1

Rethinking Comparison

An Introduction

Erica S. Simmons
University of Wisconsin–Madison

Nicholas Rush Smith
City University of New York–City College

Comparative methods have seen a double movement over the past two decades. On the one hand, political scientists point to the "enduring indispensability" of comparative methods, particularly controlled comparative methods (Slater and Ziblatt 2013), to explain political outcomes – an investment that has been deepened amid the increased influence of scholarship that relies on subnational comparison (Giraudy, Moncada, and Snyder 2019) or natural experiments (Dunning 2012) to improve causal inference. On the other hand, in the humanities and humanistic social sciences, comparative methods have come under intense scrutiny. Across these disciplines, particularly with the rise of postcolonial theory, such methods are frequently seen as "old fashioned at best, retrograde at worst" (Felski and Friedman 2013b, 1). Such critiques have gained particular traction because comparative methods were often used to compare "civilizations" through hierarchical, evolutionary social scientific paradigms and were tied to European colonial expansion (see also van der Veer 2016). Therefore, depending on one's vantage and disciplinary training, comparative methods might appear either as an indispensable tool for understanding the world or as intellectually and politically dubious.

Rethinking Comparison appears amid this bifurcated view of the comparative method, tackling some of the challenges raised by both perspectives. To do so, the book asks two fundamental methodological questions: (1) why do we compare what we compare and (2) how do the methodological assumptions we make about why and how we compare shape the knowledge we produce? In the process of addressing these questions, the chapters collectively set out comparative practices that diversify the repertoire of comparative methods available to students of politics while being cognizant of their history as tools to politically and economically dominate (see

Riofrancos, Chapter 6, this volume) or intellectually discipline (see Cheeseman, Chapter 4, this volume; and Wedeen, Chapter 13, this volume).

In pursuing these goals, the volume is intentionally aimed at a broad scholarly audience. Our hope is that the chapters offer tools for scholars of politics regardless of their epistemological or ontological assumptions, regions of interest, or scale of inquiry. Some of the contributors to this volume consider themselves "positivists"; others might use the label "interpretivists." Some compare across regions, others within countries, and others across time. Irrespective of these starting points, they all share a commitment to the importance of a pluralistic approach to comparison. Rooted in this pluralism, the chapters that follow simultaneously allow us to expand our understanding of the world even as we are cognizant that our comparative tools are themselves freighted with complex intellectual and political histories. This commitment is evident in the chapters' pragmatic goal of rethinking how comparison is practiced, their theoretical goal of rethinking why we practice comparison, their disciplinary goal of rethinking who is authorized to compare, and their political goal of rethinking the ends to which we compare.

COMPARISON AND POLITICAL SCIENCE

Our interest in engaging with questions of comparison emerges amid a revival of qualitative methods in the discipline of political science generally and a renewed interest in controlled or paired comparisons specifically.[1] Building on foundational work laying out the value of controlled comparison and the closely related strategies of paired comparisons, subnational comparisons, and natural experiments, this recent work shows how the method can combine the best of both the qualitative and the quantitative traditions.[2] Specifically, controlled comparisons allow scholars to trace out dynamic causal processes while accounting for or negating the effects of possible confounding explanations, ultimately enabling generalizable arguments.[3] Not surprisingly, controlled comparative approaches dominate current best practices in case

[1] On qualitative methods in political science generally, see Wedeen 2002; Mahoney and Rueschemeyer 2003; George and Bennett 2005; Gerring 2006; Schatz 2009; Brady and Collier 2010; Mahoney 2010; Ahmed and Sil 2012; Goertz and Mahoney 2012; and Ahram et al. 2018. On controlled or paired comparisons, see Snyder 2001; Tarrow 2010; Dunning 2012; Slater and Ziblatt 2013; Gisselquist 2014; and Giraudy, Moncada, and Snyder 2019.

[2] On controlled comparison, see Przeworski and Teune 1970; Lijphart 1971, 1975; Skocpol and Somers 1980; Brady and Collier 2010; and Slater and Ziblatt 2013. On paired comparison, see Tarrow 2010. On subnational comparisons, see Snyder 2001 and Giraudy, Moncada, and Snyder 2019. On natural experiments, see, e.g., Dunning 2012.

