A Manual for Cultural Analysis

Luke J. Matthews, Ryan Andrew Brown,

David P. Kennedy

RAND
CORPORATION

www.rand.org

Preface

The RAND Corporation seeks to improve policy through research and analysis. Within RAND, the Center for Qualitative and Mixed Methods acts primarily as an internal forum for dissemination of mixed methods. Mixed methods are a set of techniques used to process both qualitative and quantitative data. One approach to mixed methods is to begin a research process with highly qualitative data—such as documents produced by people in the course of their lives or notes from ethnographic interviews—and then process these data into fully quantitative statistical analyses.

We recognized that, within policy and academic research, dealing theoretically and empirically with culture (i.e., socially learned information that influences beliefs and behaviors) is one of the primary motivations to engaging in a qualitative-to-quantitative research process. As with much academic social science, some policy research is conducted within the framework of utility maximization (i.e., the notion that actors are agents who optimize some function for outcomes that result from their behaviors). Taking the concept of culture seriously, however, means that researchers must consider that individuals do not have perfect information when they optimize utility. Both the outcomes and the range of behaviors to achieve them often are learned by observing other individuals, and the variable distribution of this socially learned information is what we recognize as *culture*. Culture thus becomes a crucial part of the information context in which utility is maximized.

Integrating culture within social science in general requires mixed methods because culture comes to the researcher in a qualitative form but must be quantified for research paradigms that involve formal tests of theory. The authors of this manual come from different traditions within anthropology, including cognitive, evolutionary, and biocultural approaches, and we have worked in a variety of applied research roles. We discovered that we each had independently encountered many of the concepts and methods in this manual and found them useful in providing a formal approach to dealing with culture. Currently, these concepts and methods are scattered among various articles from disconnected traditions within anthropology and related areas of social science. Thus, we felt *A Manual for Cultural Analysis* was needed.

RAND Ventures

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest. RAND Ventures is a vehicle for investing in policy solutions. Philanthropic contributions support our ability to take the long view, tackle tough and often-controversial topics, and share our findings

Contents

Figures and Tables

Figures

Tables

Summary

Distinct academic traditions—including anthropology, biology, and sociology—have come to formally theorize culture as socially learned information that can influence peoples' beliefs and behaviors. Cultural processes therefore impact a myriad of applied issues, ranging from national security strategy that attempts to anticipate responses by partners and adversaries to health policy that strives to disseminate best practices through physician communities or to promote more scientifically informed medical beliefs among the general public. In all these cases, actors may strive to maximize some utility, but they do so in a limited information environment, much of which they have sampled from local social networks and longstanding inherited traditions.

Understanding cultural processes in applied contexts often necessitates formal analytic methods capable of describing and predicting cultural phenomenon within statistical frameworks. Chapter 1 of this manual discusses the definition of culture as socially learned information. Chapter 2 discusses some of what is known from existing literature about the adaptive reasons animals and people rely on such information as well as what is known about how higher-level cultural patterns emerge from each individual engaging in potentially simple social-learning processes.

Drawing from cognitive and evolutionary anthropology traditions, we next describe a set of tools capable of dealing with cultural data at various emergent levels, ranging from variation among individuals within local subcultures to small- and large-scale network topologies and finally to longstanding lineages of inherited cultural information. Cultural consensus analysis, an approach closely related to principal component analysis, is a key tool at the most granular level of variation among individuals within groups (Chapter 3). Social network analysis provides tools for examining how cultural variations are structured within and across groups (Chapter 4), and autoregression approaches provide means to deal with longstanding patterns of cultural variation among groups (Chapter 5).

These various techniques and their application to different levels of culture emergence are known in the published literature, however, to our knowledge, have never been organized into a single manual structured around a formally theorized notion of culture. *A Manual for Cultural Analysis* strives to do this, and we hope it can be a helpful resource for students, academics, and applied researchers who need to work with cultural data at multiple levels of analysis.

Acknowledgments

Abbreviations

ABM	agent-based model
AIDS	acquired immune deficiency syndrome
CBQ	Childbearing Questionnaire
CCA	cultural consensus analysis
EEA	environment of evolutionary adaptation
HIV	human immunodeficiency virus
PC1	first principal component
PCA	principal component analysis

1. Introduction: The Need for an Operational Concept of Culture

Anthropologists are not unknown in the world of public policy, but what anthropology as a science has to contribute is often unclear to nonanthropologists. Much of this ambiguity stems from many academic anthropologists discarding attempts to operationalize the discipline's core theoretical concept: culture.

Academic disciplines usually are founded around either a subject matter or a theoretical perspective. Anthropology started as the first kind of discipline, because it originated from the comparative biological studies of 19th-century naturalists. These naturalists, of which Charles Darwin is one famous representative, were the first to attempt a systematic survey of the biological diversity of our planet. In doing so, they identified the fundamental evolutionary mechanisms that govern why the various forms of species are distributed as they are on Earth.

Anthropology arose from the recognition that this evolutionary history must include humans and that there must be some kind of continuous and unbroken path leading from our biological evolution through to the diversification of our behaviors, beliefs, and technologies. Early anthropologists set out, as had their naturalist precursors, to systematically survey the full diversity of humankind (Langness, 1987).

Although initially oriented in the study of human diversity, anthropology, as it became established, was soon centered around a theoretical perspective: the concept of culture (Langness, 1987). The use of the word "culture" in anthropology shares a linguistic origin with "culture" in the context of bacteria. In a bacterial culture, the same strains of microorganisms can grow in different ways, depending on the chemicals that surround them in a petri dish. Analogously, human culture is thought to be the set of social influences that alter an individual's behaviors and beliefs over the course of a lifetime. Culture gained and retains prominence in anthropology because it is a plausibly powerful explanation for the incredible diversity of human behavior across different societies (Nash, 1998, Chapter 2).

The scientific study of culture requires operationalization, as with any other theorized mechanistic process. Just as economists developed a set of methods to investigate their core concept (maximal utility decisions under scarcity), so too have a minority of anthropologists quietly developed operational tools for the study of culture as a mechanistic process. This fact is not broadly known because much of the discipline of anthropology shifted away from operationalized models for scientific research, including research on culture. Much of mainstream cultural anthropology rejected the approach of systematic, operational, and scientific methods in favor of an interpretive understanding of human diversity grounded in academic traditions for the interrogation of texts (Bernard, 2018, pp. 1–14).

This minority group of anthropology researchers, however, continued with the scientific study of culture. Over the past several decades, evolutionary biologists, cross-cultural

psychologists, and cultural neuroscientists have increasingly recognized that culture is indispensable to understanding human diversity. All these researchers have made important theoretical and empirical contributions to scientific cultural research.

A Definition of Culture

Although theorists disagree about many aspects of culture, they agree that culture refers to something learned rather than biologically inherited (Brumann, 1999). While the concept of culture is usually invoked to understand the behavior and thought patterns of groups, only individuals can learn culture, and individuals are the only source of cultural data (Handwerker, 2001, p. 10). Therefore, any definition of culture must begin by acknowledging that human beings possess it and identify how individual human beings learn and processes information. However, since culture is learned primarily through other people, it is also the result of social interaction and so is shared collectively by members of groups. Given this, culture is understood to be something that is both socially and individually constructed. *We can therefore define culture as the shared information that results from social-learning interactions among a group of people.* This shared information influences their behaviors and beliefs. Culture is molded by feedback from the individuals' own experiences, which they use to modify their beliefs and behaviors, thus resulting in a revision of the information they pass on socially in a dynamic feedback.

