7

# Comparing Complex Cases Using Archival Research

Jonathan Obert
*Amherst College*

Social scientists have long considered the thorny question of how the large-scale phenomena we are frequently interested in explaining can both be produced by and shape behavior at the level of relatively local individual interaction. This question poses a particular problem for comparative analysis. Since we know that the kinds of phenomena we compare are often spatially or temporally related, how do we account for the seeming interdependency of much of social life as well as the fact that our comparisons often involve an implicit and untheorized assumption that those elements operate at the "same" level of analysis (Sewell 2005)? Treating cases as though they are independent and equivalent allows for controlled comparison and, hence, more rigorous causal inference and increased internal validity. However, this priority comes at the serious cost of providing a broad description of the multiple causes affecting an outcome (causal verisimilitude) as well as being unable to translate findings from research settings into the real world (ecological validity). Indeed, it is one of the great ironies of social science research that many of the phenomena we are most interested in understanding – the big, important events such as democratization, state formation, war and so on – are actually the most difficult to study using traditional tools of scientific research.

A major reason for this dilemma is that these types of social phenomena are *complex systems* – that is, they involve micro-behavior producing macro-level patterns that in turn shape micro-behavior (and so on). Social systems such as wars, nations, and institutions are constituted by activity at multiple scales of analysis and involve complex forms of interdependence among various actors, groups, and the sub- and supra-systems making them up. However, precisely because the cases we study are often interdependent at higher-order scales of analysis, traditional controlled comparisons have a difficult time actually identifying causal factors that might be responsible for explaining outcomes we care about. Thus, while on the one hand, we might wish to compare



https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

revolutions or wars to each other to help test our explanatory frameworks, we need to also consider the ways in which those events are tied into a larger, macro-historical system.

This chapter suggests one way to address this problem is to use historical archives to disaggregate the temporal and spatial properties of the phenomena we hope to compare while also tracing connections among those disaggregated elements. This allows us to identify the boundaries around subsystems that can be treated as relatively independent (thus allowing them to be compared directly), at the same time identifying the hierarchical connections tying those subsystemic activities together. Archives often present us with a complex assortment of documents, ephemera, images, and other materials, frequently compiled in unsystematic ways or drawn from multiple sources. By *classifying*, *contextualizing, layering,* and *linking* archival materials, analysts have an opportunity to think theoretically about how the phenomena they investigate operate at particular temporal and spatial scales. As a result, they can identify comparisons that consider the embedding of local and temporally truncated phenomena in large, slow-moving processes (and vice versa). Other methodologies can be leveraged to explore dependent and multi-level phenomena as well, but archival analysis allows the analyst to juxtapose different spatial and temporal scales simultaneously and to do so from a global rather than a situated perspective.

This chapter will proceed as follows. First, it will argue that social scientists should take seriously the multi-level and interdependent qualities of the historical phenomena we study, arguing that causal identification strategies (particularly controlled case analysis) only alleviate some of the problems these qualities present. Second, situating various research methodologies and their approaches to studying complexity in social life, it will argue that archival historical research offers scholars an opportunity to remedy some of the issues raised in this discussion. Finally, it will provide a short demonstration of how archival practice can enrich our explanations by analyzing the multi-level and interdependent comparative and causal claims in William Tuttle's (1970) important account of the 1919 Chicago race riot.

## THE USES AND LIMITS OF CONTROLLED COMPARISON

Social scientists have long known that adopting the scientific model of explanation to address social phenomena is extremely difficult in part because of how we typically link causal reasoning to comparative analysis. Specifically, much social science research attempts to identify whether or not a particular factor can be causally linked to an observable outcome. As a result, studies deploy comparison to isolate and identify whether that independent factor does, indeed, logically contribute to the consequence in question.

Increasingly, social scientific research explicitly addresses the inferential problems involved in identifying causal factors – that is, the logic by which

https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

causal claims are made and validated. Since the growth in popularity of the Neyman-Rubin counterfactual model in the past decades, scholars emphasize that the "gold standard" for causal explanation involves randomized controlled trials, in which the only factor systematically related to the outcome of interest among a given sample of comparable cases is an intervention made by the analyst (a treatment).[1] Randomizing who or what gets allocated to the control and treatment groups across a large and representative sample ensures that the effects of any drug, procedure, or test given by the analyst can be isolated from other systematic factors that might explain or influence the outcome (Rubin 1974). Of course, because it is morally or practically impossible to start a war or initiate a process of democratization to test a particular hypothesis, scholars instead turn to "quasi-experimental" techniques to simulate the causal environment of the controlled randomized trial. These design-based inferential strategies include so-called natural experiments, regression discontinuity designs, and the like. Indeed, even more traditional forms of social science research design such as comparative historical analysis and multivariate regression with controls try to help ensure *internal validity* – the notion that the factor in question really does bear a causal relationship to the outcome because of precise ability to control for confounding factors.

At the same time, even when using these techniques, we inevitably run into stumbling blocks. These problems largely arise for two reasons. First, because most social phenomena are deeply complex and multi-dimensional, comparing multiple cases involves invoking many assumptions about whether the objects we are comparing are actually of the same case. Second, because the causes for events inside and across complex systems are themselves nearly impossible to isolate effectively, it can be difficult to understand the overall importance of the causes we do identify in their actual empirical settings.

One problem is the issue of *interdependency*, in which an event or outcome affects subsequent or other spatially related events or outcomes. Interdependency might involve, for instance, complex causal links among various explanatory variables (such as a latent process causing two or more seemingly independent phenomena or feedback loops or confounds in the link between cause and effect), or it might involve a kind of reverse causality, in which a seeming effect (such as a rise in the stock market) actually shapes the hypothesized causal process (such as future expectations of market performance). Interdependency can also take the form of statistical endogeneity, in which some independent variable is correlated with the error term in a model specification, frequently due to omitted variables and unobserved heterogeneity, measurement error, or simultaneity among the dependent and independent variables. Without correcting for these biases,

[1] For an excellent description of the logic underlying experimental design in the social sciences, see Dunning (2012, 5–8).

https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

endogeneity makes it difficult for analysts to clearly state that a given factor is causally related to the outcome.

