then modeled actor participation into forums using hierarchical Bayesian modeling.

Another important application of NLP methods is translating the Institutional Grammar Tool into the context of polycentric governance (Crawford and Ostrom 1995; Olivier 2019; Siddiki et al. 2019). The Institutional Grammar Tool is a promising approach to systematically measuring the structure of rules across policy forums. Weible et al. (2020) document how the four main policy forums related to energy development in Colorado vary in terms of their rate of rule creation, and the actors, issues, and actions targeted by the rules. Along similar lines, Heikkila et al. (2021) classify 22 California energy policies from four decision-making venues in terms of the types of actors they share and the content of the rules (e.g., enforcement, information). The studies clearly demonstrate the multifunctional nature of polycentric systems because different forums are producing

rules related to different types of collective-action problems related to energy development.

Data Analysis Methods

In this section, we outline the main qualitative and quantitative methods that are currently prominent in EGF research. These comprise case study analysis and interview coding and analysis, as well as multivariate regression methods and quantitative network analysis. Agent-based models have also been usefully applied to EGF ideas, which allow the researcher how key processes evolve over time under different parameter configurations.

Qualitative Methods

Case Study Analysis

The EGF literature relies on case study analysis to portray the characteristics of the specific governance system and the perceptions of the actors involved in the governance system (Lubell et al. 2017). Institutions and norms are important in shaping actors' incentives, priorities, and participatory strategies. Perceptions are important because they shape actors' behavior in the system: the forums they attend, the actors they establish ties with, the allocation of resources to collaborative activities, and other behaviors. Overall, qualitative data provide insight into the key variables that are important to understand a particular system and the causal process driving actor behavior, institutional change, and policy outputs/outcomes.

In EGF research, qualitative data are typically an essential input into quantitative analysis, both by helping design effective quantitative measurement instruments like surveys and through interpreting the extent to which quantitative results reflect key causal processes occurring in polycentric systems. Although thus far EGF case studies that are purely qualitative without any quantitative methods are rare, there is scope for more qualitative work in the EGF literature. For example, detailed comparative studies of collaborative governance processes occurring within the same EGF but leading to different outcomes could shed light on the determinants of forum success and formulate new hypotheses to be tested in future research.

Coding and Analysis of Interviews

Interviews can provide a wealth of information not only on the perceptions of governance actors but also on the effective size of the system, the cleavages existing between coalitions, and the key agenda items of contention, among other variables of interest. Moreover, interviews prove important seeds for snowball sampling of key informants on the governance setting.

Lubell (2017, report) relies on analysis of 40 interviews with informants from different affiliations and levels of governance to identify the key

governance challenges that actors face in adaptation to sea level rise in the San Francisco Bay Area. On the basis of this analysis, Lubell (2017) derived the key themes that informed the quantitative survey for the same research project, which was fielded to a much wider set of stakeholders (Lubell, Vantaggiato, and Bostic 2019). Subsequently, the authors use the Discourse Network Analyzer software (Leifeld 2010) to code those same 40 interviews according to an inductive process that distinguished actors' frames, perceptions of governance challenges and preferred or hypothesized solutions to the challenges. The authors then transformed this coding scheme into a network structure connecting actors to their frames, challenges, and solutions to identify the key combinations that exist in the governance system of sea level rise in the San Francisco Bay Area.

Quantitative Methods

Generalized Linear Models

Generalized linear models (or "GLMs": e.g., ordinary least squares [OLS] regression, logistic regression, negative binomial regression) provide powerful tools to examine how a specific outcome measure is associated with a set of predictor variables (Stock and Watson 2015). For EGF analyses using GLMs, the unit of analysis is often one of the key system elements (i.e., actor, forum, or issue) on which the independent and dependent variables are measured.

GLMs assume that the error term is uncorrelated across observations: an assumption that is typically violated in a polycentric system where individual observations are inherently nested within distinct "levels" or "clusters" (Berardo and Lubell 2019). For instance, EGF research examining the determinants of actor evaluations of forum effectiveness have traditionally relied on survey data where respondents provide ratings for each forum in which they participate (e.g., Lubell et al. 2017; Lubell, Mewhirter, and Berardo 2020). This data generating process gives rise to multiple forms of interdependence, whereby actors who participate across multiple forums are represented in multiple observations; observations are nested within forums when multiple actors participate in the same forum; actors in the dataset often represent the same organization; actors work on the same distinct issues and/or in the same distinct region. Estimating a model that fails to take such interdependences into account can lead to artificially small standard errors, thus increasing the likelihood of committing a Type I error – i.e. finding "false positives" (Cameron and Miller 2015; McNeish and Stapleton 2016; Moulton 1990).

Addressing Interdependence Through Clustered Standard Errors

Clustered standard errors account for interdependence with a two-step approach where common regression techniques (OLS, logistic regression, etc.) are used to calculate the point estimates for slope coefficients, and then

standard errors are ex post adjusted to account for nonindependence (Primo, Jacobsmeier, and Milyo 2007). The process is a variant of White's (1980) "sandwich estimator" and offers the additional benefit of accounting for heteroscedasticity (Stock and Watson 2008). Clustered standard errors treat interdependence as a nuisance to be corrected: it allows the researcher to observe how a given variable impacts an outcome of interest and adjust standard errors for within-cluster dependence (McNeish and Stapleton 2016). It does not, however, allow the researcher to observe the extent to which variation in the dependent variable is attributed to cluster-based factors nor whether and to what extent the slope of a given variable varies across clusters.

Several EGF studies utilize multiple regression models that specify single-way and/or multi-way clustering (Cameron and Miller 2015; Cameron, Gelbach, and Miller 2012). Lubell et al. (2017) examine how actor-specific transaction costs impact actors' perceptions of forum effectiveness across the multitude of forums in which they participate. To do so, the authors estimate a series of OLS regressions with errors clustered at the actor level to adjust for interdependence caused by actors appearing multiple times in the dataset (contingent on the number of forums in which they participate). Similarly, Hamilton and Lubell (2019) examine how patterns of interactions among policy actors affect assessments of forums in the Lake Victoria region in East Africa. The authors utilize an ordered logistic regression model with multi-way clustered standard errors at both the actor level (addressing the same issue as Lubell et al. 2017) and forum level (as numerous actors participate in the same forum).

Addressing Interdependence Through Multilevel Modeling

Multilevel models allow a researcher to directly model clustering in the data through random coefficients (Laird and Ware 1982; McNeish and Stapleton 2016). Multilevel models are composed of "fixed" and "random" components. The fixed part of the model represents the relationship between predictor variables and the outcome of interest regardless of the cluster (forum, organization, etc.) to which an observation belongs. The interpretation of the fixed part of the model – the intercept and slope coefficients – is interpreted in a fashion similar to a standard regression model. Random effects can be estimated for the intercept and the slope coefficients, though both are not necessary. Random intercepts allow for the intercept to vary across clusters in the data, thus directly adjusting for cluster-specific autocorrelation. Random coefficients allow for the effect of a predictor variable to vary across clusters (Hedeker, Gibbons, and Flay 1994; McNeish and Stapleton 2016).

Relative to regression with clustered standard errors, multilevel models are advantageous in that they allow the researcher to directly assess the extent to which variation in the dependent variable is attributable to cluster-specific factors (through random intercepts) as well as the extent to which slopes vary across clusters (through random coefficients). Post estimation, the researcher

can examine how the intercept and/or slope coefficients vary across clusters in the model, which can be of theoretic and/or practical relevance (Rabe-Hesketh and Skrondal 2008). While multilevel models allow the researcher to specify a number of random intercepts (forum, organization, actor, etc.), they require the researcher to identify and subsequently model the nesting structure of the data: i.e., establish whether the two or more clusters are hierarchically ordered or cross-nested (McNeish and Stapleton 2016).

To date, multilevel models have primarily been used to examine factors that impact within-forum influence. Mewhirter and Berardo (2019) use survey data to examine how the geographic diversity of forums in which one participates impacts within-forum influence and how such a relationship is moderated by the structure of one's ego-network. The authors estimate a cross-nested multilevel regression with random intercepts at the actor level (accounting for multiple observations per actor) and forum level (accounting for multiple observations per forum), finding that a significant amount of variation is caused by both forms of clustering. Mewhirter, Coleman, and Berardo (2019) use a multilevel regression to examine the factors that promote actor influence in the forum that most readily impacts their interest (their "primary" forum) relative to other forums in which they participate. The authors utilize a multilevel model specifying random intercepts at the actor level and forum level, and a random slope coefficient for "prime": a dummy variable indicating that a forum serves as one's prime forum. The results indicate that significant variation in influence is attributable to actor-level and forum-level clustering and that the effect of "prime" varies across actors in the dataset.

Network Analysis

A core assumption of the EGF is that decision-making processes are interdependent. Attention to interactions – e.g., among actors, between actors and forums – highlights the value of the tools and perspectives of network science. The multidisciplinary field of network science is unified by attention to the complex relational structure among entities (Robins 2015). For example, a social network is composed of a set of social actors and the relationships among them. A network can be represented as a graph comprising nodes and edges. The nodes represent the social actors, and the edges represent the relational ties existing between them. Network analysis offers a toolkit of methodologies for characterizing polycentric governance systems and for evaluating hypotheses about the processes that generate patterns of interactions as well as the outcomes of these patterns.

Governance Networks

Here, we use the term "governance networks" to refer to a network composed of different types of nodes (actors, issues, or forums) and edges that indicate some type of relationship between the nodes. As discussed earlier,

the set of relevant nodes depends on key research design considerations of selecting the system and the boundaries, which then delineates the policy actors, forums, and issues that can potentially serve as nodes in the governance networks. Likewise, analysts must decide how to measure the edges that denote relationships or interactions among nodes. For example, EGF research commonly focuses on patterns of information exchange, cooperation, and/or collaboration among policy actors and on patterns of actor participation in forums.

Once the potential nodes and relationships are established, the analysis may focus on networks of one, two, or even three different types (i.e., modes) of nodes. One-mode networks of actor-to-actor relationships (e.g., information exchange, cooperation, and collaboration) in which the nodes are all of the same type and relationships can exist between any of them are the most common type of governance network studied in the broad policy network literature (Berardo and Scholz 2010) and have been featured in EGF research as well (Angst and Hirschi 2017; Hamilton, Lubell, and Namaganda 2018).

In contrast, actor to forum ties form so-called 'two-mode' networks, meaning that they focus on patterns of interaction between two types of nodes (e.g., actors and forums). Bipartite networks are a special type of two-mode networks in which there cannot be ties between nodes of the same type but only between nodes of different types, i.e., between actors and forums. Actor–forum networks are the most commonly studied in EGF research (Berardo and Lubell 2016; Lubell, Robins, and Wang 2014; Scott and Thomas 2017), because the existence of multiple policy forums is one of most fundamental features of polycentric systems. Actor–issue networks, which one could call an "issue network" (Heclo 1978), are less studied in the EGF literature (but see Brandenberger et al. 2020; Hedlund, Bodin, and Nohrstedt 2020).

Recent technical advances have enabled the study of three-mode networks that link issues, policy actors, and forums simultaneously. Policy actors may work to address certain issues and may also participate in forums. Forums may have jurisdiction over particular issues. Three-mode networks are an important research frontier because they embody all the fundamental concepts of the EGF in a single relational structure.