[3] This kind of comparison, often called "most similar with different outcomes" or "most different with similar outcomes," or the method of agreement and the method of difference, continues to reference Mill ([1843] 1882), although scholars often fail to acknowledge Mill's own discussion of the limitations of the approach (for an exception, see George and Bennett 2005). Regardless, what are often invoked as Mill's methods of difference and agreement are ubiquitous in

study political science research. Graduate students and professors alike look to select cases that hold potential alternative explanations constant or leverage variation in initial conditions or outcomes.[4] Indeed, in an article on the value of controlled comparisons, Slater and Ziblatt (2013, 1302) note the "enduring ubiquity" of the strategy in qualitative comparative research.

Contributions (and a Few Critiques) of Controlled Comparisons

Thus, even as qualitative comparative methods – and specifically controlled qualitative comparisons – have come under fire outside of political science, within the discipline, they have been central to some of the most enduring and influential scholarship.[5] Controlled comparisons drive canonical studies of phenomena as varied as the preconditions of social revolution (Skocpol 1979), the effects of social capital on state effectiveness (Putnam 1993), and party system stability and regime type (Collier and Collier 1991). Indeed, controlled comparison is such a dominant force in political science methods training that as two leading methods scholars note, "Nearly all graduate courses on comparative politics commence with a discussion of Mill's methods of 'difference' and 'agreement,'" which serves as the foundation for controlled comparative studies (Slater and Ziblatt 2013, 1302).

We agree that controlled comparisons have important utility for scholars engaging in small-n work. Contemporary scholars have effectively used controlled comparison to shed light on state capacity (Slater 2010), ethnic violence (Wilkinson 2006), and indigenous mobilization (Yashar 2005), just to name a few subjects. Yet, even as controlled comparisons have produced some of our most influential theories of politics, some scholars have noted their limitations. Those working within a quantitative epistemology have argued that research based on controlled comparison has limited ability to generalize (King, Keohane, and Verba 1994; Geddes 2003), a problem that scholars utilizing a mixture of quantitative and qualitative methods have tried to anticipate by implementing "nested" research designs (Lieberman 2005). Scholars working from various qualitative traditions, by contrast, have argued that projects deploying controlled comparisons tend to overemphasize their ability to address confounding explanations while necessarily underemphasizing processes of diffusion (Sewell 1985, 1996a) and interaction (Lieberson 1991, 1994). Still others suggest that controlled comparisons may unintentionally efface context by limiting the importance of people's lived experiences and the

qualitative comparative work (for a discussion, see Slater and Ziblatt 2013) and remain central to the ways in which we question and evaluate comparative case research.

[4] The approach to comparison and process tracing that George and Bennett (2005) lay out and the qualitative comparative analysis methods that Ragin (2014) pioneered are important exceptions here.

[5] Slater and Ziblatt (2013) make this point persuasively.

specific meanings they attach to political phenomena (Simmons and Smith 2017).

As an example of these challenges, take subnational comparisons, which scholars have argued are particularly well positioned for controlled comparative approaches because they allow scholars to hold so many potentially confounding variables constant (Snyder 2001; Giraudy, Moncada, and Snyder 2019). As Soifer (2019) carefully shows, however, serious problems emerge when we treat subnational units as independent because drawing appropriate subnational boundaries is challenging and the compound treatment problem plagues even the best subnational studies.[6] The consequence is that even with seemingly ideal subnational comparisons, as with any controlled comparison, it is hard to defend the claim that there is only one relevant difference between units and, therefore, difficult to establish causality in the way that scholars pursuing these comparisons often hope to do (Soifer 2019). Ultimately, the concern is that studies that rely on controlled comparisons may not be as predictive and testable as claimed (Burawoy 1989), pushing scholars to ignore research questions that do not immediately evidence variation that can be explained through logics of control (Ragin 2004, 128).