For instance, the problem of interdependence has long characterized the debate over whether or not democratic institutions cause economic development or whether or not development leads to the growth of democracy. Since many factors (such as state formation process or resource endowment) may shape both democratization and development, disentangling the direction of the causal arrow is difficult. Of course, social science scholars have made great strides in using, for instance, instrumental variables or natural experiments to try to identify variables related to the causal factor but not otherwise related to the outcome to tease out these relations (e.g., Acemoglu, Johnson, and Robinson 2001). But even in these accounts, interdependency can make it difficult to make causal arguments through comparison since there always may be unnoticed systematic factors tying those cases together, as well as intrinsic measurement problems (Morck and Yeung 2011). Since we insist on studying things that cannot be experimented on directly through randomization, like democracy and development, we end up leaving our work vulnerable to the charge that we are omitting something important.

Relatedly, interdependency can also pose a problem when a phenomenon of interest itself causes or has some effect on other cases with which we might like to compare it. For example, take the study of social revolutions. Scholars have long debated whether ideology, psychological factors, social structural factors, geopolitics, or micro-level collective mobilization was most responsible for causing epoch-transforming events such as the French, Chinese, or Russian Revolutions (Goldstone 2001). To explain social revolutions in general means to try to identify the factors present in the successful outcomes (actual revolutions) while also trying to distinguish them from factors that might be causally related to revolution but were not present in these cases (though they may have been present in other states that did not undergo revolution). Indeed, this is exactly the path laid out by Theda Skocpol (1979) in her important book, which leverages direct comparisons among France, China, and Russia (as well as the "negative" cases of England, Japan, and Germany) to make the case for a structural account of revolutionary change.

One problem with this procedure, as Skocpol (1979, 39) herself admits, is that the French, Chinese, and Russian Revolutions cannot really be treated as independent cases, in which causal factors exogenous to the events themselves are crucial in explaining their occurrence. Indeed, as Sewell (2005) suggests, as a transformative and novel process, the French Revolution was more than a case: it was an "event," one that fundamentally altered the world in such a way as to make future revolutions both thinkable (through changing ideologies of popular sovereignty and violence) and practically possible (by opening up a new era in nation-state development). In other words, it might be impossible to imagine the Russian and, consequently, the Chinese Revolutions without the French experience. Because these cases are *not* independent – at some level, the French

Revolution has a causal effect on all subsequent world history (even if subtle and imperceptible) – relying on straightforward controlled comparison can be misleading.

In their classic account of the logic of comparative social inquiry, Przeworksi and Teune (1970) argue that the fact that social phenomena are always part of a larger interdependent system is actually an advantage for making generalized theories. Since, in their view, social systems are relatively stable and bounded, it is possible to claim that "behavior of any component of a system is determined by factors intrinsic to the system and ... relatively isolated from influences outside the system" (12). This move, in turn, means that one can redescribe the properties of a given system (such as the "nation" of France) in terms of more general variables (such as *population* or *economy*) to compare it with other, similar systems (other nations).

The problem is that such interdependence is only contained within a system when we isolate particular systemic scales of analysis, such as nations. At a higher order, however, what happens in France in a given historical moment is also shaped by and shapes the behavior of actors in other comparable systems such as those of Russia and China. Instead of allowing us to treat critical social phenomena within a given historical context as isolated, such multiple hierarchical forms of interdependence call to mind the so-called butterfly effect identified by chaos theorists. The butterfly effect gets its name from the fact that because weather systems are so interdependent, a butterfly flapping its wings in some remote area can alter a local air current, setting in motion a chain of events leading to a hurricane many miles away (Hilborn 2004). In a real sense, the butterfly is a necessary condition or cause for the hurricane. But this cause is so enmeshed in a series of complex interactions that it becomes impossible to compare hurricanes to identify this cause specifically.

Of course, meteorologists would never try to explain a hurricane as solely a product of butterfly-induced wind turbulence. But without an explicit way of situating such small events in a larger complex system of physical interactions and relationships – that is, without taking scale into account of our models of weather dynamics – understanding the relative importance of butterflies in actual outcomes is impossible.

Indeed, such complexity creates a problem for the comparative study of many kinds of systems. In his brilliant critique of genetic determinism, Lewontin (2000, 95) notes the following:

> Causal claims are usually *ceteris paribus*, but in biology all other things are almost never equal. The natural differences in effects observed among organisms do not usually have sufficient regularity with respect to the natural variation of individual causes because these individually causally relevant variables are each too weak in their effects to dominate the large number of other variables.

Like organisms, real social phenomena are similarly produced through multiple causal pathways linking events in a chain of history. As a result,

empirical analysis, particularly of important moments or movements, should account for an irreducible level of interdependence among cases.

Similarly, interdependence among social phenomena means that the causes for large phenomena like revolutions involve, at some level, highly idiosyncratic and contingent phenomena at multiple levels of social life. Thus, Sewell (2005) goes on to show how the very particular historical conditions surrounding the storming of the Bastille in 1789 allowed a relatively small-scale insurgent action to scale up to full-fledged revolutionary crisis. In other words, not only are many of the kinds of phenomena we like to study interlinked, but they are also themselves complex aggregations of various scales of social action.

In part, this is a question of classification. Revolutions, for example, are distinguished from other kinds of violent rebellion like riots or *coups d'états* because of the profound effects they have on the distribution of authority within a nation-state and because of the scale of their popular mobilization (Tilly 2003). At the same time, revolutions are, at a more rudimentary level, constituted by the same kinds of local phenomena as coups and riots: street fights, pitched battles, and so forth. In this sense, revolutions are something more than the sum of their parts: their scale transforms all these small-scale actions into an event with ramifications and dimensions exceeding those of similar phenomena (Goldstone 2001).

Scientists typically use the concept of "emergence" to describe the ways in which micro-level behavior can create macro-patterns that often take on qualitatively different properties as they scale (Holland 2000). In his work on scaling processes, Geoffrey West (2017) has demonstrated how many characteristics of animals, cities, and organizations do not linearly increase with the size of the entity in question: for example, in large cities, the residents produce more patents, pollution, and wealth than we would expect if we simply extracted a trend line from a sample of smaller cities. Similarly, large animals live longer than we would expect simply by extrapolating the lifespans of smaller animals. These returns to scale, which West argues have to do with the advantages network forms of organization offer circulatory systems in animals and infrastructure in cities, reveal that larger entities are not merely aggregates of smaller sub-processes: as Durkheim (1982) revealed years ago in his description of "social facts," they possess qualitatively and quantitatively different properties.