One common assumption about culture (which has been proven incorrect) is that it is a unitary, internally consistent seamless web that contains unambiguous rules for behavior (DiMaggio, 1997, p. 267). Instead, cognitive researchers see culture as fragmented and inconsistent. At any point in time, individuals may have internalized cognitive models that are contradictory. These models, while guides for behavior, can never have a one-to-one correspondence with behavior outputs because of their heterogeneity. Rather than acting as a blueprint for behavior, culture acts like a "toolkit" of strategies that individuals use to choose among behavioral options depending on momentary external circumstances (DiMaggio, 1997).

The value of this definition is largely that it delimits a scope of study for cultural processes that is not a vague sum of all things people do (e.g., Tylor, 1871). We are specifically investigating contexts in which social learning powerfully constructs emergent patterns of shared behavior among sets of people. Cultural analysis rightly can play a secondary role for behaviors that can be fully accounted for through other paradigms, such as universal genetic tendencies or straightforward utility maximizations. This manual briefly surveys the known mechanics of how culture actually happens among sets of social learners (Chapter 2). We then proceed to explain how to operationally measure culture through basic multivariate statistical modeling of coded qualitative data (Chapter 3). Chapter 4 discusses how to use the output from analyses in Chapter 3 in standard regression models and in social network models that test which social transmission pathways govern the diffusion of a given set of cultural data. Chapter 5 addresses how to identify

and correct for the statistical dependencies that often arise from the shared and learned nature of cultural phenomenon.

2. What We Know About How Culture Operates and Why

Culture as an Adaptive Buffer

Culture provides humans with an ability to pool their cognitive resources and likely arose as an evolutionary adaptation to highly variable and fluctuating Pleistocene environments (Potts, 1998; Tomasello, 1999). Culture gives humans the ability to address practical problems they face daily in their efforts to survive, feed themselves, and reproduce. Compared with other species that interact socially, humans are uniquely adapted to pool their cognitive resources not only over populations but also through historical time (Tomasello, 1999). This creates what is called "distributed cognition," which acts as a cognitive division of labor among the members of a social network (DiMaggio, 1997). Furthermore, humans are uniquely able to pool these resources among a socially cooperative group of strangers or partial strangers, unlike other socially interacting species, which almost exclusively interact with kin (Chase, 1999). Although there is sufficient evidence that nonhuman animals also use culture for survival, the uniqueness of human culture lies in the degree to which it pervades so many aspects of human life (de Waal, 2001, p. 29; Harris, 2001, p. 125).

To understand this unique evolution of culture in the human species, we need to consider the environment under which humans evolved and developed this trait. We have limited information about the conditions of the environment of evolutionary adaptation (EEA) that served as the backdrop for human evolution during the Pleistocene (Bock, 1999, p. 197). The evidence we have points to an incredibly unstable environment, which would not have had consistent enough selective pressure for specific behavioral traits to evolve (Potts, 1998). Instead of the evolution of specific behavioral responses to specific environmental conditions, the EEA selected for plasticity in behavioral responses to changing environmental conditions. In a sense, there was no single EEA for the human capacity for culture, which evolved under variable conditions and has been such a successful behavioral strategy that humans now inhabit almost every ecological zone on earth.

Culture creates a buffer between people and their physical world. It is a type of virtual environment with which people interact. Behavioral responses that are directly tied to physical environmental cues and not through culture need to be developed over long periods of time under conditions of systematic selection of genetic variations. The variations that end up being widespread are those that allow individuals a greater probability of success at securing resources from the environment, which leads to greater probability of success at survival and reproduction. For human beings, culture fulfills a similar role as genetic evolution. Through simplified models of reality that filter out information and constrain behavior within a limited set of options, culture

provides individuals with strategies for how to live their lives in a way that assures consistent access to resources for survival and reproduction within a given environmental circumstance.

Learning Mechanisms

A number of social learning mechanisms have been studied in theoretical models and through empirical research that undergird macroscale cultural processes. These mechanisms include stimulus enhancement, local enhancement, affordance learning, imitation, teaching, and many others (Laland and Hoppitt, 2003; Thornton and McAuliffe, 2006; Tomasello and Call, 1997; Whiten and Ham, 1992). Researchers use these terms to distinguish fairly nuanced differences in how individuals process social information. Mechanisms are usually differentiated on cognitive criteria as to the types of information that is used by individuals and how that information affects the subsequent formation of behaviors and beliefs. The most important distinction to make for our purposes is between stimulus/local enhancement learning and imitation learning.

Stimulus/Local Enhancement

Enhancement learning occurs when an individual's attention is drawn to a place or object because of a social experience (Hoppitt and Laland, 2008). When the attention is drawn to a place, this learning is called *local enhancement* (Galef and Giraldeau, 2001), while attention drawn to an object or particular part of an object it is termed *stimulus enhancement* (Whiten and Ham, 1992).

For example, suppose an observer sees someone using an office appliance he previously had not noticed in a corner of the office kitchen counter. The observer never noticed the unusual-looking appliance in other times he passed through the kitchen. In subsequent trips to the kitchen, the observer sees the appliance and realizes that it is a coffee grinder. The observer proceeds to grind coffee with it by fiddling with its nobs and settings because he did not watch closely enough to understand the sequence of steps required to grind coffee with this particular grinder.

In this scenario, the social learning mechanism was stimulus enhancement because the observer's attention was drawn to the grinder. The observer did not watch closely enough to mimic operating the grinder; the grinder's operation was learned individually. Although the observer did not socially learn anything about how to operate the grinder, or even that it was for grinding coffee, the simple social process of drawing attention to the grinder effectively transmitted the behavior of grinding coffee in the office kitchen.

Imitation

Imitation differs significantly from enhancement learning because the observer socially learns something about how to perform the observed behavior (Buttelmann et al., 2007; Byrne,

2003; Whiten, 1998). Classically, imitation was defined as mimicking the observed motor actions, although sometimes it is defined as mimicking the intended actions or intended effects of the observed actions. An example of the latter would be an observer seeing someone grinding coffee, who then might copy the observed actions to grind coffee. The observer would not, however, copy the demonstrator spilling some of the beans on the counter only to pick them up. It is obvious to the observer that spilling the beans and then picking them up was not part of the intended actions.

In our coffee grinding example, the intended versus unintended actions were obvious, but this is not always the case. Whiten et al. (1996) studied imitation among both chimpanzees and human children. They found that children were very faithful imitators even of nonessential parts of a task to unlock a puzzle box. For example, children rotated a pin three times prior to removing it, as they had seen a demonstrator do even though the pin spun freely in the socket and the rotations accomplished nothing. Horner and Whiten (2005) demonstrated a similar tendency of human children to copy causally irrelevant aspects of observed actions.