Multilevel Networks

Some EGF studies analyze multilevel networks, meaning they not only measure interactions between types of nodes (as in bipartite networks) but also among the same types of nodes (e.g., accounting for how actors participate in forums as well as how actors collaborate with one another). Analyses of multilevel networks characterized by ties among policy actors (e.g., collaboration) and between actors and policy forums (e.g., participation) indicate that joint participation in forums enables collaboration in settings external to forums (Hamilton and Lubell 2018; Fischer and Sciarini 2016), possibly because joint participation reduces transaction costs. In their analysis

of actor–actor and actor–forum ties, R. McAllister et al. (2020) showed that Australian marine biosecurity governance networks were predisposed to exhibit bridging social capital, as evidenced by an overabundance of "open" network configurations (e.g., instances in which actors participate in forums in which their collaborators are absent).

Conceptually, accounting for multiple types of nodes provides opportunities to measure polycentric governance structures and processes in ways that are not possible using one-mode governance networks. Specifically, measuring actors' participation in policy forums enables researchers to evaluate the interplay between the structure of polycentric networks and social processes (e.g., direct collaborative relationships between actors) and/or policy processes (e.g., direct payoff externalities or strategy externalities between forums).

A special type of multilevel network measures environmental connectivity along with social and political interactions. Such social–ecological networks can capture the full scope of EGF dynamics by accounting for patterns by which policy actors and forums engage with distinct, but related, policy issues. Guerrero et al. (2015) use social and ecological data from a large-scale biodiversity conservation initiative in Australia to investigate whether observed patterns of stakeholder interactions mirror network configurations hypothesized to enable them to address management challenges. In an important comparative study, Widmer et al. (2019) analyze three European watersheds for prevalence of social–ecological motifs indicating both institutional fit and misfit. While both studies find motifs that indicate social–ecological fit, they also find significant evidence of misfit, which indicates significant opportunities for improving management.

Exponential Random Graph Models (ERGMs)

While network analysis encompasses numerous approaches for deriving insight from relational data, the dominant network analytic approach in EGF research is the use of ERGMs. ERGMs are a family of statistical models for social networks that permit inference about prominent configurations in the network structure, given the presence of other network structures (Robins 2015). Namely, ERGMs identify parameters by maximizing the probability of the observed network over the networks with the same number of nodes that could have been observed. This is conditional on a set of network statistics that can include node characteristics (e.g., organizational affiliation), edge characteristics (e.g., relationships between actors of the same organizational affiliation), as well as certain network structural characteristics or "motifs" (i.e., patterns of relationships among a small number of nodes; Desmarais and Cranmer 2012). The underlying assumption of ERGMs is that the observed network structure has emerged from an evolutionary process of tie formation over time, which can be explained by the combination of theoretically and empirically relevant variables as well as network dependency structures (Robins, Lewis, and Wang 2012). Hence even

a single cross–sectional ERGM represents a snapshot of the accumulation of a dynamic process over time, and theory is used to guide expectations about what types of network structures may have evolved and how they might influence individual behavior (Lubell et al. 2012).

ERGMs are appealing because they provide the means to test expectations about how social and policy processes contribute to patterns of relationships in governance networks. ERGMs can test hypotheses about how individual attributes of actors (e.g., social values, experience), forums (e.g., mandatory or collaborative, level of geographic scale), or issues (e.g., severity, uncertainty) may influence how individual nodes are linked to the network. ERGMs also test structural hypotheses by exploring the relative prominence of different network motifs, whose over/underabundance in a governance network may reflect the types of microlevel processes operating in the system. In other words, ERGMs assume that the global structure of a network of relationships is generated by microlevel processes, exemplified by motifs, which are given theoretical significance (Desmarais and Cranmer 2012).

The concepts of bridging and bonding social capital, which are central to the risk hypothesis described earlier, are an illustrative example of how different motifs may represent social processes. Figure 9.1 presents the different types of motifs that may represent bridging and bonding social capital across different types of networks. The motifs that measure bonding social capital share a structural characteristic known as "closure"
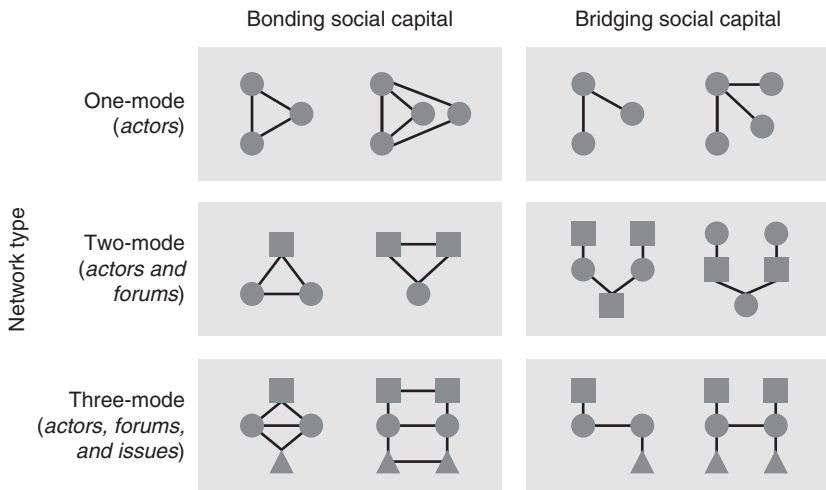
Figure 9.1 Examples of bonding and bridging social capital motifs for different kinds of networks

Note: Shapes indicate whether nodes represent actors (circles), forums (squares), or issues (triangles). Lines indicate linkages among actors, forums, and/or issues. Although network studies commonly distinguish between bridging and bonding social capital based on node attributes, for simplicity, we present motifs characterized based on structure alone.

(Burt 2005), which refers to linkages that connect nodes in ways that create clusters. For example, both motifs in the one-mode bonding social capital cell of Figure 9.1 depict triadic closure, which can enable nodes (e.g., actors) to monitor each other's behavior and otherwise encourage cooperation (Berardo and Scholz 2010).

By contrast, motifs that measure bridging social capital exhibit features that enable exposure to diverse resources. For example, the picture on the left-hand side of the one-mode bridging social capital cell of Table 9.1 is called an "open two-star"; it shows a central node with relationships to two nodes that do not interact. Thus, the central node will possess more information than either of the other two nodes. This makes the central node a desirable network partner. Possession of many ties is typically considered a measure of "popularity." The assumption is that actors prefer to connect to popular nodes to use them as sources of information about other actors, to gather information efficiently. Thus, a new node in the network will rather form a tie with the central node than to any other node because that tie allows them to access more information at once. Moreover, it is likely that any information or other resources transmitted to the central node via one of those relationships will be distinct from information or other resources accessed through any of the other relationships.

EGF studies extend this logic of using "closed" motifs to measure bonding social capital and "open" motifs to measure bridging social capital to more complex networks that account for more than one type of node. Examples of such motifs are also provided in Figure 9.1. For instance, the left-hand motif in the two-mode bonding social capital cell represents an instance in which two policy actors interact directly and jointly attend the same forum. Overlapping participation in forums helps sustain cooperation in the long run because actors who are involved in forums alongside their collaborators may perceive defection to be a particularly costly strategy to the extent that joint participation provides a mechanism for monitoring behavior (Berardo 2014a).

Bridging social capital in two-mode networks serves to foster innovation and exchange of information between actors across policy forums. This superior knowledge is capital that policy actors can leverage (Burt 2005). Typically, the intent of actors possessing bridging social capital is performing a brokerage function, thus gaining a central position in the network that may accrue them reputation, prestige, and legitimacy (Jasny and Lubell 2015). The right-hand motif in the two-mode bridging social capital cell depicts a policy actor gathering information from different, disconnected forums in which other participants are not collaborators of the actor. In such network structures, forums provide a bridging function between sets of actors who are otherwise disconnected, thereby enabling the transmission of new forms of information among policy actors.

More complex configurations arise in three-mode networks involving actors, issues, and forums and in social–ecological networks involving actors, forums, and ecological units. The logic of bonding and bridging social capital,

however, remains the same: actors who seek and possess bonding social capital invest their resources into collaboration aimed at monitoring their partners' behavior and fostering common positions (e.g., in the three-mode bonding social capital cell); actors who seek and possess bridging social capital leverage the superior knowledge they acquire by bridging across otherwise disconnected entities to gain positional and reputational advantages (e.g., motifs in the three-mode bridging social capital cell).

FURTHER DEVELOPMENTS OF ERGMS

Recent developments in statistical analysis of networks have extended ERGMs beyond models suitable for one-mode, cross-sectional network data. This enabled analysis of the increasingly complex configurations outlined in Table 9.1 via quantitative network models. Recently developed ERGMs cover bipartite (or two-mode) networks (Jasny 2012; Jasny and Lubell 2015), valued networks (i.e., networks where the ties have a weight, or value, indicating the strength or the importance of the tie), and multilevel networks (Wang et al. 2013; Wang, Robins, and Matous 2016). Moreover, Temporal ERGMs and Separable Temporal ERGMs allow for the pooling of network observations at different time points and, therefore, analysis of network evolution over time (Leifeld, Cranmer, and Desmarais 2018).

Studies in the EGF literature have made ample use of the flexibility of the ERGM family of models. For instance, Scott and Thomas (2017) use bipartite ERGMs to test the determinants of actors' choices regarding which collaborative forums to attend. Berardo and Lubell (2016) have extended the risk hypothesis to bipartite networks and found that joint actor participation in multiple forums suggests the existence of bonding social capital between them, as they meet across multiple venues and have opportunity to get acquainted with each other's preferences; in contrast, the tendency of single actors to attend multiple forums suggest that they possess bridging social capital, which enables them to leverage information across forums, whether to further their own goals (R. R. J. McAllister, McCrea, and Lubell 2014) or to improve coordination by remedying the fragmentation of the institutional system (Jasny and Lubell 2015).

Scott (2016) used valued ERGMs to model relationships of varying strength among a regional network of organizations involved in 57 collaborative groups and test the effect of government sponsoring of collaboration on collaborative outcomes of coordination and cooperation. Multilevel ERGMs are widely used in the social–ecological network literature, including contributions addressing debates in the EGF literature specifically (Guerrero et al. 2015) as well as in contributions modeling collaboration between actors who operate at different levels of governance (Nohrstedt and Bodin 2020).

Finally, temporal network models have been used to test the assumption of the risk hypothesis that policy networks evolve from being driven primarily by coordination – as actors use the network to gather information about the policy issue and about other actors' preferences – to being driven

primarily by collective-action problems of cooperation – as actors, having learned each other's preferences, form coalitions that can be more or less adversarial in nature (Angst and Hirschi 2017; Berardo and Scholz 2010).

Agent-Based and Computational Models

Agent-based models are a useful approach to studying complex systems because they allow the specification of all the components of the system and the rules governing how those components interact. Agent-based models allow observation over time and tracking of multiple parameters and state variables that often experience nonlinear effects and feedback. The researcher can easily conduct experiments with agent-based models by exogenously adjusting parameters and analyzing the results for the system at multiple scales (e.g., microlevel individual behavior, dynamic network structures, macro-level performance). Unlike real-world polycentric systems, it is possible to "rewrite history" by constructing multiple versions of models to compare performance and explore counterfactuals. Of course, all agent-based models are necessarily simplifications of the real system and there is some limit to model complexity before researchers no longer understand how the model works and consumers gain no insights from the results.