Moreover, the macro-patterns produced by micro-level activity can also provide feedback on and shape that micro-level activity in an effectively exogenous or independent way. Cities and large organisms not only aggregate micro-processes differently; they also affect the evolution of those processes over time by creating new physical or social environments through which those processes operate. This notion – *coevolution* – is a fundamental attribute of many social phenomena as well (Padgett and Powell 2012). Thus, many scholars of institutions and organizations implicitly argue that such entities are worth studying precisely because they have a seemingly independent effect

on the way actors behave, even as it is the activity of those agents that makes institutions and organizations "real" to begin with.

Emergence and coevolution arise because many social phenomena are simultaneously micro and macro – that is, they are multi-level. Indeed, the multi-level quality of social processes has as serious an implication for comparative analysis as interdependence. Since the phenomena we are interested in studying cannot always simply be derived by a straightforward aggregation of individual behavior. In addition, since those entities are seen as worthy of comparison and explanation on their own terms, we must be careful to consider not only whether our comparisons are able to control for possible causal variation but also to identify the specific level at which those comparisons are being deployed. Indeed, unless we specify the level at which we are making our comparisons and attend for the ways multiple scales together constitute cases of some social phenomenon, even sophisticated causal identification strategies are limited in actually explaining outcomes of interest (Kocher and Monteiro 2016).

Again, take social revolutions. As emergent phenomena, revolutions do not merely aggregate collective violence; they also fundamentally change the nature of authority in a state. Moreover, as examples of coevolution, revolutions often concatenate from small-scale street battles or conflictual interaction (such as the storming of the Bastille) while cascading into a larger, coherent social project that informs the interests and desires of participants. Coevolution and emergence allow us to treat revolutions as coherent social entities (or systems) worthy of explanation.

At the same time, if we compare social revolutions to explain them, we must be careful to distinguish levels of comparison. Even when separated by decades, revolutions are interdependent, leading us to try to use techniques such as natural experiments or instrumental variables to tease out specific factors that might be absent in different cases from those that are present to control for them. But as multi-level phenomena – that is, simultaneously combining local interpersonal fights and rebellion, city- or region-wide patterns of violent conflict, and national-level alterations in the distribution of political authority – at what level are those specific factors actually operating? And what role do they then play in the causal model we propose?

We might, for instance, follow Sewell (2005) in noting that the particular historically contingent uprising at the Bastille was able to scale up to the level of a revolution because of local factors like where the event happened (Paris), how the king responded (by fleeing the city), and how the symbolism of the Bastille allowed it to become a focal point for otherwise disparate opposition to the Old Regime. These factors are historically local and unlikely to ever replay in their exact sequence in other revolutions, even as they had crucial effects on the overall social structure of France at the time. However, it is precisely these local and contingent factors that are also likely to be most independent of events happening elsewhere. That is, the storming of the Bastille *itself* is unlikely to

have any direct causal relationship to other local catalytic revolutionary activity in Russia or China, for instance, even as the French Revolution *as a whole* was, indeed, causal for later revolutions. In identifying comparisons, it is worth thinking about which parts of the entities we are comparing are actually comparable in a controlled setting and which are actually related to macro-level phenomena (like capitalism or modernity) linking those entities together within a larger system.

This point is very similar to that made by Lewontin (2000), as noted earlier. Like trying to trace every biological operation to genes, thereby ignoring the ways in which environment and organism coevolve, the problem with isolating a single social indicator and using it as a proxy for a whole social phenomenon is that single elements only reflect a small part of the emergent entity that is actually of interest. Without carefully considering how these elements both contribute to the whole and operate in relation to one another, however, simple controlled comparison cannot provide adequate leverage in accounting for causal differences.

In short, the multi-level (emergent and coevolutionary) and interdependent characteristics of much of social life make straightforward comparisons a difficult proposition for social scientists. While controlled comparisons can be helpful in shedding light on questions of causal inference and internal validity, these benefits often come at the cost of *causal verisimilitude* – the ability for a comparison to reveal a fuller range of true causal relationships at play – and *ecological validity* – the ability of a particular identified causal effect to also operate in the natural world outside a particular (often experimental) controlled comparison.[2]

## ADDRESSING INTERDEPENDENT AND MULTI-LEVEL COMPARISONS

Social scientists are, of course, very aware of the difficulties involved in controlled comparison and have designed many techniques to cope with them. Statistically minded scholars spend much time addressing the problem of keeping units of analysis consistent in observational analysis or modeling individual from group observation – the well-known ecological fallacy – while techniques like multi-level modeling help account for different groupings of data within the same analytic framework. Much formal rational choice modeling, in turn, focuses on identifying specific mechanisms operating at a particular unit of analysis as well as trying to trace the relationships between micro-foundations and macro-patterns (even as much of this research remains methodologically individualist in orientation). Network and spatial analyses take seriously the interdependence of observations, and multi-method research designs often combine micro-analysis of particular

[2] An important body of research in the comparative historical analytic tradition uses controlled comparisons not to establish causality but rather to help eliminate alternatives and provide a framework through which to identify mechanisms. See Hall (2003) for a review.

https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

cases with statistical analysis of cross-sectional or time series data. Comparative historical analysts and others use process tracing, sequence analysis, and qualitative forms of comparative inquiry like QCA to examine complex causal processes like equifinality (many possible distinct causes having a similar outcome) and conjunctural causation (effects requiring multiple causes). All of these efforts are attempts to address the problems of verisimilitude and ecological validity in comparative analysis.

At the same time, we lack a unified framework for understanding how more fundamental methodological choices might affect our ability to disaggregate the multi-level and dependent qualities of social life. In complex systems (such as, for instance, the human body), many of the sub-components (such as organs and cells) can be studied on their own terms, without paying much attention to the overall system in which they are embedded. Herbert Simon (1996, 171) calls these "nearly decomposable systems" in which "the details of components can often be ignored while studying their interactions in the whole systems," while, in turn, "the short-run behavior of the individual subsystems can often be described in detail while ignoring the (slower) interactions among subsystems." Since the elements of systems are simultaneously bounded into subsystems and arranged in hierarchies according to scale (e.g., multiple organs and their connections together make a larger entity known as the body), we must therefore identify which elements of a given case are actually decomposable into smaller scale units of analysis (and therefore comparable) and which are interdependent.