One reason individuals might imitate arbitrary variations is because they interpret all such variations as normative behavior for their social group (i.e., it is the way it is supposed to be done). Human children appear particularly eager to embrace this or similar kinds of social thinking that create social functions from the imitation of technical tasks (Nielsen, 2009).

Often, however, observers imitate because they are not sure what parts of a demonstrator's actions are important to the task. This is especially true as tasks become more complex.

Consider the behavior of human adults who imitate the stock purchases of Warren Buffett. When Buffett goes long on a stock or industry, people take notice, and many imitate his purchase, even though they are unsure about what might make it a good purchase. Perhaps, on some purchases, Buffett might just be guessing? Individuals imitating Buffett might often be copying arbitrary decisions not linked to Buffett's success as an investor, but, on average, they might do well to imitate Buffett because he has been highly successful. That can be a winning strategy even though the individuals doing the imitating themselves are unable to determine which purchases by Buffett are well-informed and which are guesses. Probably few, if any, Buffet imitators think his investing patterns are social norms, or the way it should be, but stock investing is a highly complex task for human cognition, so this imitation of even arbitrary behaviors and decisions by Buffett could be an adaptive strategy.

Reinforcement

Even as individuals employ social learning mechanisms, they likely are always simultaneously incorporating feedback from their individual experiences. Feedback from individual experiences often occurs through reinforcement learning, in which an individual repeats behaviors that have delivered more rewards for time and energy invested. The rewards individuals optimize are usually of some intrinsic biological benefit (e.g., food, sex, shelter) or exchangeable for intrinsic benefits (e.g., money, social approval, prestige).

6

The simultaneous occurrence of individual learning, even as individuals engage in social learning, can create important feedback loops. For example, consider a task with two alternative solutions that differ slightly in their payoff once an individual is proficient in one solution. This will be familiar to anyone who has ever switched from using one software program to another to perform the same statistical analysis. Often, the switch is made because the same analysis can be done faster in one software than another, but learning any software results in an initial reinforcement learning curve, during which the learner becomes fully proficient. Once proficient with one variant, individuals should favor avoiding the zone of lost profit by ceasing to learn new behavioral alternatives to obtain the same reward (Figure 2.1).

Figure 2.1. Idealized Learning Curve for Alternative Actions
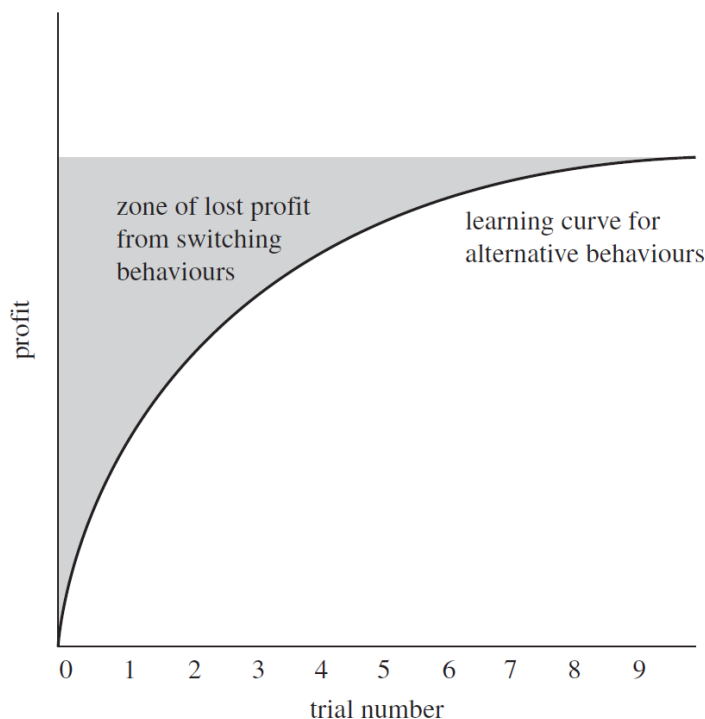
It is important to note that if an individual is already proficient in one variation, they effectively pay an opportunity cost to learn the other method. Consider a person who is conducting data analysis in Microsoft Excel. They can conduct the same analyses faster, more consistently, and likely with more nuance in software such as R or SAS. Prior to gaining proficiency with R or SAS, however, they can complete any single project faster by sticking with the Excel routines they already know.

This balancing of initial learning curve costs against opportunity costs means individuals can get stuck in equivalent or even suboptimal variations of a behavior. Individuals often use social

learning to guide their initial attempts at a behavior (e.g., they learned Excel or R from their colleagues or mentors). In this way, the dynamic interaction of reinforcement and social learning can create variable cultural traditions that typify various social groups (Franz and Matthews, 2010).

Social Learning Strategies

Given the many complexities of social learning dynamics, theorists have been interested in copying strategies that individuals might use to better ensure that social learning produces adaptive behavior rather than suboptimal or maladaptive behavior. These strategies focus on *whom to copy and when*. Note that the mechanisms discussed earlier distinguish *what* information is acquired and *how* it affects the learner. The strategies that follow focus on when the learner uses social learning mechanisms and whom the learner chooses to observe with some social learning mechanism or combination of them.

Strategy 1. Use Social Learning Only When Necessary

One way individuals can avoid being trapped in suboptimal behaviors from social learning is not to learn socially. Theorists have proposed that individuals should have evolved to engage in extensive social learning only when it pays off on average as compared with an alternative of ignoring social information and relying on personal experiences only. Some research has explored this question of when social learning pays off, and the basic finding is that social learning is very frequently a successful strategy because it is so much less costly in time and energy than is individual learning through one's own trial-and-error experiences. Social learning is usually less accurate than individual learning, but it also is usually much less costly for the individual to use because through social learning, individuals can avoid spending the time and energy on individual learning trials. It is important to note that social learning often does not enable individuals to develop more adaptive behaviors than by individual learning, but rather enables them to produce equally adaptive behaviors but with less time and energy invested in learning. Mathematical models of these learning dynamics suggest that as long as some individuals in a group have individual learning experiences to draw from, it usually pays for others to practice social learning whenever there is at least some moderate cost to acquiring one's own experiences (Laland and Kendal, 2003).

Importantly, the dynamic feedback of individual experience can often ensure that social learning typically results in adaptive behaviors even if the social learning strategy is something extremely simple (e.g., imitating a random person). Nunn et al. (2009) conducted extensive agent-based model (ABM) simulations that suggested that behaviors with adaptive consequences spread effectively even with very simple social learning mechanisms. This is because individual learning and natural selection both act to eliminate demonstrators of maladaptive behaviors from the population that social learners might observe. For example, consider traditional food taboos:

Taboos against poisonous foods could evolve without complex social learning (such as imitation or peer pressure) because the demonstrators who do not follow the taboo die from eating poisonous foods. Thus, the ill-fated demonstrators are not around for others to copy. Since those who follow taboos about poisonous foods are naturally more common (they die less), simple social learning is sufficient to maintain such taboos.