Paul Smaldino and Mark Lubell collaborated on a series of papers translating the EGF into an agent-based modeling framework. The basic idea of the model is to set up multiple public goods games in which players decide their level of resource contribution and can free ride on the benefits of cooperation. The players have fixed strategies about cooperation or defection, such as splitting their contribution budget among multiple games. The models feature a migration process where players can join and leave games based on the payoffs they experience each round or other parameters. The model can implement institutional rules to control the migration process. Smaldino and Lubell (2011) found that capacity constraints on how many actors can join a game facilitate more cooperation than player-level budget constraints on how many games they can join. Smaldino and Lubell (2014) found that reputation-based institutions that condition entry into a new game based on past cooperative behavior are even more effective at producing cooperation. The core principle driving cooperation in these EGF models is the clustering or positive assortment of cooperative strategies, whereby cooperative strategies are more likely to interact with each other.

The benefits of the simple implementation of this model come at the cost of losing important features of polycentric systems that matter in the real world. For example, the Smaldino and Lubell models all feature games with same payoffs; that is, they are all public goods games. There are many types of social dilemmas with different payoff structures that could be implemented including zero-sum games. The policy forums in real polycentric systems almost certainly feature heterogeneous payoffs (Lubell, Vantaggiato, and Bostic 2019). Similarly, the actors do not have dynamic strategies that are common in repeated game or evolutionary game theory and do not use

heuristics that apply effective strategies learned in one game to another game with different payoffs. The payoffs in the forums are not an explicit function of the types of issues considered in the forum, which vary in number and type in real polycentric systems. These are all core ideas in the EGF that could be implemented in future models.

Conclusions: Future Directions in Ecology of Games Research

The EGF seeks to assess how the attributes of policy forums, policy actors, and policy issues, and the structural relationships, among them, influence the key processes of learning, bargaining, and cooperation that are needed to address multiple, interrelated collective-action problems in polycentric governance systems. The EGF has spawned a burgeoning research community that uses a diverse set of qualitative and quantitative methods ranging from case studies to network analysis, to computational social science. The extant research, which mostly focuses on environmental governance, has started to produce valuable insights around some key EGF hypotheses as well as provide a range of other interesting results that test and expand the theory. However, the extant research has only started to scratch the surface; much more innovation in research design and analysis is needed.

From a research design perspective, we need many more comparative and longitudinal research projects. The existing comparative studies only examine a few regions, and longitudinal studies rarely feature more than two time periods. But the EGF assumes that polycentric systems experience evolutionary processes, which unfold over long periods of time and are heavily influenced by contextual features of social–ecological systems. The structure of the governance systems also coevolves with policy outputs and outcomes, which are almost never measured in current research but of central interest for theoretical and practical reasons. Many of the key processes and structures have long-term and possibly complex dynamics such as feedbacks, nonlinearities, punctuations, and thresholds that are very unlikely to be observed in short-term studies. Hence, what is really needed is a network of social–ecological research observatories that measure key elements of the EGF over long periods of time across many social–ecological systems. This requires a major investment in research infrastructure as well as cooperation among a large global research community.

The EGF also needs to further expand beyond environmental governance. Even though environmental governance has been a natural field of study given the roots of the EGF in research on polycentric governance, common-pool resources, and social–ecological systems, researchers have begun to apply the EGF to other policy sectors, even in contexts that might seem exotic to policy scientists, such as international soccer (Madison 2021) and Norwegian handball (Bjørndal and Gjesdal 2020; Bjørndal, Ronglan, and Andersen 2017). Simple introspection should be enough to realize that polycentric systems are a fundamental feature of all policy domains: health,

education, criminology, agriculture, among others. The various research design, measurement, and analysis methods summarized in this chapter can be easily exported to other policy domains.

The EGF could also benefit from advanced social science methods that are making important contributions to understanding causality in different fields: mathematical models, causal inference methods, and experiments. Mathematical models would seek to represent key EGF processes with sets of equations and could present closed-form solutions based on equilibrium or optimization concepts. Mathematical models are typically extremely simplified representations of the system, which is desirable from the perspective of parsimony but is challenging in the face of the complexity of polycentric systems. Since the very name of the EGF implies strategic interaction, formal game theory models would be a particularly attractive approach. Indeed, some of the agent-based models described earlier do embody a game theory approach by modeling a network of multiple public goods games.

Twenty-first-century social science has experienced a revolution in the analysis of causality and the development of causal inference methods (Winship and Morgan 1999). Causal inference methods approximate experimental methods by attempting to identify some type of "treatment effect" in comparison to a statistically constructed counterfactual. This is conceptually easy to understand when there is some type of endogenous "treatment" like participating in a policy program. But as examples of complex adaptive systems, polycentric systems are characterized by complicated causal relationships and feedback that unfold at multiple time scales, including the coevolution of network and institutional structure with the social processes of learning, bargaining, and cooperation. It is difficult to imagine a "treatment" in the context of the EGF, for example, isolating policy participation in one policy forum as a treatment while holding other policy forums constant at a particular point in time. In a particular system, such as the California Delta, one finds it difficult to imagine a counterfactual such as what would have happened without the large-scale collaborative program like CALFED (Lubell, Gerlak, and Heikkila 2013). Computational and mathematical models provide some basis for analyzing causality by tracing processes and equilibrium outcomes under different assumptions and parameter settings. ERGM and other network methods are also amenable to simulations that could explore how different types of network configurations might plausibly link to EGF hypotheses. However, more creative research design is needed to apply such modern causal analysis methods to empirical systems.

Experiments are another method for analyzing causality but are also challenged by the complexity of polycentric systems. There have been some useful small laboratory experiments where subjects participate in multiple games or made decisions in the context of exogenously imposed network structures or network structures that evolve endogenously in response to subject decisions. Many of these experiments at least implicitly consider EGF concepts, but more thought could be put into designing experiments to explicitly test EGF hypotheses. Other possibilities include

survey experiments that manipulate aspects of the EGF, for example, by including "treatments" that propose different types of forums for addressing a new collective-action problem. Perhaps even field experiments are possible, for example, by randomly selecting the intervention of a new actor like a development nongovernmental organization across a set of regions facing similar collective-action problems in which field research has measured the extant structure of the EGF. Similarly, one might consider randomly selecting regions to receive the "treatment" of a new collaborative governance process or other type of planning process like a vulnerability assessment and its resulting effects on the structure and function of a polycentric system.

The EGF approach to the study of polycentric systems has developed a rich research tradition especially in the context of environmental policy. EGF research relies on a diverse set of qualitative and quantitative methods that have been triangulated to test some core hypotheses and develop initial consensus on some common knowledge about the structure and function of polycentric systems. But even with existing methods, a no-brainer conclusion of this chapter is that we desperately need more research – comparative and longitudinal, and extended to a broader range of policy sectors. Future research will benefit from a stronger integration of more formal social science approaches such as mathematical models, causal inference, and experiments. We are optimistic that this chapter can help stimulate such research, which is especially important given the demonstrated ubiquity of complex, polycentric systems in real systems throughout the world. If society hopes to make these systems more resilient, sustainable, effective, and equitable, we need a stronger understanding of how they actually work.

References

Aligica, Paul D., and Vlad Tarko. 2012. "Polycentricity: From Polanyi to Ostrom, and Beyond." *Governance* 25(2): 237–262.

Angst, Mario. 2020. "Bottom-Up Identification of Subsystems in Complex Governance Systems." *Policy Studies Journal* 48(3): 782–805.

Angst, Mario, and Christian Hirschi. 2017. "Network Dynamics in Natural Resource Governance: A Case Study of Swiss Landscape Management." *Policy Studies Journal* 45(2): 315–336.

Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast. 2000. "The Analytic Narrative Project-Analytic Narratives." *American Political Science Review* 94(3): 696–702.

Bell, Emily, and Tyler A Scott. 2020. "Common Institutional Design, Divergent Results: A Comparative Case Study of Collaborative Governance Platforms for Regional Water Planning." *Environmental Science & Policy* 111: 63–73.

Berardo, Ramiro. 2014a. "Bridging and Bonding Capital in Two-Mode Collaboration Networks." *Policy Studies Journal* 42(2): 197–225.

Berardo, Ramiro. 2014b. "The Evolution of Self-Organizing Communication Networks in High-Risk Social-Ecological Systems." *International Journal of the Commons* 8(1): 236–258.

Berardo, Ramiro, Tanya Heikkila, and Andrea K. Gerlak. 2014. "Interorganizational Engagement in Collaborative Environmental Management: Evidence from the South Florida Ecosystem Restoration Task Force." *Journal of Public Administration Research and Theory* 24(3): 697–719.

Berardo, Ramiro, and Mark Lubell. 2016. "Understanding What Shapes a Polycentric Governance System." *Public Administration Review* 76(5): 738–751.

———. 2019. "The Ecology of Games as a Theory of Polycentricity: Recent Advances and Future Challenges." *Policy Studies Journal* 47(1): 6–26.

Berardo, Ramiro, Tomás Olivier, and Anthony Lavers. 2015. "Focusing Events and Changes in Ecologies of Policy Games: Evidence from the Paraná River Delta." *Review of Policy Research* 32(4): 443–464.

Berardo, Ramiro, and John T. Scholz. 2010. "Self-organizing Policy Networks: Risk, Partner Selection, and Cooperation in Estuaries." *American Journal of Political Science* 54(3): 632–649.

Bjørndal, Christian Thue, and Siv Gjesdal. 2020. "The Role of Sport School Programmes in Athlete Development in Norwegian Handball and Football." *European Journal for Sport and Society* 17(4): 362–384.

Bjørndal, Christian Thue, Lars Tore Ronglan, and Svein S. Andersen. 2017. "Talent Development as an Ecology of Games: A Case Study of Norwegian Handball." *Sport, Education and Society* 22(7): 864–877.

Bodin, Örjan, S. M. Alexander, et al. 2019. "Improving Network Approaches to the Study of Complex Social–Ecological Interdependencies." *Nature Sustainability* 2(7): 551–559.

Bodin, Örjan, and Daniel Nohrstedt. 2016. "Formation and Performance of Collaborative Disaster Management Networks: Evidence from a Swedish Wildfire Response." *Global Environmental Change* 41: 183–194.

Bodin, Örjan, Daniel Nohrstedt, et al. 2019. "Working at the 'Speed of Trust': Pre-Existing and Emerging Social Ties in Wildfire Responder Networks in Sweden and Canada." *Regional Environmental Change* 19(8): 2353–2364.

Brandenberger, Laurence, et al. 2020. "Boundary Spanning through Engagement of Policy Actors in Multiple Issues." *Policy Studies Journal.* https://doi.org/10.1111/psj.12404

Brewer, Devon D. 2000. "Forgetting in the Recall-Based Elicitation of Personal and Social Networks." *Social Networks* 22(1): 29–43.

Burt, Ronald S. 2005. *Brokerage and Closure: An Introduction to Social Capital.* Oxford: Oxford University Press.