To return to the example of the butterfly effect, Holland (2000) notes that even though accurately describing every component of a dynamic, nonlinear system like weather is impossible, "because meteorologists do *not* know the value of all the relevant variables, they do not work at a level of detail, or over time spans, in which chaos would be relevant." Instead, he argues, "the key to deeper understanding, as with weather prediction, is to determine the level of detail and the relevant mechanisms" needed to accurately describe weather with a broad brush (44).

Using this as a guide, I suggest classifying methodological choices based on the breadth of perspective through which they identify and verify mechanisms as well as how well they are able to embed those mechanisms in multiple temporal and spatial scales. Table 7.1 presents these classifications. Such a typology is only a heuristic, meant to capture broad generalizations for how to think about comparative analytic choice; undoubtedly, hybrid, multi-method, and computational approaches bridge these classificatory boundaries. Nevertheless, this provides at least a first cut at understanding how the causal logics of our methodological sources address the problems of mechanism and scale.

Experimental designs, for example, are usually very good at comparing *situated* mechanisms. Within the ambit of the controlled environment, that is, an experimental design can identify how a causal factor operates, usually

TABLE 7.1 *Methodological choice and complex comparisons*

| | | Perspective on Mechanisms | |
|---|---|---|---|
| | | **Situated** | **Global** |
| Approach to Scale | Ordered | Experiments | Large-N Empirical Analysis |
| | Simultaneous | Formal Rational Choice Modeling; Comparative Statics | Agent-Based Models; Network Analysis; **Archival Analysis** |

at the level of the individual. At the same time, most experiments have a difficult time accounting for the multi-level nature of social phenomena, precisely because they are designed to isolate those mechanisms in particular contexts rather than link them across scales. In this sense, they provide an *ordered* approach to considering levels of analysis by only examining one scale at a time. In turn, most cross-sectional and time-series statistical modeling of observational data compares multiple mechanisms at once, including those involving groups, individuals, communities, and so forth. In this sense, rather than a situated perspective, such modeling involves global comparisons. Similar to experimental designs, though, even when it deploys hierarchical controls, observational statistical analysis is usually designed to produce correlative comparisons at a particular scale of analysis – dependent variables involve a particular level of analysis rather than chaining small-scale and large-scale phenomena, for example. It too is ordered.

Some approaches – such as much formal rational choice modeling – instead try to model the links between micro- and macro-level phenomena explicitly, with the goal of allowing the analyst to compare multi-scalar phenomena like tipping points, collective action, and so forth. In this sense, these approaches adopt a *simultaneous* analysis of scale. Also, much like experiments, with a few important exceptions (e.g., Bacharach 2006), many formal rational choice models are ultimately only interested in describing mechanisms emerging from individual, rational behavior. In this sense, one only compares mechanisms that are situated rather than thinking about how multiple sorts of actors and agents might coproduce social phenomena.

Scholars have also proposed techniques such as agent-based modeling or multiple network analysis designed to compare the operation of multiple mechanisms across cases and to juxtapose different scales of analysis directly. However, while there are important exceptions (e.g., Padgett and Powell 2012), these approaches are frequently used for exploratory and theory-building purposes, rather than for analyzing historical, empirical cases.

One of the oldest methodological choices – historical archival research – is surprisingly one of the best options when it comes to comparing actual empirical examples of complex cases usually because of structural factors involved in such work rather than the preferences of the researcher: archives frequently include a wide variety of materials, including letters, reports, notes, photographs, ephemera, and objects, frequently compiled in unsystematic ways. This very non-systematicity requires the analyst to situate and link those materials to interpret them.[3]

## THE ADVANTAGES OF ARCHIVES: A FRAMEWORK

I suggest that any user of historical primary material must perform at least four tasks that also serve as a conduit to taking a global perspective on comparing mechanisms and juxtaposing multiple scales of analysis simultaneously. This is not to say most users of historical archives actually engage in these activities. Indeed, one of the hallmarks of traditional historical scholarship is deep focus on a particular historical event, individual, or phenomenon rather than a comparative one. At the same time, comparative social scientists have much to learn, I argue, from using primary material in an analytic way.

The most basic task any archival research process requires is *classification*. Scholars almost always process and sort archival objects into groups sharing a given characteristic. Archives are often pre-classified by archivists or collectors into record groups or collections, albeit to varying degrees. Within a given repository, for example, collections might range from being well-inventoried selections of material from a clear provenance (e.g., government documents) to relatively scattershot collections of uncertain origin (e.g., personal papers). However, scholars almost always need to refit this classification to serve their own ends. This allows them to create what Zerubavel (1996) has termed an "island of meaning," a way of suppressing individual difference and forging uniformity in service of cognitive organization. When examining a series of letters, for example, a researcher might classify by date, topic, provenance, sender/recipient, or some combination of these to address some specific questions or identify some particular chain of events. In turn, scholars also typically classify the materials they review across collections or archives, grouping materials that are institutionally segmented.

Classifying is essential in any research process, but archival classifying provides some intriguing advantages for scholars examining multi-level and interdependent cases. Some archival material (such as letter collections) can be "high bandwidth," including information about specific events or people, as well as reflections on larger macro-level phenomena, like wars or elections. Scholars attempting to classify this material must think through how the

[3] For an insightful discussion of the pleasures and complexities of archival inquiry, see Farge (2013).

https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

multiple forms of social life captured in the archival document are situated at the level of the output they hope to produce (usually historical narrative).

Another, related task is *contextualizing* archival materials – that is, situating their production and reception in particular places and times. Contextualizing is also common in a variety of research designs; however, because of the frequently obscure processes by which many archival collections were gathered in the first place, as well as the frequently scattershot way in which the materials within those collections are organized, archival researchers inevitably focus much attention on the task of identifying the provenance for particular documents, people, allusions, and events found in those materials.

Like classifying, contextualizing archival materials is an important way to examine complex social phenomena. By trying to situate often cryptic materials (such as scraps of ephemera, undated account books, etc.) in a time and place, and to interpret what a particular document was and meant, scholars engage in a form of theorizing about the entity producing the text; how the document might have circulated; and, most crucially, why it was preserved and others, perhaps, were not. This, in turn, draws scholars to identify that different archival objects possess their own scales of analysis – personal journals or family documents shed light on different kinds of mechanisms and agents than do published reports or reprinted photographs.