This example is not merely hypothetical. Henrich and Henrich (2010) showed that Fijian food taboos during pregnancy selectively targeted potentially toxic marine species and allowed consumption of relatively nontoxic ones. The simulation findings from Nunn et al. (2009) suggest that, specifically because natural selection is at play (i.e., toxic foods that can kill you or your baby), complex social learning mechanisms are not necessary for the evolution of such taboos. Complex social learning might be a factor because humans have complex cognitive worlds, but humans probably did not evolve this complexity to deal with straightforwardly selected traditions (e.g., taboos against foods that actually are dangerous).

Fogarty et al. (2012) obtained similar results by creating an ABM that hosted a competition for social learning strategies. Other teams of researchers submitted strategies encoded as mathematical rules for when to innovate, copy, or play an already learned behavior. The winning strategies focused primarily on the age of the agent's current information—that is, when it got too out of date, they copied. In fact, in the second most-effective strategy, agents innovated only when they were the first individual in a new simulated environment. This second most-effective strategy engaged solely in social learning, however, in all other situations. After the first round, the agents in this second most-effective strategy merely copied whatever individuals around them did. Because natural selection was an active part of the simulations, most surviving individuals did reasonably adaptive behaviors: the ones who did not, died, so they were not around to be copied. Thus, for a behavior that truly has adaptive consequences, relatively simple social learning strategies are often sufficient essentially because other ecological processes guarantee that demonstrations of maladaptive behaviors are relatively rare.

Strategy 2. Prestige Bias

One way a social learner can try to improve on a random copying method would be to mimic individuals who are highly successful in terms of money, career success, offspring produced, or any other outcome of some intrinsic value to the observer. This strategy is called *prestige* or *indirect bias* in the literature (Boyd and Richerson, 1985; Henrich and Gil-White, 2002). The idea is that such a strategy enables the social learner to copy many features of a successful individual without needing to assess exactly which features are causally related to their success. This will, of course, result in copying many arbitrary features of the prestigious individual, but, under some conditions, it will result in copying of the causally salient features that lead to the desired success. The copying of investment decisions of Warren Buffett is likely an example of prestige bias.

Prestige bias is thought to have evolved as an effective social learning strategy in the environments in which our psychology evolved (Henrich and Gil-White, 2002). Among hunter-gatherer groups and horticultural tribes, social arrangements that characterized most of our evolutionary past—the behavior of informal leaders—may be modelled by others principally because these leaders are personally known to the observers to be highly adept at particular technological and social skills (Service, 1966, Chapter 4; Sahlins, 1968, Chapter 5). The present times are very different because modern communication technologies enable celebrities, sports stars, and famous public intellectuals to be viewed as prestigious by orders of magnitude more people than could ever have occurred in traditional environments.

Strategy 3. Conformity Bias

Another social learning strategy that has received a fair bit of research is conformity bias. One familiar social context for conformity bias is peer pressure. The potential maladaptive outcomes of conforming to a group, for instance, are well known to every parent of a teenage child. As with prestige bias, however, in the settings during which our psychological tendencies evolved, conforming to what most people did was likely a highly adaptive strategy (Boyd and Richerson, 1985; Henrich and Boyd, 2001). Even today, conforming to a group may be often a good bet on average.

Framing conformity bias more technically, we can state that any social learning mechanism that statistically overweights the majority of observed behaviors and beliefs in influencing future behaviors and beliefs can be considered as conformity bias. Thus, when 60 percent of observed individuals do behavior X, a conformity-biased social learner exhibits more than 60 percent probability of performing X (Figure 2.2).

Figure 2.2. Social Learning Curve for Individuals Practicing Conformity Bias
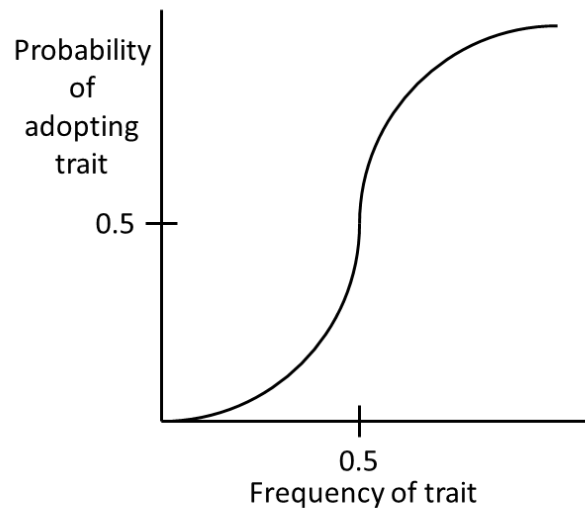
SOURCE: Redrawn with modifications from Boyd and Richerson, 1985.
NOTE: The probability of the conformity-biased social learner adopting a trait is a nonlinear function of the trait's frequency among the sample of individuals the learner observes.

Conformity bias can often serve as a valuable heuristic tool to cheaply arrive at an adaptive outcome. Under situations in which the observed individuals face the same payoffs for their behavior as the observer, and assuming there is any degree of individual reinforcement learning, on average, the majority of individuals should perform the more adaptive behavior among alternative options (Franz and Matthews, 2010). Thus, biasing behaviors in such situations to those of the majority is an adaptive strategy.

Emergent Higher-Order Patterns from Social Learning

Social learning mechanisms operate at the level of individual psychology, but the operation of social learning produces many emergent patterns in the knowledge ecology of the interacting, socially learning individuals. For example, even simple social learning mechanisms can produce durable traditions that typify social cliques within groups and entire social groups (Franz and Matthews, 2010; Matthews, 2009; Whiten et al., 1999). The ecology of knowledge maintained by social learning can exhibit inheritance properties similar to genetic inheritance. For example, the cultural diversity of tool assemblages and language phonemes decreased in parallel with genetic diversity during human migration events that serially reduced population sizes (Atkinson, 2011; Lycett and von Cramon-Taubadel, 2008). Just as smaller populations contained less genetic diversity, so too did they contain less diversity of ideas and behaviors.

This situation is analogous to the widely recognized interaction of micro- and macroeconomics. Most economic theory is grounded in the notion that individuals make

decisions under scarcity, with an aim toward optimizing some outcome measure. At a broad scale, however, macroeconomic patterns emerge from all these individual decisions. Macroeconomic patterns themselves are not expressible as individual choices, even though they are causally reducible to interactions of individual choices. Consider the price of a commodity such as gasoline. No individuals choose to balance supply and demand to set a system-wide price per gallon. Yet, individual choices about this specific gas station versus that one, combined with constraints of extraction and transport costs, bumps from speculators, and the like. interact to create a reliable and often predictable adjustment of price based on the balancing of supply and demand.

Analogously, cultural dynamics emerge from interactions of social learners, none of whom intend to create the emergent properties of the system. As discussed earlier, a variety of social and asocial mechanisms interact to create the frequent emergence of adaptive traditions. This dynamic interaction means cultural systems can effectively evaluate the worth of ideas for the participants in them, rather than those same participants having to evaluate the ideas independently. Thus, culture is often regarded as a form of distributed cognition or memory; a kind of social adaptation that bootstraps human cognitive capacity beyond what it is capable of on an individual basis (DiMaggio, 1997; Henrich and Boyd, 2001; Laland and Rendell, 2013; Provenza, 1995; Richerson and Boyd, 2005). Next, we briefly review some of the more important emergent cultural properties that have been well-characterized by researchers.