Cairney, Paul. 2012. "Complexity Theory in Political Science and Public Policy." *Political Studies Review* 10(3): 346–358.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2012. "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics* 29(2): 238–249.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2): 317–372.

Crawford, S. E. S., & Ostrom, E. (1995). A Grammar of Institutions. *The American Political Science Review*, 89(3), 582–600.

Desmarais, Bruce A., and Skyler J. Cranmer. 2012. "Micro-level Interpretation of Exponential Random Graph Models with Application to Estuary Networks." *Policy Studies Journal* 40(3): 402–434.

Fischer, Manuel, and Simon Maag. 2019. "Why Are Cross-Sectoral Forums Important to Actors? Forum Contributions to Cooperation, Learning, and Resource Distribution." *Policy Studies Journal* 47(1): 114–137.

Fischer, Manuel, and Pascal Sciarini. 2016. "Drivers of Collaboration in Political Decision Making: A Cross-Sector Perspective." *The Journal of Politics* 78(1): 63–74.

Fisher, Dana R., and Philip Leifeld. 2019. "The Polycentricity of Climate Policy Blockage." *Climatic Change* 155(4): 469–487.

Guerrero, Angela M., Örjan Bodin, Ryan R.J. McAllister, and Kerrie A. Wilson. 2015. "Achieving Social-Ecological Fit through Bottom-up Collaborative Governance: An Empirical Investigation." *Ecology and Society* 20(4): 41.

Hamilton, Matthew, and Mark Lubell. 2018. "Collaborative Governance of Climate Change Adaptation across Spatial and Institutional Scales." *Policy Studies Journal* 46(2): 222–247.

Hamilton, Matthew, and Mark Lubell. 2019. "Climate Change Adaptation, Social Capital, and the Performance of Polycentric Governance Institutions." *Climatic Change* 152: 307–326.

Hamilton, Matthew, Mark Lubell, and Emilinah Namaganda. 2018. "Cross-Level Linkages in an Ecology of Climate Change Adaptation Policy Games." *Ecology and Society* 23(2): 36.

Heclo, Hugh. 1978. "Issue Networks and the Executive Establishment." In *The New American Political System*, edited by Anthony King. Washington, DC: American Enterprise Institute (pp. 413–422).

Hedeker, Donald, Robert D. Gibbons, and Brian R. Flay. 1994. "Random-Effects Regression Models for Clustered Data with an Example from Smoking Prevention Research." *Journal of Consulting and Clinical Psychology* 62(4): 757.

Hedlund, Johanna, Örjan Bodin, and Daniel Nohrstedt. 2020. "Policy Issue Interdependency and the Formation of Collaborative Networks." *People and Nature*. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1002/pan3.10170 (January 6, 2021).

Heikkila, T., Weible, C.M., Olofsson, K.L., Kagan, J.A., You, J. and Yordy, J., 2021. The structure of environmental governance: How public policies connect and partition California's oil and gas policy landscape. *Journal of Environmental Management*, *284*, p.112069.

Jasny, Lorien. 2012. "Baseline Models for Two-Mode Social Network Data." *Policy Studies Journal* 40(3): 458–491.

Jasny, Lorien, and Mark Lubell. 2015. "Two-Mode Brokerage in Policy Networks." *Social Networks* 41: 36–47.

Klasic, Meghan, and Mark Lubell. 2020. "Collaborative Governance: From Simple Partnerships to Complex Systems." In *Handbook of US Environmental Policy*, edited by David M. Konisky. Cheltenham: Edward Elgar Publishing (pp. 257–274).

Laird, Nan M., and James H. Ware. 1982. "Random-Effects Models for Longitudinal Data." *Biometrics 38*(4): 963–974.

Leifeld, Philip. 2010. *Discourse Network Analyzer (DNA) Manual*. Bonn: Max Planck Institute for Research on Collective Goods. www.philipleifeld.de

Leifeld, Philip, Skyler J. Cranmer, and Bruce A. Desmarais. 2018. "Temporal Exponential Random Graph Models with Btergm: Estimation and Bootstrap Confidence Intervals." *Journal of Statistical Software* 83(6). www.jstatsoft.org/v83/i06/ (January 3, 2019).

Levin, Simon, Tasos Xepapadeas, Anne-Sophie Crépin, Jon Norberg, Aart de Zeeuw, Carl Folke, Terry Hughes, et al. 2013. "Social-Ecological Systems as Complex Adaptive Systems: Modeling and Policy Implications." *Environment and Development Economics* 18(2): 111–132.

Levy, Michael A., Mark N. Lubell, and Neil McRoberts. 2018. "The Structure of Mental Models of Sustainable Agriculture." *Nature Sustainability* 1(8): 413–420.

Long, Norton. 1958. "The Local Community as an Ecology of Games." *American Journal of Sociology* 64(3): 251–261.

Lubell, Mark. 2013. "Governing Institutional Complexity: The Ecology of Games Framework." *Policy Studies Journal* 41(3): 537–559.

———. 2017. *The Governance Gap: Climate Adaptation and Sea-Level Rise in the San Francisco Bay Area*. Davis, CA: University of California.

Lubell, Mark, Andrea Gerlak, and Tanya Heikkila. 2013. "CalFed and Collaborative Watershed Management: Success Despite Failure?" In *Making Space for the River: Governance Experiences with Multifunctional River Flood Management in the US and Europe*, edited by Jeroen Frank Warner, Arwin van Buuren, and Jurian Edelenbos. London: IWA Publishing (pp. 63–78).

Lubell, Mark, A. D. Henry, and M. McCoy. 2010. "Collaborative Institutions in an Ecology of Games." *American Journal of Political Science* 54(2): 287–300.

Lubell, Mark, Jack Mewhirter, and Ramiro Berardo. 2020. "The Origins of Conflict in Polycentric Governance Systems." *Public Administration Review* 80(2): 222–233.

Lubell, Mark, Jack Mewhirter, Ramiro Berardo, and John T. Scholz. 2017. "Transaction Costs and the Perceived Effectiveness of Complex Institutional Systems." *Public Administration Review* 77(5): 668–680.

Lubell, Mark, and Matthew Robbins. 2021. "Adapting to Sea-Level Rise: Centralization or Decentralization in Polycentric Governance Systems?" *Policy Studies Journal* Forthcoming.

Lubell, Mark, Garry Robins, and Peng Wang. 2014. "Network Structure and Institutional Complexity in an Ecology of Water Management Games." *Ecology and Society* 19(4). www.ecologyandsociety.org/vol19/iss4/art23/

Lubell, Mark, John Scholz, Ramiro Berardo, and Garry Robins. 2012. "Testing Policy Theory with Statistical Models of Networks." *Policy Studies Journal* 40(3): 351–374.

Lubell, Mark, Francesca Pia Vantaggiato, and Darcy J Bostic. 2019. *The Governance of Sea Level Rise in the San Francisco Bay Area: Results from a Survey of Stakeholders*. Davis, CA: University of California. https://bcdc.ca.gov/cm/2019/0516SLRSurveyReport.pdf

Madison, Michael. 2021. "Fair Play: Notes on the Algorithmic Soccer Referee." *Vanderbilt Journal of Entertainment & Technology Law*. https://scholarship.law.pitt.edu/fac_articles/299

Mancilla García, María, and Örjan Bodin. 2019. "Participation in Multiple Decision Making Water Governance Forums in Brazil Enhances Actors' Perceived Level of Influence." *Policy Studies Journal* 47(1): 27–51.

McAllister, Ryan, Heleen Kruger, Nyree Stenekes, and Robert Garrard. 2020. "Multilevel Stakeholder Networks for Australian Marine Biosecurity: Well-Structured for Top-down Information Provision, Requires Better Two-Way Communication." *Ecology and Society* 25(3): 18. www.ecologyandsociety.org/vol25/iss3/art18/ (December 18, 2020).

McAllister, Ryan R.J., Rod McCrea, and Mark N. Lubell. 2014. "Policy Networks, Stakeholder Interactions and Climate Adaptation in the Region of South East Queensland, Australia." *Regional Environmental Change*. https://agris.fao.org/agris-search/search.do?recordID=US201400057535 (January 27, 2021).

McAllister, Ryan R.J., Catherine J. Robinson, Alinta Brown, Kirsten Maclean, Suzy Perry, and Shuang Liu. 2017. Balancing Collaboration with Coordination:

Contesting Eradication in the Australian Plant Pest and Disease Biosecurity System. *International Journal of the Commons* 11(1): 330–354.

McLaughlin, Danielle M., Jack M. Mewhirter, and Mark Lubell. 2021. "The Spread of Conflict in Polycentric Governance Systems." *Journal of Public Administration Research and Theory*, 2021, muab045, https://doi.org/10.1093/jopart/muab045.

McLaughlin, Danielle M., Jack M. Mewhirter, James E. Wright II, and Richard Feiock. 2020. "The Perceived Effectiveness of Collaborative Approaches to Address Domestic Violence: The Role of Representation, 'Reverse-Representation,' Embeddedness, and Resources." *Public Management Review* 23(12): 1808–1832. https://doi.org/10.1080/14719037.2020.1774200

McNeish, Daniel M, and Laura M Stapleton. 2016. "The Effect of Small Sample Size on Two-Level Model Estimates: A Review and Illustration." *Educational Psychology Review* 28(2): 295–314.

Mewhirter, Jack, and Ramiro Berardo. 2019. "The Impact of Forum Interdependence and Network Structure on Actor Performance in Complex Governance Systems." *Policy Studies Journal* 47(1): 159–177.

Mewhirter, Jack, Eric A. Coleman, and Ramiro Berardo. 2019. "Participation and Political Influence in Complex Governance Systems." *Policy Studies Journal* 47(4): 1002–1025.

Mewhirter, Jack, Mark Lubell, and Ramiro Berardo. 2018. "Institutional Externalities and Actor Performance in Polycentric Governance Systems." *Environmental Policy and Governance* 28(4): 295–307.

Mewhirter, Jack, and D. McLaughlin. 2021. "The Pitfalls Associated with More Intensive Engagement in Collaborative Forums: The Role of Behavioral Spillovers and Cognitive Load." *Journal of Behavioral Public Administration* 4(1). http://dx.doi.org/10.30636/jbpa.41.166

Mewhirter, Jack, Danielle M. McLaughlin, and Manuel Fischer. 2019. "The Role of Forum Membership Diversity on Institutional Externalities in Resource Governance Systems." *Society & Natural Resources* 32(11): 1239–1257.

Moulton, Brent R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *The Review of Economics and Statistics* 72(2): 334–338.

Murphy, John T. et al. 2014. "Water Relationships in the U.S. Southwest: Characterizing Water Management Networks Using Natural Language Processing." *Water* 6(6): 1601–1641.

Nohrstedt, Daniel. 2018. "Bonding and Bridging Relationships in Collaborative Forums Responding to Weather Warnings." *Weather, Climate, and Society* 10(3): 521–536.

Nohrstedt, Daniel, and Örjan Bodin. 2014. "Evolutionary Dynamics of Crisis Preparedness Collaboration: Resources, Turbulence and Network Change in Swedish Municipalities." *Risk, Hazards & Crisis in Public Policy* 5(2): 134–155.