Scholars working with historical archives also frequently engage in another task: *layering*. To produce a coherent account of a given social phenomenon, researchers often see the documents and texts they find in the archives as pieces (or layers) added onto an underlying structure of existing historical scholarship. These pieces might revise or alter that existing scholarship, but archival work always involves a kind of dialogue between secondary interpretations and primary materials. Not only do scholars use existing works as roadmaps to archival collections, they also try to make novel contributions by examining overlooked or unused sources. Moreover, since layering involves aggregating, disaggregating, or uncovering archival evidence in relation to other texts, it also entails the researcher in a process of rethinking the overall phenomenon of interest in light of specific evidentiary discoveries. With each new archival finding – a letter containing an unknown detail or a pattern of events revealed by a statistical summary of previously scattered data – historical researchers inevitably identify the ways these new archival layers alter the conceptualization and measurement of the case they study.

A final task for archival researchers is *linking* materials together to forge an actual historical analysis. Because archival collections must be organized and interpreted to become part of the scholarly conversation, researchers frequently find themselves connecting pieces of evidence from different contexts to produce an evidentiary basis for their arguments. Linking evidence (whether or not archival) is particularly crucial in examining the multi-level and interdependent nature of historical cases. Gaining purchase on the ways in which multiple mechanisms operate together means drawing connections

across different cases, reducing social complexity, and identifying multiple resolutions through which to analyze the evidence. That is, archival work requires embedding evidence from a variety of materials together in often-uneasy ways; massive government reports make different claims than those of newspapers because the conditions of their production differ. Analysts trying to link these materials often find themselves zooming between big-picture and close-up analyses, allowing for multiple perspectives on the scales of social action involved in the making of a given occurrence or fact.

Classifying, contextualizing, layering, and linking are not merely techniques for processing archival materials. They also are essential in comparing those materials across cases and contexts to gain theoretical insights. Classifying archival material also means situating particular islands of meaning in a larger ocean of historical evidence; contextualizing involves thinking about seemingly unrelated micro- and macro-processes that might condition a particular case (or, conversely, be absent from that case); layering involves tying a particular case to a larger edifice of scholarly interpretations of similar and different cases; and linking involves direct comparisons and reductions among pieces of evidence and, potentially, different cases to create an actual analysis.

In this sense, even when it putatively deals with a single case, archival work requires scholars to think hard about how the complex social phenomena they study are defined and situated in relation to other historical phenomena. In the words of Ann Laura Stoler (2002), archives are not merely "objects" containing materials to be accessed, they are also "subjects," capable of revealing much about how the producers of these materials thought about and organized their conceptual worlds. Thus, for precisely this reason, archival inquiry can be helpful in identifying a global set of mechanisms and thinking about multiple scales of analysis simultaneously. Engaging the archive as a subject helps us disaggregate and recombine spatial and temporal dimensions of complex social phenomena so as to identify more clearly how comparisons across cases might reveal the relationships among them. It also provides us a way to model complex social cases by giving us an opportunity to think about which elements of cases can be treated as bounded subsystems and the ways in which those bounded subsystems relate to one another hierarchically. I now turn to an examination of William Tuttle's sophisticated account of the Chicago Race Riot of 1919 to demonstrate how archival research can help scholars account for multi-level and interdependent phenomena in a comparative framework.

## ARCHIVAL ANALYSIS AND THE CHICAGO RACE RIOT OF 1919

The Chicago Race Riot of 1919 was a key moment in the history of American racial violence. Angered by a fatal attack by a white bystander on a Black boy swimming in a contested zone of a beach in the highly segregated city on the

https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

afternoon of July 27, a group of African Americans demanded that a white police officer arrest the man they believed responsible for the attack. Failing to comply with this demand, the white officer instead arrested an African American bystander, leading to rumors about police complicity in the death of the swimmer. This, in turn, sparked a series of conflicts between crowds of African Americans, whites, and police officers, cascading into a general series of brawls and fights. As Tuttle (1970, 8–10) describes it: "Once ignited on July 27, the rioting raged virtually . . . uncontrolled for the greater part of five days. Day and night white toughs assaulted blacks, and teenage black mobsters beat white peddlers and merchants in the black belt. As rumors of atrocities circulated throughout the city, members of both races craved violence." By the time the rioting was finally quelled, at least thirty-eight people were dead and hundreds were injured, one of the bloodiest episodes in a summer that witnessed tremendous levels of bloodshed.

Defining "riots" (like their close cousins "revolutions") is fraught with difficulty. In his typology of violent interaction, Tilly (2003, 18) eschews the term completely, arguing that it necessarily "embodies a political judgment" and is rarely used as a term of self-identification. Yet it is precisely because riots involve such political judgments that they take on a "life of their own," capable of shaping other social phenomena (including other riots). Riots are politically integrated and consequential events, treated by participants and audiences as different from "mere" brawls, fights, or lynchings. As a result, they are multi-level and interdependent phenomena. In his work on ethnic rioting, which tries to set out a more encompassing definition, Donald Horowitz (2001, 17–28, 522–24) thus stresses the *process* of rioting, which is intense, intergroup, violence concentrated in time and space. Focusing on this process means that race riots aggregate micro-level violence differently from other macro-level protests and genocides, with the goal of purposively organizing conflict to preserve or unsettle racial social boundaries.

Thus, the Chicago Race Riot itself was made up of many individual acts, but those acts were parts of an integrated spatial and temporal pattern. In his account, Tuttle stresses how the initial impulse of violence diffused outward from the original, concentrated points of contact near the "dead line" (the street separating the highly segregated African American Black belt from white neighborhoods in the south side of the city) to more isolated individual attacks of African Americans and whites in areas controlled by the "other side" in the following days (see Figure 7.1). In particular, Tuttle (1970, 42–43) notes, this was related to police tactics: as the police concentrated in the Black belt, conflict moved into the white neighborhoods surrounding the stockyards. Meanwhile the rumor mill worked overtime, spreading misinformation about the scope and scale of the violence both in the press and among neighbors, friends, and family members. Eventually, the mayor asked the governor (with whom he had an intense political rivalry) for the aid