Group Typical Traditions

As individuals interact through social learning, they can create locally adapted behavioral traditions via feedbacks from the reinforcement. If people are living in different environments that favor subsistence on different crops, for example, cultural traditions can become locally adapted much more rapidly than would be enabled by genetic evolution (Richerson and Boyd, 2005, Chapter 2).

Culture also can produce multiple equilibrium states that are more or less equivalent yet different solutions to an ecological problem. This is because social ties are not distributed homogenously; rather, our social ties are clustered within families, social groups, workplaces, countries, and the like. This means that even neutral cultural traditions will tend to homogenize within the clustered boundaries of these nested levels of human social interaction.

Biases toward cultural conformity interact with the heterogeneous distribution of social ties and increase the homogeneity of behaviors within groups (Henrich and Boyd, 2001). This becomes clear if we consider the contrast of what happens with genes and culture when people from different populations intermarry. The children of such marriages carry equal proportions of genes from both populations, but they frequently learn only one of the parental languages or religions or are more proficient in one than another (Cavalli-Sforza and Feldman, 1981). The result can be that populations that are genetically inseparable can still maintain strong cultural differentiation. The populations of France and Germany are a real-world examples of this: They

are essentially indistinguishable genetically and yet have longstanding and contrasting cultural practices.

Conformity bias, however, is not always necessary even to fix alternative behaviors typical of different social groups because the interactive effects of avoiding trial-and-error learning costs mean even the simplest social learning mechanisms are capable of producing multiple equilibrium states (Franz and Matthews, 2010). Franz and Matthews (2010) conducted a thorough set of simulations to validate this point, but the finding is intuitive to anyone who has learned a software program. Once you are proficient in Excel, R, or Python, you have already paid the upfront cost to learn a software and so can perform needed tasks relatively quickly. Switching to a new software for any particular task requires more initial learning cost, so individuals tend to stick with the software they know and the software, most likely, they were introduced to by someone else. Thus, individuals who know R users tend to become R users, and they stick with R because they do not want to pay the cost to learn something else. This can occur just because of the interaction of social networks with individual learning costs and without any cultural norms, peer pressure, or collective identities being formed around such things as "R-userness." This mechanism most likely causes largely equivalent software products to cluster by academic discipline (e.g., sociologists use Stata, biologists use R, medical researchers use SAS, engineers use MATLAB).

The emergence of multiple equilibrium states presents new opportunities for social signaling of group co-membership among individuals who are unrelated and do not know each other from past interactions. This ability to recognize anonymous individuals as members of one's tribe, religion, or nation creates evolutionarily novel ways for individuals to solve cooperation problems that otherwise prevent cooperation among anonymous individuals. Without these cultural mechanisms, individuals find themselves in a Hobbesian state of nature in which they are less able to predict whether an individual will cooperate or defect in any given interaction (Sahlins, 1968).

Sahlins (1968) made the case that most hunter-gatherer groups do in fact live in a Hobbesian state of "Warre; and such a Warre, as is of every man, against every man" (Hobbes, 1909). By this, Sahlins did not mean that rates of actual violence are high among hunter-gatherers (they are not), but that without the higher-level institutions of chiefdoms, states, and the like, there always exists among hunter-gatherers an imminent risk of violence between any anonymous individuals. To quote Sahlins (1968):

> In the social condition of Warre, force is a resort legitimately available to all men. There need not be violence, but neither is there assurance to the contrary.

An example of this phenomenon was reported by the early social anthropologist Radcliffe-Brown upon entering an Australian Aboriginal camp where his Aboriginal guide could not find even the most distant of relatives (Radcliffe-Brown, 1913). The guide refused to sleep in the camp and reported to Radcliffe-Brown that he was frightened because if those men were not his relatives, then they must be presumed enemies.

13

The emergence of multiple cultural equilibria is thought to be one mechanism that removed humans from this condition. Arbitrary cultural equilibria enable social signaling among anonymous individuals through arbitrary variations in liturgies, accent, food preparation, national anthems, and the like. This in turn yields anonymous individuals some probabilistic degree of confidence that, in a given situation, those with similar cultural backgrounds are likely to cooperate or not cooperate similarly as themselves (Atkinson et al., 2008; Hamilton, 1975; Matthews, 2012; Sahlins, 1968). This removes the uncertainty of outcomes from interactions with anonymous other individuals that leads to the risk of preemptive violence described by Radcliffe-Brown and Sahlins. Note that just as many people, or likely more, may die or be seriously punished among modern societies as compared with hunter-gatherer groups. In the context of recognizable cultural institutions, however, these acts of violence are predictably structured along fracture points between ethnicities, religions, and states or mediated by these institutions' use of force internally.

These traditionally recognized forms of culture are thought to function in part to mediate cooperative interactions among anonymous individuals. Culture, however, has many other functions as well; for example, we already discussed the toxin-avoidance functions of food taboos. This manual will cover still more potential functions and effects of culture, but when culture creates multiple stable equilibria, it often is reasonable to at least ask whether these equilibria might function to create social identities that mediate anonymous interactions.

Cumulative Culture

Group-typical traditions that incorporate reinforcement feedback potentiate another emergent property of human culture termed *cumulative culture* by researchers. Compared with other animals that use simple social learning, this unique feature of human culture is enabled by our having high-fidelity social learning mechanisms, such as imitation and artefactual methods for storing cultural data outside of individual minds (e.g., pictures, architecture, writing, books, libraries) (Laland and Rendell, 2013). Cumulative culture occurs whenever a cultural construct increases in efficiency or complexity over time (Dean et al., 2014; Laland and Rendell, 2013).

Culture can accumulate efficiency because individuals continuously improve their socially acquired behavior through their own experiences. These improved behaviors then become the new models for the next set of social learners, which results in a ratchet-like effect such that a cultural tradition can continuously improve upon itself (Boyd and Richerson, 1985).

Consider the many technologies each of us uses routinely that are much more efficient than the same technologies used even just a decade ago. No single individual, no matter how brilliant, could independently invent an iPhone or a laptop computer without building on the many incremental learning steps of prior innovators.

It is not only our technologies that exhibit this type of cumulative dynamic. Our systems of philosophy and theology also ratchet in a cumulative manner. So do our legal systems and artistic productions. In some cases, there is a clear efficiency gain toward some end in these

systems. For example, one could plausibly argue that our modern legal system is at least somewhat more efficient at distributing equitable justice than its ancient Roman predecessor. Again, it is hard to imagine our current legal frameworks being invented wholesale without incremental and historic improvements, such as the writings of Cicero, the Magna Carta, or the Bill of Rights. No single legal genius could sit down and generate the whole thing de novo.