———. 2020. "Collective Action Problem Characteristics and Partner Uncertainty as Drivers of Social Tie Formation in Collaborative Networks." *Policy Studies Journal* 48(4): 1082–1108.

Olivier, Tomás. 2019. "How Do Institutions Address Collective-Action Problems? Bridging and Bonding in Institutional Design." *Political Research Quarterly* 72(1): 162–176.

Olivier, Tomás, and Ramiro Berardo. 2021. "Birds of a Feather Fight Together: Forum Involvement in a Weakly Institutionalized Ecology of Policy Games." *Policy Studies Journal*. https://onlinelibrary.wiley.com/doi/abs/10.1111/psj.12418 (January 6, 2021).

Ostrom, Elinor. 2009. "A General Framework for Analyzing Sustainability of Social-Ecological Systems." *Science* 325(5939): 419–422.

Primo, David M, Matthew L. Jacobsmeier, and Jeffrey Milyo. 2007. "Estimating the Impact of State Policies and Institutions with Mixed-Level Data." *State Politics & Policy Quarterly* 7(4): 446–459.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: STATA Press.

Robins, Garry. 2015. *Doing Social Network Research. Network-Based Research Design for Social Scientists.* London: SAGE Publications Ltd.

Robins, Garry, and M. Alexander. 2004. "Small Worlds among Interlocking Directors: Network Structure and Distance in Bipartite Graphs." *Computational & Mathematical Organization Theory* 10(1): 69–94.

Robins, Garry, Lorraine Bates, and Philippa Pattison. 2011. "Network Governance and Environmental Management: Conflict and Cooperation." *Public Administration* 89(4): 1293–1313.

Robins, Garry, Jenny M. Lewis, and Peng Wang. 2012. "Statistical Network Analysis for Analyzing Policy Networks." *Policy Studies Journal* 40(3): 375–401.

Schlager, Edella, and Elinor Ostrom. 1992. "Property-Rights Regimes and Natural-Resources – A Conceptual Analysis." *Land Economics* 68(3): 249–262.

Scott, Tyler A. 2016. "Analyzing Policy Networks Using Valued Exponential Random Graph Models: Do Government-Sponsored Collaborative Groups Enhance Organizational Networks?" *Policy Studies Journal* 44(2): 215–244.

Scott, Tyler A., and Craig W. Thomas. 2017. "Winners and Losers in the Ecology of Games: Network Position, Connectivity, and the Benefits of Collaborative Governance Regimes." *Journal of Public Administration Research and Theory* 27(4): 647–660.

Scott, Tyler A., Nicola Ulibarri, and Ryan P. Scott. 2020. "Stakeholder Involvement in Collaborative Regulatory Processes: Using Automated Coding to Track Attendance and Actions." *Regulation & Governance* 14(2): 219–237.

Siddiki, S., Heikkila, T., Weible, C. M., Pacheco-Vega, R., Carter, D., Curley, C., Deslatte, A., & Bennett, A. (2019). Institutional Analysis with the Institutional Grammar. *Policy Studies Journal*, early view.

Smaldino, Paul E., and Mark Lubell. 2011. "An Institutional Mechanism for Assortment in an Ecology of Games." *PLOS ONE* 6(8): e23019.

———. 2014. "Institutions and Cooperation in an Ecology of Games." *Artificial Life* 20(2): 207–221.

Stock, James H, and Mark W Watson. 2008. "Heteroskedasticity-robust Standard Errors for Fixed Effects Panel Data Regression." *Econometrica* 76(1): 155–174.

———. 2015. *Introduction to Econometrics*. London: Pearson.

Ulibarri, Nicola. 2015. "Collaboration in Federal Hydropower Licensing: Impacts on Process, Outputs, and Outcomes." *Public Performance & Management Review* 38(4): 578–606.

Ulibarri, Nicola, and Tyler A. Scott. 2017. "Linking Network Structure to Collaborative Governance." *Journal of Public Administration Research and Theory* 27(1): 163–181.

Vantaggiato, Francesca Pia, and Mark Lubell. 2021. "Functional Differentiation in Governance Networks for Climate Adaptation: Sea Level Rise in the San Francisco Bay Area." Forthcoming manuscript.

Wang, Peng, Garry Robins, and Petr Matous. 2016. "Multilevel Network Analysis Using ERGM and Its Extension." In *Multilevel Network Analysis for the Social Sciences: Theory, Methods and Applications*, Methodos Series, edited by Emmanuel

Lazega and Tom A.B. Snijders. Cham: Springer International Publishing (pp. 125–143). https://doi.org/10.1007/978-3-319-24520-1_6 (January 27, 2021).

Wang, Peng, Garry Robins, Philippa Pattison, and Emmanuel Lazega. 2013. "Exponential Random Graph Models for Multilevel Networks." *Social Networks* 35(1): 96–115.

Weible, Christopher M., Jill Yordy, Tanya Heikkila, Hongtao Yi, Ramiro Berardo, Jennifer Kagan, and Catherine Chen. 2020. "Portraying the Structure and Evolution of Polycentricity via Policymaking Venues." *International Journal of the Commons* 14(1): 680-691.

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica: Journal of the Econometric Society* 48(4): 817–838.

Widmer, A., Herzog, L., Moser, A., & Ingold, K. (2019). Multilevel water quality management in the international Rhine catchment area: How to establish social-ecological fit through collaborative governance. *Ecology and Society*, 24(3), 27.

Winship, C., and S. L. Morgan. 1999. "The Estimation of Causal Effects from Observational Data." *Annual Review of Sociology* 25: 659–707.

10 The Evaluation and Advancement of Policy Process Research

Christopher M. Weible and Samuel Workman

Introduction

This volume arises from the argument that advancing policy process research requires the interplay of theories and methods and their ongoing and iterative advancement. We use "methods" generally to refer to the systematic techniques or tools (e.g., research design, data collection, data analysis) to apply theories of the policy process and their concepts.[1] This volume is also about the justifications or rationales for using such techniques and tools, often termed "methodology" (Gerring, 2001; Blaikie, 2007).

As an equally important text to *Theories of the Policy Process,* the *Methods of the Policy Processes* brings attention to how we "apply" theories. When we talk to students or any scholar new to the study of theory-based policy process research, they are often unsure of what it means to "apply" a theory. To help convey what it means to apply a theory, we have used metaphors to help the reader understand how this occurs. Paul Sabatier, for example, used the term "lenses" to suggest that we use different theories to view and simplify aspects of the world through languages of concepts, assumptions of causalities, and arguments of how these concepts interact under various conditions (Sabatier, 1999a).

Thus, applying a theory means using one to help make sense, understand, describe, explain, and contribute knowledge about policy processes and, perhaps, inform their politics and governance. Of course, we hold much of the knowledge gained through these efforts as individuals and groups. However, even more importantly, the knowledge gained from empirical applications, we at least try to incorporate back into our theories (including their ontologies and related epistemologies) and methods (including their associated methodologies). In other words, our theories and methods – at least in the portion of policy process research represented in this volume – serve as reservoirs of knowledge. Hence, we pay attention to, and are conscious of, how we apply and develop them.

Policy process research seeks to contribute knowledge about the interactions that occur over time between public policies and the surrounding actors, events, contexts, and outcomes. When any of us observes policy processes, we partly form our understandings from our life experiences,

mental faculties, or what some might call "common sense." However, the scientific method sets policy process research apart as a social science. As we discuss in this chapter, the scientific method has provided a means to conduct research and justify knowledge claims. Yet, the way we use the scientific method often departs from the textbook models, creating a gap between what we say and what we do. This chapter aims to close this gap.

We begin this chapter by summarizing what it means to design policy process research by recapping the chapters. Next, we return to this field's scientific method, evaluate it, and describe some of its challenges. We then ask the pointed question: how can we do better? Finally, we respond by stating the principles for good scientific research for this field and offering ideas for advancing it. In all, policy process research has contributed to our knowledge about policy processes. However, while we brought the need for better theories to the foreground, we also pushed to the background the need for better methods. Our argument is not that we should elevate one over the other. On the contrary, we need to promote both, which inevitably entails returning to the foundations of our science. Thus, to ensure continued contributions in knowledge, we need to engage our ontological, epistemological, methodological, and methods orientations.

Summarizing Methods Across the Theories

We began this volume with a roadmap for reasoning, essentially a set of strategies for thinking about the chapters. This roadmap includes (1) research questions and scope, (2) research design as comparison, (3) measurement, and (4) tools of analysis. In surveying the chapters based on these strategies and the subcategories within them, several themes emerged within and across chapters. We summarize these themes immediately below.

First, we see commonalities across the chapters. Despite some of the theories emerging in some topical areas and favoring those areas in the past, today, all theories show comprehensive coverage of different topics. For example, although the Advocacy Coalition Framework has numerous studies involving environmental and energy issues, it now covers the gamut.

Second, we see differences across chapters in their level of governance. Punctuated Equilibrium Theory, for example, focuses on agendas and outputs of policymaking venues with generalizations drawn to the political system (and sometimes subsystems) (Workman et al., 2021). Innovation and Diffusion (Karch, 2021) and the Ecology of Games Framework (Lubell et al., 2021) represent other approaches that focus on the policymaking venues and how they operate in concert. In contrast, the Institutional Analysis and Development Framework (Schlager et al., 2021), the Advocacy Coalition Framework (Henry et al., 2021), and Multiple Streams Framework (Herweg et al., 2021) all deal with the loci of political activity of the action situation, the policy subsystem, and the policy community, respectfully. For these three approaches, the analysis can be at almost any level of governance. The Narrative Policy Framework deals with micro, meso, and macro levels of

analysis, though most research deals with the former two (Jones et al., 2021). Similarly, Policy Feedback Theory spans governing institutions (including actors and elites), interest groups and organizations, and mass publics (SoRelle and Michener, 2021).

Third, the methods vary in dealing with comparisons with some relationship between scope and comparability. For example, Punctuated Equilibrium Theory focuses on a policymaking venue mainly at the national level of governance, enabling relatively easy and direct comparison across countries, as reflected in its established methods in the Comparative Agendas Project.[2] In contrast, approaches that deal with multiple venues simultaneously, policy subsystems, policy communities, and action situations conduct work comparatively. However, given that their units of analysis are harder to identify and vary across contexts, they face additional hurdles in generalizing their results. Similar arguments emerge about Policy Feedback Theory; its focus on understanding the resources and interpretive effects on society makes it hard to conduct comparatively using the same methods.

Fourth, one of the principal challenges in policy process research, akin to many social sciences, is creating, developing, and communicating concepts and, ultimately, measuring them. We see these challenges across all theories. However, the challenge is starkest in all the theories except Punctuated Equilibrium Theory, which tends to measure topics and budgets to draw inferences about other concepts. The challenge in measurement reflects the need to share survey, coding, and interview instruments. Indeed, over the last decade, we have seen significant advances in coding narratives (Shanahan et al., 2018) and institutions (Siddiki et al., 2019). While there are many ways to advance our science, developing best and established practices in methods is undoubtedly one of them, as we elaborate further below.