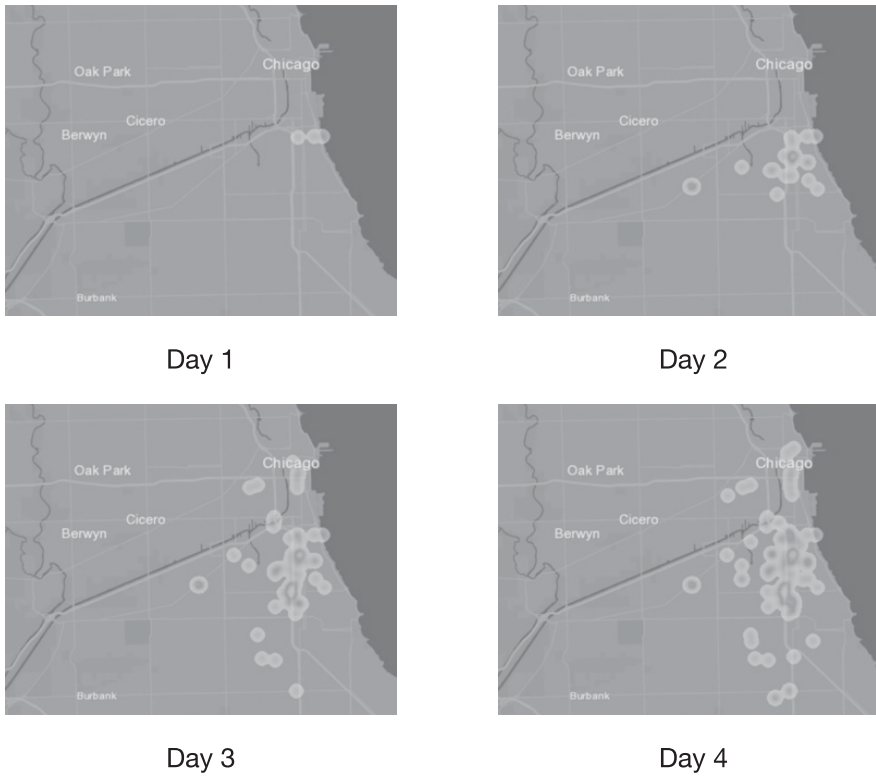


FIGURE 7.1 Diffusion of violence in Chicago Race Riot (July 26–29, 1919)

of the militia. Combined with a couple of rainy days, which forced many of the riot participants indoors, the state guard helped bring an end to the violence.

Tuttle's ability to treat the riot as both an integrated event and as a series of street-level interactions at least partially flowed from the close attention he paid to multiple kinds of archival sources, including interviews, newspaper articles, photographs, military reports, and (most importantly) the report of the Chicago Commission on Race Relations released in the aftermath of the event – *The Negro in Chicago* (1922). In a revealing discussion on his use of sources, Tuttle (1970, 269–89) provides an exhaustive description of how he organized his collections within themes (e.g., labor relations, politics, the Great Migration) as well as by type of resource (e.g., archives, newspapers, bibliographies). This dual scheme allows the reader to reconstruct how Tuttle organized his book (with distinct chapters dedicated to labor, politics, etc.) as well as how he used different sources to different ends (drawing largely on secondary sources for background or context-setting sections of the book and using primary documents for his own reconstruction of the riot itself).

The materials Tuttle used thus explicitly prompted him to *classify* his material by constructing a time line of events involving multiple groups of actors – street gangs, laborers, police officers and officials, and military commanders – acting across several levels of organization. At the same time, he built a narrative *linking* those events into a shared story of a tragic, albeit relatively routine act of violence (an assault upon an African American by a white man defending his "turf") scaling outward through processes of rumor and social mobilization to involve a large section of the city.

The Chicago Riot is thus a clear example of the kinds of complex cases on which this chapter has focused. On the one hand, it (like all social entities) was an aggregation of micro-behavior; on the other hand, the use of the label "riot" by the press and politicians as a shorthand for all the individual acts of violence and the response of the city government to this violence as a large-scale, integrated phenomenon (by, for instance, mobilizing the police and later the militia) directly shaped those very micro-behaviors (Brass 1997). This, in turn, complicates our ability to compare fights or beatings within the riot in any kind of controlled way, since many of those events were directly influenced by the emergence of the multi-level phenomenon of the riot as an integrated social phenomenon.

Tuttle, however, was not only interested in explaining the micro-level of violence in Chicago itself. His archival work also led him to situate the riot at two other levels of analysis. First was a meso-level, national cross-sectional network of towns and cities. As Tuttle points out, the Chicago riot was only one in a large number of racially motivated episodes of violence in 1919: between April and November that year, at least eleven riots took place across the nation in which at least two people were killed for racial reasons, in addition to a large number of other racially motivated violent events (see Table 7.2). In some cities – in particular, Charleston and Washington, DC – the presence of recently demobilized sailors and soldiers (some of whom were African American) meant that local interracial fights quickly scaled into larger confrontations between organized units. In other areas – such as Longview, Texas, and Omaha, Nebraska – riots accompanied vigilante mobilization and attempts at lynching alleged African American criminals. Still other events – in particular the Elaine, Arkansas, and Bogalusa, Louisiana, riots – concerned economic and labor competition.

For Tuttle (1970, 10–14), these events were part of a pattern of violent resistance to the presence of African Americans in those towns and cities most affected by the Great Migration, which brought many Black residents north for the first time. Most significant was the newfound competition over jobs, particularly in midwestern industrial centers and mid-Atlantic cities. African Americans were often hired as temporary scabs or as a low-cost alternative to union labor, leading to resentment on the part of striking workers and presenting a major challenge to unionization efforts. In addition, the movement of African Americans into cities frequently led to shifts in

TABLE 7.2 *Racially motivated violence in the United States (1919)*

| Town/City | State | Date | Type of Violence |
|---|---|---|---|
| Millen | GA | April 13 | Riot |
| Warrenton | GA | May 1 | Lynching |
| Philadelphia | PA | May 9 | Riot |
| Charleston | SC | May 11 | Riot |
| San Francisco | CA | May 14 | Riot |
| Vicksburg | MS | May 14 | Lynching |
| Milan | GA | May 24 | Riot |
| New London | CT | May 29 | Riot |
| Ellisville | MS | June 26 | Lynching |
| Annapolis | MD | June 28 | Riot |
| San Francisco | CA | June 30–July 1 | Riot |
| Bisbee | AZ | July 4 | Riot |
| Dublin | GA | July 6 | Lynching |
| Longview | TX | July 11 | Riot |
| Port Arthur | TX | July 14 | Riot |
| New York | NY | July 19 | Riot |
| Washington | DC | July 19–July 23 | Riot |
| Norfolk | VA | July 21 | Riot |
| Newberry | SC | July 26 | Lynching |
| Chicago | IL | July 27–July 31 | Riot |
| Syracuse | NY | July 31 | Riot |
| Lexington | NE | August 5 | Riot |
| Mulberry | FL | August 18 | Riot |
| Cadwell/Ocmulgee | GA | August 26–August 27 | Riot |
| Bogalusa | LA | August 30–August 31 | Lynching |
| Knoxville | TN | August 30 | Riot |
| Jacksonville | FL | September 7 | Lynching |
| New York | NY | September 21 | Riot |
| Omaha | NE | September 28–September 29 | Riot |
| Montgomery | AL | September 30 | Lynching |
| Elaine | AR | September 30–October 1 | Riot |
| Baltimore | MD | October 2 | Riot |
| Gary | IN | October 4 | Riot |
| Monticello | GA | October 7 | Lynching |
| Donora | PA | October 9 | Riot |
| Hubbard | OH | October 10 | Riot |