Some of our cultural ratcheting may increase complexity more so than efficiency. For example, our modern philosophical and artistic traditions present an incredible breadth of thought unknown to prior generations. An important factor in their evolution is the role of biases introduced by prior cultural information. For example, Eastern and Western philosophy and theology display notable differences in approach that result from stepwise acceptances of different premises. Once one premise is adopted, it makes other premises or conclusions more reasonable. In this way, Eastern and Western cultures have arrived at different attitudes, for example, toward whether individual liberties or collective good should have priority (Fincher et al., 2008; Hofstede, 1984; Hofstede, 2001; World Values Survey Association, 2016; Inglehart and Baker, 2000; Inglehart and Welzel, 2010). This process of walking down alternative cultural paths is termed *secondary bias* by some anthropologists, and it can result in the accumulation of multiple cultural equilibria but along very different cultural trajectories (Durham, 1991).

Cultural Phylogenetics: When Like Begets Like

Thus far, we have discussed mostly how culture behaves over short time scales of months, years, or decades. At these time scales, social learning processes can generate variation, establish subcultures with similar socially learned behaviors, and enable culture to be an adaptive buffer between individuals and their environments. Sometimes, however, cultural variations can be passed on over much longer time scales that involve multiple human generations. Our languages, religions, and governments all persist over centuries or millennia. That kind of time duration is remarkable in that the same behavioral phenotype is being maintained by social learning even as all the individuals involved die and are replaced over and over.

An important implication of highly durable cultural traditions is that they may exhibit nested patterns of relationships that are analogous to patterns we see among biological species. These nested patterns of descent, called *phylogenies*, reflect how recently different cultural traditions shared a common ancestor prior to their divergence. For example, languages often exhibit treelike patterns of relationship to one another (Figure 2.3). In the case of languages, splits may occur between populations largely because of geographic separation. Geographically separated populations of speakers then diverge in the arbitrary correspondences of the sounds of words and the things they signify. Such events that happened even thousands of years in the past still define how languages are similar or different to one another in the present (Figure 2.3).

15

Figure 2.3. Cultural Phylogeny from Bayesian Analysis of Linguistic Data (Cognate Word Lists) for Indo-European Languages

Hittite
Luvian
Lycian
Tocharian_B
Tocharian_A
Greek_Mod
Greek_ML
Ancient_Greek
Armenian_List
Armenian_Mod
Classical_Armenian
Albanian_Top
Albanian_G
Albanian_K
Albanian_C
Avestan
Wakhi
Digor_Ossetic
Iron_Ossetic
Old_Persian
Waziri
Afghan
Tadzik
Persian_List
Baluchi
Kurdish
Vedic_Sanskrit
Kashmiri
Romani
Singhalese
Nepali_List
Bihari
Bengali
Assamese
Oriya
Marathi
Gujarati
Sindhi
Marwari
Hindi
Lahnda
Urdu
Old_Prussian
Latvian
Lithuanian_ST
Slovenian
Old_Church_Slavonic
Serbocroatian
Bulgarian
Macedonian
Russian
Polish
Ukrainian
Byelorussian
Lusatian_L
Lusatian_U
Czech
Czech_E
Slovak
Gothic
Old_English
English_ST
Frisian
Dutch_List
Flemish
Old_High_German
Luxembourgish
German_ST
Old_Norse
Faroese
Icelandic_ST
Danish
Riksmal
Swedish_List
Swedish_Up
Swedish_VL
Spanish
Portuguese_ST
Catalan
French
Walloon
Provencal
Italian
Friulian
Romansh
Ladin
Sardinian_C
Sardinian_N
Sardinian_L
Romanian_List
Vlach
Latin
Umbrian
Oscan
Breton_SE
Breton_ST
Breton_List
Cornish
Welsh_N
Welsh_C
Irish_A
Scots_Gaelic
Old_Irish

8000 7000 6000 5000 4000 3000 2000 1000 0

SOURCE: Redrawn from Bouckaert et al. (2012), with correction ("Corrections and Clarifications," 2013).
NOTE: X axis labels indicate years before present. Contributed by Quentin Atkinson.

16

Some evidence points to language divergence being not just a matter of isolated speaker populations gradually mutating at random, but also driven by social processes that favor isolating speaker populations to smaller group sizes that can more readily form cooperative units (Atkinson et al., 2008). This process of social isolation for the adaptive function of forming effective cooperative units is thought, in part, to drive the treelike divergence of religious denominations (see Figure 2.4) (Matthews, 2012). Whether in language or religion, one of the characteristic signatures for the social signaling function is punctuated change in characteristics specifically at schism points, as without social functions, characteristics typically display a more clocklike pattern of change across all branches.

Figure 2.4. Christian Phylogenetic Tree Depicting Denominational Schisms Known from the Historical Record

Roman Catholic

Baptist

Anglican

Methodist

Presbyterian

Lutheran

Orthodox

Coptic

1784 - Christmas Conference in Baltimore founding Methodist Episcopal Church

1612 - Helwys founds the General Baptist church in London

1534 - English Parliament Act of Supremacy

1520 - Luther's three reform treatises

1541 - Calvin returns to Geneva

Year (Anno Domini)

— 1800

— 1600

— 1400

— 1200

1054 - Cardinal Humbert and Patriarch Cerularius excommunicate one another

— 1000

— 800

— 600

454 - Council of Chalcedon

SOURCE: Redrawn from Matthews, 2012.

Phylogenetic methods provide a rich analytic toolkit that is applicable to cultural systems with sufficient time durability. Cultural phylogenetic approaches have been used to reconstruct how people dispersed into new environments (Buchanan and Collard, 2007), to reconstruct stories of shipwreck and find a missing vessel (Dunn and Kirsner, 2011), to assess which types of political change and state failure are more probable (Currie et al., 2010; Matthews et al., 2016), and to study which religious groups are more likely to adopt violent ideology (Matthews et al., 2013).

Summary

We briefly but systematically expanded on our definition of culture as the shared information that results from social interaction. This chapter reviewed the research on social learning mechanisms and social learning strategies and how these lead to emergent properties of cultural dynamics. We hope to have conveyed that culture, while complex, is not a completely vague concept of which anthropologists and other researchers have gained no understanding. There is a strong foundation of research establishing different forms of social learning mechanisms. Social learning strategies have been explored through models (Franz and Matthews, 2010; Nunn et al., 2009), empirical research (Henrich and Gil-White, 2002), and competitive computer tournaments (Fogarty, Rendell, and Laland, 2012). The emergent patterns of culture are the subject of much research in cultural consensus theory (Romney, Weller, and Batchelder, 1986) and cross-cultural comparative research (Atkinson, 2011; Currie et al., 2010).

It is important to note that our understanding of cultural processes does not necessarily conflict with rational actor models for behavior. Microeconomic theory espouses that people act rationally to maximize a set of utility functions. Traditional economic theory assumes that individuals have full and accurate information when making decisions, and it is this assumption that provides the link from microeconomic models to macroeconomic patterns, such as the balancing of prices based on supply and demand.

Our theoretical framework modifies this only by proposing that individuals do not have full or accurate information. Instead, individuals acquire much of their information socially and from within their particular religious and ethnic communities. Some of this information may be inaccurate or incomplete. Within the context of what individuals know or believe, however, they may still be rational actors in the economic sense.