Fifth, given the value of quantitative and qualitative methods in providing more comprehensive forms of knowledge, we find the potential to conduct both in all theories. Still, some theories lean toward one over the other. The most balanced approaches tend to be Institutional Analysis and Development Framework, the Advocacy Coalition Framework, and the Narrative Policy Framework. These three have many qualitative and quantitative applications. Leaning more qualitative and needing more quantitative methods are the Multiple Streams Framework and Policy Feedback Theory. Leaning more quantitative and needing more qualitative methods are Ecology of Games Framework, Punctuated Equilibrium Theory, and Innovation and Diffusion. Of these, Innovation and Diffusion might be of the most significant need for qualitative analysis of its posited causal mechanisms. All theories could do better quantitatively and qualitatively, yet we know that the gaps are wider in some than others. Additionally, all theories call for more and better quantitative and qualitative methods; the question now is who will answer the call.

Finally, the theories differ in their tendency to balance reliability and validity. Corresponding to the theories' leaning toward quantitative and qualitative approaches, theories that lean toward quantitative methods favor reliability over validity. Theories that lean toward qualitative methods favor

validity over reliability. Theories that lean toward both quantitative and qualitative methods tend to balance them. Given that reliability and validity serve as indicators of quality, the lack of growth in areas that run counter to the leaning of any given research program might be because of their favoring reliability over validity or vice versa in the conduct of their research or in reviewing others' research. We return to the importance of cultivating diverse research programs later in this chapter.

Scientific Principles for Contributing Knowledge to Policy Process Research

Social science involves using scientific processes or methods for forming understandings and knowledge about people and their interactions with each other and the world. Recognizing the multiple ways of conducting social science and the tendency for some to approach the underlying assumptions with indifference, hesitation, and ambivalence, we offer what we consider four general scientific principles for contributing to knowledge about policy process research (see also Table 10.1).

1   *Transparent methodological pluralism based on justified best practices (at a minimum) and established practices (at the maximum).*

Methodology refers to the rationales for choosing particular methods for empirical inquiry (Gerring, 2001; Hay, 2002; Blaikie, 2007). For policy process research, our fundamental methodological question is: *How can we know policy processes, and what are our rationales for doing so?*[3]

To answer this question, the research programs associated with each policy process theory employ methodological pluralism. As described by Little (2016, p. 231),

> For any given empirical question, there will be a variety of methods on the basis of which to investigate this problem. And, ideally we should select a set of tools that are well suited to the particular characteristic of the problem at hand. In other words, analysis of the situation of scientific research into the unknown would suggest methodological pluralism. We should be open to a variety of tools and methods and should design

Table 10.1 Scientific principles for contributing knowledge in theory-based policy process research

1. *Transparent methodological pluralism based on justified best practices and established practices*
2. *Bounded epistemological pluralism transparent in justifying contributions to knowledge*
3. *Bounded ontological pluralism with transparency in its creation, structure, and continuation*
4. *Diversity in theory-based research programs*

research in a way that is closely tailored to the nature of the empirical problem.

More simply, this first principle points to the openness to various methods and the need to justify their uses (see also Gerring, 2001).

The methods associated with policy process theories include embracing varied relationships between researchers and their phenomena, from basic science where researchers detach themselves from their subjects to various forms of engaged scholarship, such as co-producing knowledge, ethnographic approaches, and advocacy (Van de Ven, 2007). Our methods incorporate various research designs, from natural experiments to one-shot case studies or field research to computational simulations, such as agent-based modeling. They involve quantitative to qualitative approaches done abductively, inductively, and deductively. Data sources can vary from primary sources, such as observations, surveys, and interviews, to secondary sources, such as public documents, social and news media, and other official records. Our analyses correspond with our data, from quantitative to qualitative techniques. Lastly, the research programs centered on the theories in this volume have conducted their research using a combination of the above through mixed methods and efforts of triangulation.

The foundational stone that grounds methodological pluralism is transparency, which requires methods open to scrutiny and replication. Transparency echoes longstanding norms where "the procedures are public" (King et al., 1994, p. 8) to "be clear enough to be proven wrong" (Sabatier, 1999b, p. 5). The importance of transparency lies in its potential to invite evaluation, criticism, and replication. In other words, it enables us to learn from each other, including our rights and wrongs. It includes efforts of falsification and error seeking. Thus, we should be clear in stating the assumptions and rationales for a research design. We should make as public as possible any adjustments while implementing our research and how we created and used interview protocols, survey instruments, or code forms and analyzed the data.

While the choice and justification of methods depend partly on the research question or objective, resource availability, and the researchers' values, skills, and preferences, methodological pluralism must occur with a justification of the choice of methods in conducting any empirical inquiry. Our justifications can rely on established social science procedures. For example, the selection of case studies in a research design might invoke Gerring (2006). Alternately, our justification might rely on a more realistic version of the scientific method that adheres to the established procedures of doing science while mitigating its known flaws, inconsistencies, and liabilities.

Our justifications also might rely on – or even better improve on – a theory's best practices and established practices. "Best practices" refer to the general guidance, ideas, strategies, and approaches currently used in applying a theory in a given context that is considered adequate, at least for now. Every theoretical chapter in this volume offered best practices. "Established practices" refer to well-specified methods used in theory for

data collection and analysis in a given context. Punctuated Equilibrium Theory's Comparative Agendas Project provides an excellent example of an established practice. Not every theory offers established procedures for application, yet all of them should strive for them.

This volume offers the first compilation of the best and established practices for the most established theories of the policy process. By elevating these practices, we want to promote communication on methodology and methods and, thus, provide opportunities to contribute knowledge in these areas. At the same time, we want to avoid establishing rigid orthodoxies in methodologies or methods. The spirit behind making these best and established practices public is sharing what we know about applying the theories to do better in the future.

One inevitable challenge in championing transparency is avoiding dogmatism that might limit our choice of methods and, thus, our forms of knowledge. We often invoke transparency in supporting some research while denouncing others, especially studies that differ in balancing the tradeoffs between validity and reliability. For example, we might reward the reliable coding techniques for policy narratives under the Narrative Policy Framework (Shanahan et al., 2018) and institutional statements under the Institutional Analysis and Development Framework (Siddiki et al., 2019). Contrariwise, we might mistakenly overlook or critique approaches that emphasize validity over reliability, such as some research under Policy Feedback Theory. Our point is not to forgo reliability for validity or vice versa. Our point is to recognize the tradeoffs and know that some research might favor one over the other and maintain receptivity to different methodologies and methods.

2   *Bounded epistemological pluralism transparent in justifying contributions to knowledge.*

Epistemology is the study of knowledge (Hay, 2002; Blaikie, 2007). It directs us to ask a fundamental question about what it means to advance policy process research: *what can we know about policy processes, and how can we justify our understanding (beliefs) about policy processes?* More generally, epistemology raises questions of what we know and how we know it. We can think about our knowledge about policy processes by the words in our research questions: who, where, what, how, and why. Ultimately, elevating challenges in epistemology entails what it means to build or contribute knowledge about policy processes as individuals and groups. Consistent with methodological pluralism, a form of epistemological pluralism ensues for supporting multiple ways of contributing to understanding and, hence, numerous ways of knowing. Such conditions of epistemological pluralism correspond to similar developments in interdisciplinary research that support diverse forms of knowledge, which are research efforts paralleling our own (e.g., see Miller et al., 2008).

The epistemological orientation of policy process theories traditionally relies on a combination of empiricism and falsificationism as part of their

traditional scientific method. Often equated with the positivist paradigm, this traditional scientific method entailed iterative or cyclical iterations between theories and derived hypotheses, empirically testing hypotheses, confirming or refuting those hypotheses, and confirming, updating, or refuting our theories.[4] The underlying assumption was that our truth claims would be tentative as long as theories and hypotheses remained unrefuted.

This epistemological orientation influenced our methodologies and methods as well as our norms of the field, as captured in the phrase to "be clear enough to be proven wrong" (Sabatier, 1999b, p. 5). Accordingly, the methodology favored objectivity of the researcher with distance from their subjects and overcoming biases with inter-subject reliability tests. Using this scientific method, we relied on its conceptual language and causal assumptions. We posited conceptual interactions in our theories to simplify the world, describe and explain it, and make inferences or sense of it.

In practice, the epistemological orientations deviate from the traditional positivist paradigm by drawing on ideas from a combination of neorealism, rationalism, even some features of constructivism and conventionalism, along with the traditional empiricism and falsificationism (Blaikie, 2007). This deviation holds in policy process research as a whole and even within the same theory.

These diverse epistemological orientations have emerged for several reasons. They result from the inevitable changes in practices from the flaws in verification and falsificationism (see, for example, Hawkesworth, 1992).[5] They result from deliberate modifications to critical rationalism, such as implementing Lakatos' (1970) ideas in supporting our research programs. They result from the aforementioned changes in the relationship between researchers and the researched phenomena, including engaged scholarship and reflexivity (Van de Ven, 2007). They result from the recognized absence of covering laws yet the possibilities for some patterns, such as the likely tendency to find stochastic regularities rather than deterministic point predictions (Jones and Baumgartner, 2005; Little, 2016). They result from the growing unease with objectivism in our science, from challenges in descriptive inferences in our measures to the theory-laden dependence of our knowledge. Moreover, they result from the increasing recognition in the presuppositions we bring to our research and the influence our theories bring in shaping the language we use, what we then see and think, and how the effects of our language feedback into our theories and knowledge.

In the philosophy of epistemology, knowledge claims have been associated with a statement or idea that is factual, believed with confidence, and justi-fied (Blaikie, 2007). However, the underlying arguments about the notion of justified-true-beliefs remain unsettled and debated (e.g., Gettier, 1963), and waiting for resolution of these issues is not our best strategy. Exacerbating this problem, theories of the policy process serve as reservoirs of knowledge and a means for communication to students and new and experienced scholars inside and outside the field. We update these theories based on the accu-mulation of empirical evidence and, sometimes, reasoned argumentation, as

found in the multiple iterations of *Theories of the Policy Process*. More critically, how this updating happens is often unclear and typically subject to the theoretical leaders found among the authors of this volume and *Theories of the Policy Process*. One of our recommendations is to clarify how this updating occurs.

Time to bring attention to our epistemological orientations and the question of what we can know about policy processes, and how can we justify our understanding (beliefs) about policy processes? Relevant to policy process research, our knowledge is implicitly or explicitly comparative and contributes to knowledge claims through three major pathways: (1) multiple direct observations of a phenomenon, (2) descriptive and causal inferences of unobserved phenomena based on multiple indirect observed evidence, and (3) making sense of the world by various constructs or constructions of a phenomenon (King et al., 1994). In all three, transparency entailing an openness to scrutiny and some degree of replicability should guide the methods. We can also describe the knowledge claims by their tentativeness, uncertainty, and generalized or localized tendencies. In all three pathways, knowledge claims become more convincing with evidence conducted by a diversity of researchers and drawn from a plurality of methodologies and methods that span varied contexts.