In this volume, Jason Seawright (Chapter 2) gives us a trenchant critique of the inferential capabilities of controlled comparisons. He does so by applying the comparative method to a class of individuals who would seem comparable and, therefore, about whom it should be easy to make general claims: billionaires. Within the United States, at least, billionaires (particularly politically conservative ones) would seem ideally comparable because they have a common political culture, overlapping social networks, and a shared elite status. Yet, despite these similar characteristics, as Seawright shows, attempts to make general claims about how conservative billionaires engage in politics quickly fall apart, as they have divergent political styles, they fund different kinds of organizations, and they often have varying concrete political goals despite the broadly shared "conservative" label. The problem of comparing billionaires raises a troubling question, though: if controlled comparative methods are of little help in understanding the political practices of such seemingly comparable individuals, how useful are controlled comparative methods for understanding political practices of more complex units of analysis like organizations, social movements, or states?

If political scientists are asking what the value of controlled comparisons is for generalizable causal inference, in the humanities and the humanistic social sciences, scholars have gone much further by questioning the value of comparison in the first place. Specifically, in the wake of postcolonial theory and amid recent demands to decolonize the academy, scholars have examined

[6] Giraudy, Moncada, and Snyder (2019, 36) define the compound treatment problem as emerging "when the treatment encompasses multiple explanatory factors, thus making it difficult to pinpoint which factor actually causes the effect."

the historical conditions under which comparative methods emerged and the political ends that they served (see Riofrancos, Chapter 6, this volume). Scholars writing in this tradition have traced the contemporary use of comparative methods to Europe's colonial encounters, seeing such methods as part of the evolutionary theories of civilizational "development" that helped justify colonial domination (see Cheah 1999, 3–4; van der Veer 2016, 1–2).

This dubious history, some scholars argue, means that comparison is always politically suspect. As Radhakrishnan (2013, 16) argues, "Comparisons are never neutral: they are inevitably tendentious, didactic, competitive, and prescriptive." Even further, for scholars in this vein, "Comparisons work only when the 'radical others' have been persuaded or downright coerced into abandoning their 'difference'" (Radhakrishnan 2013, 16). Indeed, for some, this need to create units that can be compared with other like units – say, a state or an ethnic group – makes comparison itself a violent process as the material and discursive conditions from which a unit is created *as* a unit are effaced (Cheah 2013, 178; see also Cheah 1999 and Cheeseman, Chapter 4, this volume). Worse, the material and political conditions under which these units are produced, scholars argue, are never equal (Spivak 2013, 253). So, even as most social scientists may consciously reject evolutionary theories of society and the colonial projects to which they were harnessed, critical scholars argue that this violent past is never fully dead; it continues to haunt present practices of comparison (see Riofrancos, Chapter 6, this volume).

Such haunting takes myriad forms: Parkinson's (Chapter 8, this volume) contribution to this volume offers an excellent example. Studying Palestinian refugee camps in Lebanon, Parkinson quickly found that the organizational dynamics she hoped to understand could not be easily disentangled into independent and dependent variables. Instead, she found that any understanding of the dynamics within the camps needed to be embedded in the historical context that had constructed camps as "camps" that could be compared to one another in the first place. The camps were hardly natural units, after all, given that they were products of violent political processes. Yet, the disciplinary training that most political scientists receive makes such units appear as natural and unproblematic, even as Parkinson discovered in her research that they were anything but.

Beyond Control

While it is important to recognize some of the limitations of controlled comparisons, our primary goal with this volume is *not* to critique controlled comparison (though critiques do inevitably emerge in some of the chapters). Instead, it is to both draw our attention to and better elaborate the logics behind some alternative ways of comparing. Even as controlled comparisons have produced lasting insights and continue to dominate research designs, they are not the only form of comparison that scholars utilize.