Having established a conceptual model of what culture is and how it works, the remaining chapters focus on operational methods for the study of culture. These methods can find many applications in policy analysis, but it was necessary to first explain the culture concept because anthropological methods are not merely a grab-bag of mixed methodologies for policy; they have been shaped by the theoretical notion of culture. The relationship of these methods to cultural theory is similar to that of econometric techniques to economic theory; indeed, econometrics could be viewed as just a particular instantiation of statistics, but one that is shaped by qualitative

theory about maximal utility decisions under scarcity. Similarly, the methods described in what follows hang together because they all were developed as tools to understand how socially transmitted beliefs and behavior (culture) can be described and predicted.

In the next chapter, we explain how cultural consensus analysis (CCA) is used—from understanding individuals as data points to establishing emergent cultural constructs. CCA can establish where each individual lies in terms of cultural similarity or dissimilarity to the other individuals measured.

3. Methods to Understand Culture: Cultural Consensus Analysis

The theory of culture presented in Chapter Two is in terms of conceptual variables—such that the socially transmitted information itself is rarely observed directly (Bernard, 1994, pp. 25–32). Culture used as a conceptual variable is a theory, which cannot be tested directly. Although theories cannot be tested directly, they imply hypotheses, which are predictions of what a researcher expects to find given the veracity of the underlying conceptual theory. Hypotheses, unlike theories, can be tested but only when the conceptual variables from the theory have been given operational definitions. For the study of culture, we can observe cross-sectional or longitudinal patterns in beliefs and behaviors across individuals and groups, but social transmission of these is necessarily an inference.

If cognitive models are shared among members of a group, implying a cultural pattern of thought and behavior, this sharing needs to be empirically demonstrated in order to be used in a test of a hypothesis. Members of social groups often have similar cognitive models because of similar life experience and tend to agree with one another about the meaning of certain domains of knowledge and behavior. However, within social groups, there are subgroups of members who have slightly or not so slightly different life experiences, which causes disagreement. Age, gender, and socioeconomic status are examples of characteristics that can influence members of the same language group, religion, or geographic area to have varying cultural models of particular domains. Sometimes this varying life experience results in intra-cultural variation in which there is variation within an overall level of agreement. Sometimes it may result in inter-cultural variation in which there are groups within the population that have high internal levels of agreement but do not agree significantly with other groups within the same social group. And sometimes there may be very low levels of agreement among all members of the population (Handwerker, 2001). Thus, to validate claims of agreement on cognitive models enough to classify them as cultural, levels of agreement must be measured, and cultures must be demonstrated empirically. Since the mid-1980s, cognitive anthropologists have developed ethnographic techniques to collect field data, which can be analyzed statistically to demonstrate this agreement empirically and can be used to test the internal validity of theoretical claims about culture.

The methodology of CCA began as an effort to limit the effect of an ethnographer's personal biases on the final product of ethnography and to determine an objective measure of informants' cultural knowledge (Romney, Weller, and Batchelder, 1986, p. 314). Noting well-known controversies over the conclusions of famous ethnographies as well as the tendency of informants to disagree with each other, Romney, Weller, and Batchelder developed a formal method that would determine which informants were more culturally knowledgeable than others. Experiments showed that informants who knew the correct responses to questions tended to

agree with each other more than informants who did not know the answers (Boster, 1985). Noting that, mathematically, the reverse of this—that those who agreed with each other would be more likely to have knowledge about some unspecified truth—would also be true, Romney, Weller, and Batchelder's formal model provides a means of determining who knows more than whom as well as the answers to the questions themselves.

CCA uses the multivariate technique of exploratory factor analysis to develop these measures. Factor analysis enables the indirect measure of conceptual variables that cannot be measured directly. These unobserved "latent variables" (Bollen, 2002) can be measured indirectly through one or more directly measured variables that are affected by the latent variable. Factor analysis aims to explain correlations among a set of observed variables that do not have direct causal relationships with each other but that are spuriously associated through their common dependence on one or more other unobserved variables (Bartholomew et al., 2002). The typical goal of factor analysis is to reduce the number of variables by finding which variables appear to measure the same underlying thing and replace them in analysis with a smaller number of latent variables (Neter et al., 1996). Interpretation of factor analysis requires examination of several metrics, including (1) factor loadings, which are measures of the correlation between each directly measured item with the underlying factor, and (2) factor scores, which aggregate the variable loadings of a particular observation on a factor to place or rank the observation on a specific factor (DiStefano, Zhu, and Mindrila, 2009).

Factor analysis serves different objectives in CCA than its typical use in identifying patterns among variables. In CCA, the latent variable of interest is culture. Culture cannot be measured directly but has influence over individuals in a group. This produces a pattern of similar thoughts, behaviors, and emotions among the members of the group on a particular domain. Therefore, the latent variable of culture is not a summary of patterns among variables but rather of patterns among individuals. To identify these patterns, CCA transposes the traditional profile data matrix so that the analysis is focused on individuals by items rather than items by individuals. Rather than factor analysis identifying latent factors to explain correlations in a variable-by-variable correlation matrix, CCA identifies latent patterns in an informant-by-informant agreement matrix. This approach (McKeown and Thomas, 2013) operationalizes culture as a latent variable. Culture itself is not directly measured but is measured indirectly on a set of informants, and this provides an indirect measure of culture. A factor analysis of a transposed profile matrix requires reversal of typical interpretation of factor analysis metrics. Factor loadings produced with CCA, for example, are correlations between informants and the underlying cultural pattern identified with factor analysis (the factors). They are no longer interpreted as correlations between variables and factors. Factor scores, on the other hand, are interpreted for variables rather than individuals. They are aggregations of the factor loadings across the set of informants for each variable. Ranking variables from low to high on their factor scores enables interpretation of the primary drivers of the pattern of factor loadings.

By analyzing the variable by informant matrix, CCA effectively estimates how well informants agree with the main variations exhibited among the variables (Romney, Weller, and Batchelder, 1986; Weller and Romney, 1988, p. 75). CCA can be used when certain assumptions are met, including the assumptions that there is a specific correct answer for each question asked of an informant, each set of answers given by an informant is independent of other informants' answers, and that the questions are drawn from one domain of knowledge (Handwerker and Borgatti, 1998; Romney, 1999, p. S107). Given these assumptions, the original formulation of CCA would analyze loadings of individuals on the single factor accounting for the greatest variation. Typically, the variables are responses to survey items or coded themes from interview data.

This original formulation of CCA assumes that the researcher knows ahead of time that all the separate variables are from one domain of knowledge (i.e., they are from the same underlying cultural construct). Although variables for CCA can come from any source, such as existing scales, CCA is best used as a mixed-methods companion to exploratory data collection and analysis, such as ethnography or cultural domain analysis (Weller, 2007; D'Andrade, 1995; Weller and Romney, 1988). Systematic, exploratory data collection assures the identification of a coherent set of relevant elements of a domain for a population of interest. This is why applying factor analysis to the transpose of the individual-by-variable matrix makes sense: The items are preselected for their potential to demonstrate cultural salience in the population. This contrasts to the typical approach from psychometrics and most other social sciences to use factor analysis to identify which variables "go together." In the classic formulation of CCA, the researcher already knows that the variables they have measured go together in a single construct, and so all that is being assessed is how well individuals "go together" with respect to their values on the variables.