(*continued*)

https://doi.org/10.1017/9781108966009.007 Published online by Cambridge University Press

TABLE 7.2 (*continued*)

| Town/City | State | Date | Type of Violence |
|---|---|---|---|
| Corbin | KY | October 30 | Riot |
| Macon | GA | November 2 | Lynching |
| Wilmington | DE | November 13 | Riot |
| Bogalusa | LA | November 22 | Riot |
| Lake City | FL | November 29 | Lynching |

housing values: many whites did not want to live near African Americans or feared the effects of African American neighbors on the values of their properties. Speculators took advantage of this fear to drive prices down, buy properties at a discount, and then rent to African Americans at exorbitant rates. In other areas, white property owners formed housing associations to prevent African Americans from moving into their neighborhoods. In Tuttle's reading, understanding the riot also means situating it in a regional story of interdependent economic and demographic change. Without these meso-level factors, which explain how racial animosity could become so widespread in Chicago to begin with, it is impossible to make sense of how a routine violent interaction could scale to the level of a citywide riot.

How did Tuttle identify this meso level? Partly this was due to the availability of two particular archival resources: newspaper reports and, especially, magazine articles, both of which prompted him to *contextualize* Chicago's experience within the nation and compare it with other areas experiencing racial violence. Newspaper commentators routinely asserted that the events of 1919 were part of an integrated phenomenon of racial conflict (New York Times 1919), while periodicals produced by the National Association for the Advancement of Colored People (in particular, *Crisis* magazine) catalogued and interpreted multiple events across the nation during that year. Tuttle *layered* these findings on an interpretive framework connected to scholarship on the Great Migration, which helped him interpret the local violence as embedded in a larger, structural change in American social life.

Tuttle also, however, situates the Chicago riot at a third level of analysis: a macro-level story about how the global revolutionary movement and social transformations induced by wartime unsettled the racial order in the United States and led to a mass panic (commonly called the "Red Scare") over the threat of radicalism and social disorder. The Red Scare, in turn, "was an extension of the atmosphere of the war, with its cult of patriotism, its generalized climate of violence, and its need for an enemy" (Tuttle 1970, 17).

In other words, it was not simply that street-level violent events were situated in cities and towns undergoing economic and demographic change; those changes were seen as particularly threatening because of the historical conjunctural moment confronting the United States in 1919. The Bolshevik Revolution in Russia, nationalist and revolutionary struggles throughout Eastern Europe, growing disorder and disintegration of the social fabric in Germany and in Western Europe in general, and the presence of a newly militarized American national identity all made violence appear as a much more palatable solution to the problem of localized disorder in Tuttle's (1970, 15–16) view. As African Americans grew increasingly assertive in their willingness to defend themselves and their communities (a phenomenon typically called the growth of the "New Negro" sentiment, emblematic of the increasingly popular resistance rhetoric deployed by figures like W. E. B. DuBois), a generalized middle-class fear of economic radicalism focused on the racialized intransigence of these newly assertive African American voices (Tuttle 1970, 208–22). This was compounded by a growing concern about labor radicalism: for many patriotic Americans, US cities seemed poised on the edge of a generalized revolution. Tuttle's macro-level framework thus allows him to identify the way in which the riot in Chicago was socially constructed as a crucial event in a longer historical process of American race and labor relations.

In constructing this level of analysis, he *layered* his meso- and micro-level findings on sociological literature on racial violence in general, while *contextualizing* his primary material from 1919 to a broader comparison with riots in the 1960s. In his final chapter, which addresses this comparison explicitly, Tuttle (1970, 258–68) was able to *classify* violence in both periods to draw attention to the similarities and differences, arguing that 1960s violence shared with the riots of 1919 similar structural conditions (e.g., rising expectations for African Americans, general social disorder related to warfare) as well as situational and contingent precipitants (e.g., hot weather, a catalytic fight or confrontation). By *linking* his historical material with (at the time) cutting-edge sociological research, Tuttle hoped to craft a longer narrative about the recurrence of racial strife in American life.

## THINKING ABOUT COMPLEX COMPARISONS

Does Tuttle's complex and sophisticated analysis of the Chicago riot provide a comparative technique for understanding the causes of race riots? Tuttle, I contend, is able to give us a sophisticated story about specific mechanisms critical to explaining violence in Chicago across multiple levels of analysis simultaneously. He, like Horowitz and Sewell, focuses on the process of collective violence formation – the ways in which micro-level isolated fights intersect with meso-level social cleavages to produce widespread local social mobilization in places affected by rapid demographic and economic change, and how these events were interpreted as elements in a macro-historical

moment of national disorder and crisis. In this sense, he does provide an explanatory framework capable of being compared across cases.

At the same time, it makes no sense to directly compare the riot in Chicago with events in, say, Washington, DC, with any hope of controlling for every confounding explanation. Both events were tied up in regional- and national-level phenomena (the Great Migration, demobilization after World War I, etc.) affecting many areas in the country that were, in Tuttle's view, crucial in explaining how a riot could aggregate out of both an otherwise routine fight and a shared background condition. Moreover, it is clear that events in places like Chicago and Washington, DC, did, in fact, help precipitate violence in other cities: news of earlier riots (often carried by national papers like the important African American paper *The Chicago Defender*) likely helped make them "thinkable" for participants in violence elsewhere. Because the riot is a multi-level phenomenon interdependent with other cases with which we might want to compare it, we need to disaggregate these levels of analysis and trace the connections between those different elements instead.