Scholars from virtually every subfield of political science have used forms of comparison that are not controlled to make central contributions to our understanding of politics, even as the logics behind these comparisons are rarely elaborated. In political theory, the tradition of systematic but uncontrolled comparison saturates the Western canon. Indeed, Aristotle's consideration of what was the best political regime was arguably the first uncontrolled comparative political science study (see, e.g., Tilly 2010, 8–10). Yet, if creative comparisons undergird the political theory canon, they also show up in both the classics and the cutting edges of modern political science. In American politics, such work ranges from Douglas McAdam's (1982) foundational study of African American political mobilization to Jamila Michener's (2019) recent work on inequality and civil law. In international relations, political sociologist Immanuel Wallerstein (1974) and his followers developed world systems theory to examine the uneven dynamics of capitalist accumulation, while Paul Amar (2013) has more recently studied how free-market economic policies have required novel security arrangements across the globe through a surprising comparison of such practices in Brazil and Egypt.

In comparative politics – the subfield most closely associated with comparative methods – the examples are legion. Foundational works by divergent scholars in the subdiscipline include Samuel Huntington's (1968) canonical study of political order, Benedict Anderson's (1983) classic text on nationalism, and Charles Tilly's (1990) agenda-setting study of state formation. These studies have been read by almost every graduate student at some point and continue to profoundly shape the discipline, despite the fact that they do not rely on controlled comparisons. More recent – yet still highly influential works – by Mahmood Mamdani (1996) on the logics of colonial states across Africa; Doug McAdam, Sidney Tarrow, and Charles Tilly (2001) on the dynamics of contentious politics; James Scott (1998) on high modernist state ideology; and Richard M. Locke and Kathy Thelen (1995) on labor politics – have all used modes of comparison that do not utilize controlled comparative logics. The same is true for some of the most recent work in the subdiscipline including Lisa Wedeen's (2019; see also the Epilogue, Chapter 13, this volume) study of Syrians' desire for autocracy amid the country's civil war, and our own work on social movements in Latin America (Simmons 2016) and vigilantism amid democratic state formation in South Africa (Smith 2019; see also Chapter 12, this volume).

Somehow, though, despite this legacy of classic studies and a profusion of recent scholarship that relies on non-controlled comparison to make claims about politics, surprisingly little methodological guidance is available to graduate students for how to design or execute comparisons that do not rely on control as a central element.[7] Worse, there is little epistemological insight on

[7] The works of Tilly (1984); Locke and Thelen (1995); Sewell (1996a, 1996b, 2005); Ragins (2004); Ahram et al. (2018); and Boswell, Corbett, and Rhodes (2019), which we discuss later in the chapter, are important exceptions here.

why such comparisons might be compelling in the first place. As a result, scholars often eschew comparative research designs that are not premised on controlled comparisons, or they shoehorn cases into controlled comparison frameworks that may not fit. And even when scholars do employ a non-controlled comparative approach, they rarely explain the utility of or logic behind the comparisons that they employ or how other scholars might perform similar comparisons. The consequences for our understanding of politics are severe. When we limit the kinds of comparisons we make, we necessarily constrain the kinds of questions we ask, limit the kinds of knowledge we produce, and foreclose our ability to imagine politics differently (Ragin 2004, 128).

Building the Foundations for an Expanded Comparative Method

Scholars within political science have most certainly developed valuable critiques of controlled comparisons, and those outside political science have raised important questions about the goals of comparison as a whole. Yet, social scientists have been less effective in laying out alternative approaches to comparison. The methodological logics behind the comparisons at the heart of their studies are rarely described, even as a wide range of approaches to comparison appears in some of the most influential work across the subfields of political science (see our earlier discussion).

Think, for example, of Benedict Anderson's (1990, but see also 2016) important work comparing ideas of power in Javanese and European political thought. Comparison between Java and Europe violates virtually every tenet of how a controlled comparison should be performed. Anderson writes across different scales (an island versus a continent), different regime types (a monarchy and subsequent dictatorships versus a wide variety of regimes), and different religious traditions (an Islamic system with animist elements versus largely Christian systems). Despite the lack of control, Anderson uses the friction between the conflicting concepts of power to illuminate how ideas inversely structure political practice in both settings. Had Anderson approached the comparison through the logic of control, he would not have been able to generate these insights. At the same time, it is not clear what the epistemology underlying these comparisons is or why they are persuasive or insightful. This explanatory gap leaves graduate students and faculty without the tools to explain why similar research designs will produce important insights.