Punctuated Equilibrium Theory provides a good illustration of the first two pathways. It offers transparent methods with multiple instances of observations of patterns of punctuated and incremental change in agendas and public policies over time. It provides inferences about unobserved causal mechanisms via slip-stick dynamics and bottlenecks in attention drawn from empirical evidence outside of this field about the faculties of human and organizational information processing. It is also conducted by a diversity of scholars in varied contexts. This is an example of a theory that "discovers" something about policy processes. Multiple Streams Framework provides a good illustration of the third with its concept of "policy windows" or "windows of opportunity." The three streams and multiple indicators within each plus the policy entrepreneur constitute the definition of a policy window, making its general hypothesis more definitional than causal. We do not observe windows of opportunity directly or infer them from data. Instead, we identify components of the different streams and policy entrepreneurs to make claims about the presence of a window of opportunity, usually post hoc. In essence, the Multiple Streams Framework "creates" windows of opportunity to make sense of an otherwise complex policy process.

While we agree with Sabatier (1999a) about the unfalsifiability of the general hypothesis in the Multiple Streams Framework, we disagree that this is a significant flaw. Indeed, many hypotheses in *Theories of the Policy Process* are unfalsifiable. Instead, the Multiple Streams Framework's policy window concept contributes to our knowledge by helping us make sense of complex policy processes that otherwise baffle our minds and limit our attempts to engage. Such a contribution is good enough for us.

We, thus, build knowledge from a practical scientific method that uses hypotheses, propositions, principles, and other relational forms involving our concepts to falsify, verify, or help make sense of our surroundings. In other words, our theories primarily aid us in communicating something about the world that allows others to argue for or against, to confirm or refute, and view as more or less helpful in making sense of policy processes. This practical scientific method then becomes more about being transparent and open to scrutiny and replication to others in the research community than about strict adherence to the assumptions inherent in the traditional scientific method.

Becoming open to a pluralistic epistemology that supports multiple ways of knowing also helps keep us from learning in only one way. Yet, its boundedness reminds us that understandings are possible in generalizable and localizable forms while avoiding strict or extreme relativism. Our epistemological transparency means we need to be clear in our knowledge-making processes. We can then sometimes conduct our science for science's sake in contributing to theories and methods in making bounded yet generalizable claims and for practical purposes in helping understand localized knowledge claims in a particular locale and possibly to inform its policy processes.

Translating justifiable knowledge of our research into reservoirs of knowledge in theories and methods remains an important challenge. In other words, how are the chapters in this volume and the *Theories* volume revised and updated? First, it should involve more than one person (but probably not everyone). Second, it should lay out the evidence supporting the knowledge claims used to update a theory or method. Third, it should describe the processes by which this happens. Lastly, our epistemological norms must be transparent open to scrutiny, not just to the research program that supports a theory but also to people outside it. Indeed, the best way to mitigate threats to the scientific method and ensure contributions to knowledge is transparency and engagement by a diversity of scholars who continue to ask why they should believe the claims of study or studies and what justifications support such claims.

3   *Bounded ontological pluralism with transparency in its creation, structure, and continuation.*

Ontology is the study of being, what is, or what exists (Hay, 2002; Blaikie, 2007). We ask the following ontological question: *What is the nature of policy processes?* The ontology of the policy process includes issues of what exists and what are their attributes and interactions. Policy process theories have served as this field's ontologies but without referencing it as such (exception includes Poteete et al., 2010). However, by not exploring some of the more basic ontological questions, our theories have not developed as much as they could have.

As examples of midrange theories (Merton, 1967), each policy process theory specifies their most valuable types of questions and contexts

for applications as part of their scope. In their descriptions of the nature of policy processes, theories inevitably involve a language in concepts that enable communication within a research program. Such languages also influence what we observe in conducting our research and are informed and revised by our research. Our concepts also do not all serve the same functions. Some serve as phenomena to measure and study, others serve to help direct research design, and still, others serve to help infer patterns without direct measurement. We postulate interactions among some concepts in hypotheses or other types of relational forms, such as propositions or simply in argumentation. Ontologies include causal drivers, like assumptions of agency in the "model of the individual" and the interplay between agency and structure. Following Little (2016, p. 2), "Ontological thinking is really a form of empirically informed theorizing, at a fairly abstract level."

In policy process research, we see a combination of various forms of objectivism or realism (critical, shallow, conceptual, cautious, depth, and subtle) and hints of moderate idealism (Blaikie, 2007). These ontologies generally assume the existence of an external world but vary in this world's characteristics with the researcher (e.g., dependence vs. independence), its observables and unobservables, its underlying structures and mechanisms, and its commonalities and differences in its generalizations. Additionally, ontological pluralism helps avoid boxing in the field by a rigid ontology that constrains methodologies, methods, epistemologies, and knowledge about policy processes. Any differences and inconsistencies in ontologies then serve as points of tension and argumentation about the nature of policy processes and help maintain skepticism about where we are and where we are going. Yet, we also bound our ontological pluralism away from extreme idealism that essentially denies existence (atheistic idealist).

Some ontological assumptions used in policy process research necessitate mindfulness and tolerance of their assumptions. For example, in contrast to a shallow realist's association with naturalism or the unity of science, the human-based systems in policy processes differ from natural systems (e.g., Hawkesworth, 1992; Flyvbjerg, 2001; Little, 2016). People's thoughts, motivations, understandings, norms and rules, and cultures separate policy processes from natural systems. Following Little (2016), differences occur in the heterogeneity within concepts, the characteristics (e.g., consistency) of any presumed causality, the ongoing plasticity over time, and the contingency of behaviors, either observed in points in space and time or in distributions in space and time. This is one of the reasons for the absence of covering laws in policy process research.

However, despite the differences between social and natural systems, some concepts, methodologies, and methods used in natural systems can help research policy processes. For example, we might adopt research techniques from the natural sciences wherein we try to detach ourselves from the phenomenon, use quantitative methods of data collection and models of analysis at elevated levels of abstraction, assume stability in our concepts over time, and allow internal heterogeneity within concepts. We might then identify

broad regularized patterns across space and time. Such a study might assume many if not all of the assumptions of a shallow realist. At the same time, for example, epistemological and methodological pluralism allows us to explore the same phenomenon using more in-depth case studies, perhaps, using a form of engaged scholarship and ethnographic qualitative methods, where we assume human heterogeneity, cognitive frailties, and tendencies to construct realities as found in a cautious realist ontology. While these arguments resemble mixed methods and triangulation, we emphasize the healthy tension between and needed tolerance of different ways of conducting our science to capitalize on ontological, epistemological, and methodological pluralism.

The original creation of theories plants the seeds of their assumptions, concepts, and purposes in their ontologies, thereby creating the foundation for any future adjustments. All the theories in this volume emerged in the United States by western scholars and, thus, have embedded democratic values, interest–group pluralism, western culture, and even the personalities and styles of scholarship of their creators. These foundations have been recognized as a source of challenges and opportunities. Even within the context of the United States, how the theories emerged often established their particular ontological structures. Prime examples include national-level analyses based in Washington D.C. in creating the foundations of the Multiple Streams Framework and its assumption of independent streams and the subnational studies primarily based outside of Washington D.C. that helped inspire the Advocacy Coalitions Framework and its definition of policy subsystems. Our point is not to use the creation stories of the theories as a reason to block their development or use but rather to continue to explore how these baseline characteristics of each theory continue to play out in their ongoing comparative applications.

The building blocks of the structure of any theory and its ontology are its concepts. For any given theory, the concepts represent phenomena of importance and, to various extents, what exists. Thus, learning a theory entails learning its concepts and the language therein. As mentioned, these concept-based languages affect what researchers see in the conduct of empirical inquiry (i.e., it forms the theory's lens), and updating and changing these concept-based languages signifies one-way theories grow and change. We sometimes measure these concepts directly and interact them to describe, explain, or make sense of policy processes. Other times, we do not measure concepts directly while using them to help explain something unobserved about the policy processes to make inferences. For example, Punctuated Equilibrium Theory does not measure its core concept of "attention" directly in cognitive fabric but instead measures topic frequencies and then infers shifts in attention through changes in these frequencies over time. In contrast, approaches, such as the Narrative Policy Framework, the Advocacy Coalition Framework, and the Ecology of Games Framework, focus on measuring core concepts, such as narratives, coalitions, and games, without much inference to other concepts. All the theories in this volume can improve by clarifying the function of the concepts in their ontologies: what concepts are

observed and measured directly, what concepts are unobserved and measured indirectly, and what concepts are used to describe or explain.

The ontology of a theory also involves structures and layering. Most theories incorporate aspects of Lakatos' (1970) hardcore (stable theoretical assumptions) and protective belt (adaptable hypotheses). Ideally, while refuting and confirming hypotheses in the protective belt, a research program separates progressive adjustments that add theoretical descriptive and explanatory power from degenerative adjustments that rationalize counterevidence and do the opposite. Similarly, some theories organize themselves based on the "framework–theory distinction" (Ostrom, 2005). Unlike the generic term "theories" used thus far, the framework-theory distinction renders frameworks and theories specific meanings. A framework provides the foundation for a research program via general assumptions, working concepts and vocabulary, fundamental conceptual interactions, and the scope of relevant questions and problems (similar to Lakatos' hardcore). Theories provide more refined concept interactions, often in hypotheses for particular settings (similar to Lakatos' protective belt). This distinction strengthens the portability of frameworks and the adaptability of theories, as found in the Advocacy Coalition Framework and Institutional Analysis and Development Framework. A framework's generality allows researchers to explore its concepts differently, incorporate new concepts, and investigate novel concept interactions, often using inductive or even a modified approach to grounded theory.

Our theories continue to grow and adapt in part from empirical research. We update them by adding new concepts and sometimes relate those concepts to new hypotheses. As discussed, the epistemological justification for contributing to our knowledge from our research has not been well specified or transparent. The same concern also lies in the continued growth of the ontologies associated with our theories. We see revisions of theories based on empirical evidence drawn from multiple cases, contexts, sources of data and forms of analysis, and researchers. However, we also see revisions to theories based on essentially no empirical evidence but instead based on reason, argumentation, and sometimes hunches. This is a serious issue as we often confuse our theoretical ontologies with our knowledge. While we can support various ways to revise our theories, we should practice transparency in how we do it.

4   *Diversity in theory-based research programs.*

People create, maintain, and advance policy process theories or destroy them through neglect and poor judgment. They define what it means to do science in this field and what counts as knowledge. This echoes notions of science as a "social enterprise" (King et al., 1994, p. 9) and the importance of people in our research programs (Laudan, 1978). It relates to ongoing challenges in how we engage the scientific process, such as questions about replication (Bird, 2020), pathologies of groupthink in the peer-review

process (Resnik and Smith, 2020), and issues associated with scientific paradigms (Kuhn, 1962). It relates to ongoing issues of science and society from democratizing science (Nowotny, 2003) to the socio–political–cultural interdependence of scientific knowledge and, hence, the emergence of post-factual politics (Durnová, 2019).

The implication is that our theories are not just out there, neither unchanging nor independent. Along with revising them based on our research and the research of others, we interact with them in teaching, research, and discussions with academic and nonacademic audiences. We, thus, bring to the theories personal ontologies based on our lived experiences, values, identities, skills, proficiencies, cultures, and more.