Rather than comparing to control for isolated variables, Tuttle's form of comparison therefore invites us to consider which elements the complex system of a particular riot can be decomposed into subsystems or sub-events and which are related to larger system connections with other riots. The more we can identify particular, bounded subsystems and the hierarchical relations among them, the more confident we can be that we can compare multiple mechanisms operating across different scales of action. Thus, the micro-level story Tuttle presents to explain the precipitating moment of the riot at its base was a simple street fight – a brawl involving members of different racial groups. Such events were relatively common in racially mixed American cities in 1919 and could be treated as a relatively stable subsystem, tied into larger dynamics of city racial hierarchy, but also directly comparable across cases. Indeed, precisely because they are bounded and discrete social events, such fights were neither multi-level nor interdependent across cases.

What made the fight of July 27 different? In Tuttle's estimation, it was the way this particular street fight engaged two other micro-level factors – the brawl took place in a racially contested physical space and a representative of state authority (the police officer) refused to intervene. These other factors are also subsystems, but their particular combination in this case meant that a simple brawl now had the potential to scale to something much larger. It would be fully possible in theory to develop comparisons between cases in other cities like Washington, DC, or Philadelphia that take these micro-level subsystems into consideration and to show how their peculiar juxtaposition provided such enabling conditions.

At the same time, however, as Tuttle reminds us, these micro-level subsystemic mechanisms were embedded in meso-level and macro-level systems that did connect riot events in Chicago to other cities and towns. These connections reveal that part of what made the brawl in Chicago actually scale to the level of

a city riot had to do with factors occurring in multiple places at once. Comparing the Chicago riot to other riots in these cases does not mean examining independent cases but instead multiple symptoms of the same underlying social process. This, in turn, might prompt us not to try to compare Chicago's riot to that in Washington, DC, directly to uncover this underlying cause, but rather to compare the *riot system* of 1919 to other moments of collectively violent activity (such as the race riots of the 1960s in many American cities). Of course, even in these comparisons, we are not examining independent events, but merely events that can be bounded into subsystems that, inevitably, at a higher level of historical analysis are themselves causally related.

## CONCLUSION

This chapter argues that archival analysis presents us with a highly useful way of demarcating the boundaries around contained and stable subsystemic-level events as well as identifying how these subsystems are themselves connected into larger hierarchical patterns. Classifying and contextualizing archival materials are also ways of grouping and situating events and processes on which we might want to focus analytic energy, drawing boundaries around what I have been calling subsystems (fights, racial geography, etc.) that are comparable and isolable from other subsystems. Conversely, linking materials and layering them in a larger scholarly conversation involves connecting those subsystems into larger narrative structures, situating them in a hierarchy of processes that help us make larger sense of the events and phenomena we care about, revealing the way individual micro-level activities are both shaped by and constitute macro-level patterns.

If we take seriously the need to examine causal mechanisms in a global way and to compare multiple scales of analysis simultaneously, scholars will be much better equipped to address the limits of traditional, variable-bound analysis, which rarely consider the complex nature of social causality. Although not a panacea, archival approaches such as that adopted by William Tuttle can powerfully reveal the way in which multiple factors together help produce real social outcomes.

## REFERENCES

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Bacharach, Michael. 2006. *Beyond Individual Choice*. Princeton, NJ: Princeton University Press.

Berkowitz, Leonard, and Edward I. Donnerstein. 1982. "External Validity Is More Than Skin Deep: Some Answers to Criticisms of Laboratory Experiments." *American Psychologist* 37 (3): 245–57.

Brass, Paul. 1997. *Theft of an Idol: Text and Context in the Representation of Collective Violence*. Princeton, NJ: Princeton University Press.
Chicago Commission on Race. 1922. *The Negro in Chicago: A Study of Race Relations and a Race Riot*. Chicago: University of Chicago Press.
Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
Dunning, Thad. 2012. *Natural Experiments in the Social Sciences*. Cambridge: Cambridge University Press.
Durkheim, Emile. 1982. *Rules of Sociological Method*. Edited by Steven Lukes. Translated by W. D. Hall. New York: Free Press.
Farge, Arlette. 2013. *The Allure of the Archives*. Translated by Thomas Scott-Railton. New Haven, CT: Yale University Press.
Goldstone, Jack A. 2001. "Toward a Fourth Generation of Revolutionary Theory." *Annual Review of Political Science* 4 (1): 139–87.
Hall, Peter A. 2003. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer, 373–404. Cambridge: Cambridge University Press.
Hilborn, Robert C. 2004. "Sea Gulls, Butterflies, and Grasshoppers: A Brief History of the Butterfly Effect in Nonlinear Dynamics." *American Journal of Physics* 72 (4): 425–27.
Holland, John H. 2000. *Emergence: From Chaos to Order*. Oxford: Oxford University Press.
Horowitz, Donald L. 2001. *The Deadly Ethnic Riot*. Berkeley: University of California Press.
Kocher, Matthew A., and Nuno P. Monteiro. 2016. "Lines of Demarcation: Causation, Design-Based Inference, and Historical Research." *Perspectives on Politics* 16 (4): 952–75.
Lewontin, Richard C. 2000. *The Triple Helix: Gene, Organism, and Environment*. Cambridge, MA: Harvard University Press.
Morck, Randall, and Bernard Yeung. 2011. "Economics, History, and Causation." *Business History Review* 85 (1): 39–63.
New York Times. 1919. "For Action on Race Riot Peril," October 5, sec. F.
Padgett, John Frederick, and Walter W. Powell, eds. 2012. *The Emergence of Organizations and Markets*. Princeton, NJ: Princeton University Press.
Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: Wiley.
Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688–701.
Sewell Jr., William H. 2005. *Logics of History: Social Theory and Social Transformation*. Chicago: University of Chicago Press.
Simon, Herbert A. 1996. *The Sciences of the Artificial*, 3rd edn. Cambridge, MA: MIT Press.
Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
Stoler, Ann Laura. 2002. "Colonial Archives and the Arts of Governance." *Archival Science* 2: 87–109.
Tilly, Charles. 2003. *The Politics of Collective Violence*. Cambridge: Cambridge University Press.

Tuttle, William M. 1970. *Race Riot: Chicago in the Red Summer of 1919*. Champaign: University of Illinois Press.

West, Geoffrey B. 2017. *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. New York: Penguin.

Zerubavel, Eviatar. 1996. "Lumping and Splitting: Notes on Social Classification." *Sociological Forum* 11 (3): 421–33.