To be sure, some core methods texts argue that tools like process tracing can "ameliorate the limitations" of the methods of agreement and difference and promise to expand the range of questions political scientists ask (George and Bennett 2005, 214–15). But even in these cases, process tracing is seen as a second-best approach when controlled comparisons are not possible (214–15), given that process tracing often appears as a critical *component* of controlled comparative research designs as opposed to an *alternative* to them

(e.g., Luebbert 1991; Htun 2003; Slater 2010).[8] Indeed, absent a controlled comparison, process tracing runs the risk of being labelled as "mere" descriptive inference because it would be unclear how generalizable the processes highlighted in a specific instance are – a critique that it would be hard for any political scientist to withstand, even though descriptive inference has a critical role to play in our explanations of politics (see Wendt 1998).[9]

Even if we were to value the kinds of explanations that emerge from process tracing or descriptive inference equally to causal logics that emerge from controlled comparisons, questions remain about how to pursue projects asking constitutive questions in a comparative fashion. Anderson's work on power in Java and Europe, for example, might be considered a project of descriptive inference in that he gives us two ways to think about concepts of power. Process tracing's focus on causal processes, though, is not appropriate for engaging in this kind of constitutive work given that Anderson is not describing a historical process. And clearly, controlled comparisons would not make sense to justify his project given the wide divergences across context we described earlier. But if the most commonly taught tools in graduate methods courses are not up to the task of helping us produce work like Anderson's, what other options are there?

Fortunately, in the past few years some political scientists have started to develop justifications for alternative modes of comparison. Boswell, Corbett, and Rhodes (2019, 36–39), for instance, argue that scholars might be able to expand the range of comparisons they make if they focus on the similar dilemmas people face across seemingly divergent contexts. As an example, they point to the ways in which both a pensioner in the Isle of Wight and a politician in Palau face major dilemmas in how to navigate the limited transportation options for carrying out their daily tasks (shopping in the case of the pensioner and campaigning in the case of the politician). Although distant in place and experience, Boswell and coauthors argue, comparing shared dilemmas creates opportunities for seeing political and conceptual connections across widely differing contexts that might not be immediately obvious.

Similarly, policy scholars Bartlett and Vavrus (2017) emphasize forms of comparison that look not only to similarities and differences but also to linkages, hierarchies of power, and questions of scale. They describe three, mutually imbricated axes of comparison – vertical, horizontal, and transversal

[8] See Slater and Ziblatt (2013, 1304) on this point as well.

[9] Descriptive inference looks different from causal inference insofar as it does not attempt to show that A caused B. Instead, it might explain how B came to be designed the way it is or the conditions that made B possible in the first place. This kind of inference allows us to explain phenomena in the world while not having to make the often-challenging assumptions required by the logics of causal inference (e.g., variable independence). Although they do not necessarily use the term, several of the chapters in this volume make the case that descriptive inference should play a central role in our analysis of politics (see Schaffer, Chapter 3, and Htun and Jensenius, Chapter 10).

comparisons – that situate comparisons not only spatially but also across scales and within historical contexts and relations (on transversal comparison, see also Kazanjian 2016).[10] Engaging in this type of comparison, though, means ridding ourselves of some preexisting ideas about how we think about our cases – to engage in an "unbounding" (Bartlett and Vavrus 2017, 13; see also Cheeseman, Chapter 4, this volume) of our cases so that we can see unexpected connections. Unbounding, Cheeseman argues, leads us to "follow resemblances across a political topography in pursuit of a problematic category or politically salient idea" (this volume, p. 66). Even as we locate our studies in a specific setting, we use that location as a vantage to help us explore the general. Attention to these axes also invites us to think differently about cases in ways that are echoed by Soss (Chapter 5, this volume) and Riofrancos (Chapter 6, this volume). When we focus on the "importance of examining *processes* of sense making as they develop over time, in distinct settings, in relation to systems of power and inequality and in increasingly interconnected conversations with actors who do not sit physically within the circle drawn around the traditional case" (Bartlett and Vavrus 2017, 10–11), we can both see our cases from multiple vantages, as Soss encourages us to do, and situate those cases in the global contexts that Riofrancos emphasizes in her contribution.