Thus, theories act not as the sole research guide but also in conjunction with us, as individuals and groups. We see this in students who use their lived experiences to question the integrity of theories. Consequently, through their dissatisfaction and skepticism, students often conduct their research to explore existing concepts in new ways or bring new concepts from their lived experiences. In other words, our engagement with theories becomes one of the reasons for their growth, and our absence becomes one of the reasons for their demise. The point is not to forgo techniques for achieving reliability (e.g., inter-coder reliability) but to recognize the benefits of the diversity of the people in any research program to continue the growth and development in our theoretical ontologies (as well as our epistemologies, methodologies, and methods). Equally important is incorporating people outside our research programs through practicing engaged scholarship, breaking the barriers between theories, or working in interdisciplinary teams.

How Can We Better Our Science?

Given the principles above, the following list serves as some of the ways to improve our science.

1 *Develop and promote best and established practices in methods.* The raison d'être of this volume is to provide an opportunity for the leading scholars of the most established theories of the policy process to state their best and established practices in methods and – these practices yet exist – to encourage them to propose what they might be. The underlying rationale is simple: through clearly articulated methods and associated methodologies, we might provide a baseline of what is known today in hopes of doing better in the future.

2 *Develop transparent epistemological norms to justify knowledge from our research and to update our theories and methods.* Plenty of research is being conducted under the theoretical umbrellas found in this volume and published in peer-reviewed journals and books. Obscurity remains in what of this research becomes noticed or unnoticed, valued or devalued, and, ultimately, legitimate or illegitimate in contributing to knowledge and a part of updating theories and methods.

3   *Clean and clarify our theories' ontologies, particularly the functions of concepts, the sources and reasons for including concepts, and the rationale for their interactions, by reason, empirical evidence, or hunches.* Our theories have been updated and revised over time. As mentioned, these changes occur for various reasons, including responses to empirical evidence, responses to criticisms and ambiguities, hunches, and values of priorities of the theoretical leaders. The problem is that we now have theories that can make strong claims about the world based on solid empirical grounds and claims without such foundations.

4   *Engage and leverage the inherent tensions in pluralistic ontologies, epistemologies, methodologies, and methods inherent in policy process research.* This chapter characterizes policy process research as practicing pluralism in its methodologies and methods and its ontologies and epistemologies. Such assumptions provide a foundation for interdisciplinary scholarship and, thus, aptly offer a useful way to think about the foundations of our science. It creates tensions, a need for tolerance, and a source for our debates, which we encourage and welcome.

5   *Keep developing diversity in the research programs.* The engine of growth is the diversity of people in the theoretical research programs. It recognizes the varied ways of doing and knowing and how life experiences contribute to the types of questions we ask and how we answer them. Just as important, it contributes to a sense of community and belonging that encourages long-term commitment and engagement.

6   *Maintain transparency.* The bedrock of science remains transparency that coincides with an openness to scrutiny, a willingness to share rights and wrongs, and a desire to learn from each other.

Bring Ourselves into Our Research and Our Art into Our Science

American Chef Thomas Keller once quipped: "*A recipe has no soul. You, as the cook, must bring soul to the recipe.*" Keller's choice of the word "soul" is essential. It refers to the energy and personal touch that the best chefs bring to their recipes and, hence, to the meals they serve. Chefs who put aside themselves will lack a noticeable quality and distinctness in what they make.

Like the soulless chef, we often teach our students the scientific method in a way that emphasizes a restraint of themselves for fear of showing their personal biases. This leads people to apply theories dispassionately, sometimes becoming mechanical and robotic. They look to a theory's ontology for thinking about the world and its methodologies and methods for studying it. However, in doing so, they also lose the energy and personal touch that makes them human and the source of their imagination and creativity. Theories then become the "shackles" that inhibit what they observe in the world, what they bring to science, what advancements they achieve in contributing to knowledge, and what fulfillment they experience in their careers.

We mention students in the previous paragraph, but the logic also applies to new and established scholars. Take anyone and give them all the knowledge

embedded in our theories and methods – without their imagination, ideas, and creativity, they might produce, but their contributions to science will likely be stale and trifling. Indeed, one common trait of our mentors and people who maintain long productive careers is their tendency to view their science as their art.

We should not forgo the scientific method. The arguments about supporting or suppressing reflexivity are not a binary decision. It instead flows from how and in what ways we bring ourselves into our research. Yes, we can use our creativity and maintain transparency. Yes, we can bring personal experience and expertise into a theory's ontology, epistemology, and methodology while making valid and reliable contributions. The norm should be disclosure and recognition of what we, as people, inevitably bring to our science as individuals and collectives.

At the heart of bringing ourselves into our research lies the spirit of embracing diversity. It means to support the cultivation of our community and each of us. If we assume otherwise, then diversity in our research programs would be meaningless. Instead, we can practice our scientific method while supporting the diversity of expertise and experiences we bring to our research and how it complements our data collection, analysis of data, and interpretation of data. The benefits are nontrivial in ultimately more inclusive forms of knowledge.

Notes

1 We use "theories" generically to represent any research approach to focus the scope of inquiry, help specify assumptions, and define and relate concepts. This includes the possibility of establishing various relational forms, such as hypotheses or propositions. It, thus, contains the terms frameworks, theories, and models unless specified otherwise. See Weible (2018, p. 1) for elaborations.
2 See www.comparativeagendas.net.
3 This question and the ones in the other principles parallel those stated by Hay (2002).
4 Similarly, Paul Sabatier summarized good science as including public methods, clearly defined concepts with empirically falsifiable propositions that are as general as possible, and the methods and concepts evaluated with skepticism (Sabatier, 1999b, p. 5).
5 As far as we know, despite goals of being clear enough to be proven wrong rooted in the principle of falsification, not a single theory or hypothesis in this volume has been falsified. We can ponder the reasons, including many unfalsifiable hypotheses, fuzzy concepts, lack of risk-taking, the peer-review process and the biases against null findings, and more.

References

Bird, Alexander. 2020. "Understanding the Replication Crisis as a Base Rate Fallacy." *The British Journal for the Philosophy of Science*. Advance online publication. https:// doi .org/10.1093/bjps/axy051

Blaikie, Norman. 2007. *Approaches to Social Enquiry: Advancing Knowledge*. Cambridge: Polity Press.

Durnová, Anna. 2019. *Understanding Emotions in Post-Factual Politics: Negotiating Truth*. Cheltenham: Edward Elgar Publishing.

Flyvbjerg, Bent. 2001. *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*. Cambridge: Cambridge University Press.

Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. Cambridge: Cambridge University Press.

Gerring, John. 2006. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Gettier, Edmund. 1963. "Is Knowledge Justified True Belief?" *Analysis 23*(6): 121–123.

Hawkesworth, Mary. 1992. "Epistemology and Policy Analysis." In William N. Dunn and Rita Mae Kelly (eds) *Advances in Policy Studies since 1950*. New Brunswick, NJ: Transaction Publishers (pp. 295–329).

Hay, Colin. 2002. *Political Analysis*. Basingstoke: Palgrave.

Henry, Adam Douglas, Karin Ingold, Christopher M. Weible, and Daniel Nohrstedt. 2021. "Advocacy Coalition Framework: Advice on Applications and Methods." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First Edition. Milton Park: Routledge.

Herweg, Nicole, Nicholaos Zahariadis, and Reimut Zohlnhofer. 2021. "How to Conduct a Multiple Streams Study." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First Edition. Milton Park: Routledge.

Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago, IL: University of Chicago Press.

Jones, Michael D., Mark K. McBeth, Elizabeth A. Shanahan, Aaron Smith–Walter, and Geoboo Song. 2021. "Conducting Narrative Policy Framework Research: From Theory to Methods." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First Edition. Milton Park: Routledge.

Karch, Andrew. 2021. "Innovation and Diffusion: Connecting Theory and Method." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First. Milton Park, UK: Routledge.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes." In Imre Lakatos and Alan Musgrave (eds) *Criticism and the Growth of Knowledge: Volume 4: Proceedings of the International Colloquium in the Philosophy of Science*. Cambridge: Cambridge University Press (pp. 170–196).

Laudan, Larry. 1978. *Progress and Its Problems*. Los Angeles, CA: University of California Press.

Little, Daniel. 2016. *New Directions in the Philosophy of Social Science*. Lanham, MD: Rowman & Littlefield.

Lubell, Mark, Matthew Hamilton, Jack Mewhirter, and Francesca Pia Vantaggiato. 2021. "Methodological Approaches to Studying Polycentric Governance with the Ecology of Games Framework." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First. Milton Park: Routledge.

Merton, Robert K. 1967. *On Theoretical Sociology: Five Essays, Old and New*. New York, NY: Free Press (No. HM51 M392).

Miller, Thaddeus R., Timothy D. Baird, Caitlin M. Littlefield, Gary Kofinas, F. Stuart Chapin III, and Charles L. Redman. 2008. "Epistemological Pluralism: Reorganizing Interdisciplinary Research." *Ecology and Society* 13(2): 46.

Nowotny, Helga. 2003. "Democratising Expertise and Socially Robust Knowledge." *Science and Public Policy 30*(3): 151–156.

Ostrom, Elinor. 2005. *Understanding Institutional Diversity.* Princeton, NJ: Princeton University Press.

Poteete, Amy R., Marco A. Janssen, and Elinor Ostrom. 2010. *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*. Princeton, NJ: Princeton University Press.

Resnik, David B., and Elise M. Smith. 2020. "Bias and Groupthink in Science's Peer-Review System." In David M. Allen and James W. Howell (eds) *Groupthink in Science*. Cham, Switzerland: Spring (pp. 99–113).

Sabatier, Paul A. 1999a. "Fostering the Development of Policy Theory." In Paul A. Sabatier (ed.) *Theories of the Policy Process*. Boulder, CO: Westview Press (pp. 262–275).

Sabatier, Paul A. 1999b. "The Need for Better Theories." In Paul A. Sabatier (ed.) *Theories of the Policy Process*. Boulder, CO: Westview Press (pp. 3–17).

Schlager, Edella, Saba Siddiki, and Michael Cox. 2021. "Methods for Analyzing Social Dilemmas and Institutional Arrangements within the Institutional Analysis and Development Framework." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First. Milton Park: Routledge.

Shanahan, Elizabeth A., Michael D. Jones, and Mark K. McBeth. 2018. "How to Conduct a Narrative Policy Framework Study." *The Social Science Journal 55*(3): 332–345.

Siddiki, Saba, Tanya Heikkila, Christopher M. Weible, Raul Pacheco-Vega, David Carter, Cali Curley, Aaron Deslatte, and Abby Bennett. 2019. "Institutional Analysis with the Institutional Grammar." *Policy Studies Journal*. https://doi.org/10.1111/psj.12361.

SoRelle, Mallory and Jamila Michener. 2021. "Methods for Applying Policy Feedback Theory." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First. Milton Park: Routledge.

Van de Ven, Andrew H. (2007). *Engaged Scholarship: A Guide for Organizational and Social Research*. Oxford: Oxford University Press on Demand.

Weible, Christopher. M. 2018. "Introduction: The Scope and Focus of Policy Process Research and Theory." In Christopher M. Weible and Paul A. Sabatier (eds) *Theories of the Policy Process.* Milton Park: Routledge (pp. 1–13).

Workman, Samuel, Frank R. Baumgartner, and Bryan D. Jones. 2021. "The Code and Craft of Punctuated Equilibrium." In Christopher M. Weible and Sam Workman (eds) *Methods of the Policy Process*, First. Milton Park: Routledge.

Index