Contributors to a recent volume that revives comparative area studies offer a third path out of the binds controlled comparisons can create (Ahram, Köllner, and Sil 2018). Area studies have long been relegated to a second-class status relative to social science disciplines within the academy because of their supposed focus on "mere" description, rather than generalizable theories (see Cheeseman, Chapter 4, this volume). (Note, here, echoes of the critiques of process tracing and descriptive inference we raised earlier in this section.) Comparing across world regions, Köllner, Sil, and Ahram (2018, 4) suggest, is one means of solving this problem by combining the "thick" knowledge of an area specialist with the broad view of a social scientist. As with Locke and Thelen's (1995) work on contextualized comparisons, comparative area studies promise a middle path between the relatively narrow concerns of country experts and the maximizing goals of quantitative social scientists by helping case study researchers identify causal linkages that scholars in both of these other traditions might miss (Köllner et al. 2018, 5). To place area studies within a comparative framework, therefore, is in some sense infringing on disciplinary borders with the goal of balancing "a context-sensitive exploration of phenomena within individual cases with comparative analysis across cases from more than one area so as to develop portable inferences and illuminate

[10] A transversal comparison shows a connection between seemingly dissimilar places, time periods, or concepts. Coming from the Latin word *transvertere*, which combines the words for "across" and "to turn," Kazanjian (2016, 7) writes that the concept "means to turn across or athwart, to turn into something else, to turn about, or to overturn. A transverse is thus not simply a line that cuts across, but also an unruly action that undoes what is expected."

the convergence, divergence, or diffusion of practices across regions" (Köllner et al. 2018, 6).

Yet even with these recent efforts to rethink the building blocks of the comparative method, work that relies on non-controlled comparisons is often understood to be extremely risky. Because their logic is not broadly understood, these studies are often reserved for senior scholars with well-established reputations, are published in outlets not necessarily geared toward political scientists such as area studies journals, or are simply dismissed. Such work, therefore, comes with risks – risks that are not borne equally across the profession. For example, in his contribution to this volume, Soss (Chapter 5) writes about some of the risks he has taken in doing non-controlled comparative work throughout his career. Soss asks:

> Would the freedom to take this path – and the risks and eventual rewards involved in doing so – have been the same for a more socially and institutionally disadvantaged graduate student? Do scholars today have equal opportunities to take this path, regardless of gender, race, class, sexuality, tenure, or institutional prestige? I think the answer is clearly no. . . . The hope in this sort of writing is that we can broaden the ways people think about what is methodologically sound *so that it becomes less risky* to pursue alternative but equally valid ways of knowing and doing. (this volume, p. 101, italics in original)

Taking social location seriously means not only that political scientists need additional tools for designing and executing non-controlled comparisons; scholars also need those tools to become a central part of methods training. In doing this work, we would broaden who has the ability to do this nontraditional work by lowering its risks.

We see *Rethinking Comparison* as an effort to accomplish these goals by building on and expanding existing efforts in three crucial ways. First, *Rethinking Comparison* is deliberately designed to offer logics, tools, and insights for *all* scholars engaged in qualitative political science (and we hope social science generally), regardless of epistemological assumptions. Whereas Boswell et al. (2019) make a crucial contribution to our understanding of how comparison might be practiced by interpretive scholars, we aim to take their work one step further by bringing together scholars working from both positivist and interpretivist epistemologies to make the general case for the value of non-controlled approaches to comparison and, in so doing, bring into the conversation two epistemological communities that are often seen as distinct and irreconcilable. Second, while we are encouraged by efforts to bring comparative area studies back in, not all research projects lend themselves to that particular kind of focus and not all scholars are equipped with deep regional expertise (and we are not arguing that all research questions require this kind of expertise). So, having an expanded range of options for how to compare for a wide range of methods is desirable and, we hope, can help push the comparative area studies effort forward by broadening the range of