Handwerker has developed a modification of the CCA model to measure cultural differences when some of the assumptions mentioned earlier are not met. His modification is to use principal component analysis (PCA) instead of factor analysis when analyzing the intersection of agreement among informants (Handwerker, 2001, p. 185; Handwerker, 2002, p. 112). Handwerker's method still performs the statistics on the individual-by-individual agreement matrix, the transpose of the usual approach from psychometrics. PCA and factor analysis are very similar procedures: They both linearly transform the original variables into a smaller set of uncorrelated variables called *factors*. PCA differs from factor analysis in that it uses the maximum variance rather than the common variance (Bartholomew et al., 2002, pp. 167–168; Dunteman, 1989, pp. 7–9). Using the maximum variance is more important in cases where an investigator wishes to examine the total variance of the agreement among informants to determine whether there exists one, multiple, or no cultures of agreement on a particular domain. The factor analysis method used by CCA is more appropriate when the goal is to determine cultural competence on a single predefined culture. Instead of looking for the correct answers, Handwerker's method looks for the possibility of subcultural agreement on different sets of "correct" answers.

To determine subgroupings of agreement, Handwerker proposes using the factor loadings on the first two unrotated factors of a PCA in a cluster analysis (Dunteman, 1989, p. 78–79; Handwerker, 2001, p. 88; Handwerker, 2002, p. 115). A cluster analysis can be performed on the loadings of the first two factors if they account for a substantial amount of the total variance. If there is one culture whereby each informant is linked to one another through a shared agreement on a domain, the first factor will represent at least 50 percent of the variance, the difference between the eigenvalues of the two factors will be large (the first factor is three times the second), and there will be negligible negative loading on the first factor (Handwerker, 2001, p. 186). If there is not one culture, plotting the factor loadings for the informants on the two factors against each other will show patterns of agreement. These patterns can be used as clues for finding multiple subcultures as well as multiple cultures within a social group. In theory, the original sample could then be subsetted into the empirically determined subcultures, and the classical CCA would be appropriate within each of these samples.

This method provides a means of empirically justifying the claim that a group of people belongs to a particular culture. In other words, it provides an empirical means of assessing construct validity (Handwerker, 2001, pp. 183–217; Handwerker, 2002). If a group of informants is selected based on their ethnicity or national origin and there is a particular domain of knowledge, attitude, or behavior that is the focus of study, this method provides a means of determining empirically if in fact there is one culture for this group on this domain. Another way researchers have phrased this problem is that Handwerker's and related methods determine whether there is more than one answer key for the variables; as in, there is more than one set of underlying socially learned and shared notions about the correct answers to these items. If there is not a big difference between the first two factors, this method will help determine what kinds of culturally relevant groups there are. Future groupings of informants may reveal that other life experiences (such as age or gender) are more important factors in determining cultural groups for the theoretical domain in question. Thus, once an operational definition is given to a culturally based construct variable, empirical evidence can determine how well this construct holds together. This method can be used to evaluate the consistency of findings with future research that uses the same operational definitions of theoretical concepts.

Factor analytic techniques are important substitutes for classical statistical methods because of several properties of cultural data. First, cultural data are by nature multidimensional (Handwerker, 2001, p. 190). The cognitive models that form the basis for cultural models contain a multitude of elements, their relationship with each other, and how all the elements function for a particular domain (D'Andrade, 1995). To understand the whole of a cultural model, measurement must be made on the individual elements that make up the model. This requires analysis of answers given to structured questions from survey questionnaires, which form the basis for comparison of agreement among informants (Weller, 1998). The responses to the questions, or the "verbal production," are not the cultural model. Rather, consistency in

answering the questions gives evidence that the informants are all drawing from the same shared cultural model to produce their responses (Dressler and Bindon, 2000, p. 247).

Other Potential Solutions for When the Cultural Constructs Are Not Known a Priori

Handwerker's method is one solution to the assumption of the original CCA technique that there is only one answer key—that is, one set of socially learned ideas about the correct responses to the individual variables measured by the researcher.

His method does not, however, deal with the potential that the researcher does not necessarily know ahead of time which variables go together as part of a single latent cultural body of knowledge. Evolutionary anthropologists have approached this other assumption of CCA by applying PCA in the more typical manner that finds agreement among variables, with the notion that correlated variables reflect some common body of cultural knowledge. Having thus obtained a set of components that summarize the original variable set, these researchers then proceeded to analyze individual data points scores via cluster analysis in a manner very similar to Handwerker's application of cluster analysis to person loadings. An example of this approach from evolutionary anthropology is the reanalysis of artifact data from New Guinea performed by Shennan and Collard (2005) on the Welsch, Terrell, and Nadolski (1992) data set.

We propose that there may be utility in further combining these approaches by first employing a typical PCA approach to establish which variables belong to the same cultural constructs, then performing the classic CCA on the transpose separately for each construct. These steps could be performed on the same or separate subsets of data, depending on how exploratory or confirmatory a researcher intended to interpret the findings. The advantage of this modification is that it leads to loadings from 0 to 1 that reflect how much each individual agrees with their peers about a particular cultural domain. Obtaining an intuitive score for an individual's accuracy of cultural knowledge was an original goal of the classic CCA approach, and not provided by the nonintuitive principal component scores, whose range of variation often is highly arbitrary.

We can work through what this process would look like with a hypothetical problem and data that will allow us to quickly see the full process. Suppose we collected data on whether individuals believed one or more of the numerous conspiracy theories about vaccines. We could collect such data through surveys, monitoring of social media, analysis of documents such as websites, and many other techniques. Survey methods are still the most commonly used for this type of research, but methods involving social media are definitely on the rise. Once we have collected the data, they can be arranged into a person-by-belief table, as seen in Table 3.1.

Table 3.1. Hypothetical Individuals by Vaccine Beliefs

	Vaccines Cause Autism	Vaccines Do Not Stop Diseases	The FDA Conceals Data on Vaccines and Autism	Vaccines Cause ADHD	Vaccines Are Safe	Vaccines Prevent Flu	Vaccines Kill Bacteria
Individual 1	1	1	1	0	0	0	0
Individual 2	0	0	0	1	1	1	1
Individual 3	1	0	0	1	0	0	1

A matrix such as this could then be analyzed via PCA to determine which variables are related to one another (covary). These axes of covariation can be viewed as operational measurements of an underlying cultural construct. Sometimes researchers call this a "latent" variable to indicate that the putative cultural variable is measured indirectly via many direct observations that are related to it.

Following the evolutionary researchers, one could plot where along each of the axes each individual lies. Clusters of individuals in the PCA space are typically interpreted as subcultures (Figure 3.1). In the hypothesized PCA in Figure 3.1, there are two main components, with autism-related conspiracies loading heavily on the first component and all other conspiracies loading heavily on the second. This would indicate that autism-related conspiracies are being transmitted among social learners independently of the nonautism conspiracies. Individuals are located along each of these two axes, but clusters of points also can be identified. A cluster of individuals that score highly for both types of conspiracies appears to be present (circled in dashes).

Figure 3.1. A Hypothetical Principal Component Analysis of Vaccine Conspiracy Beliefs
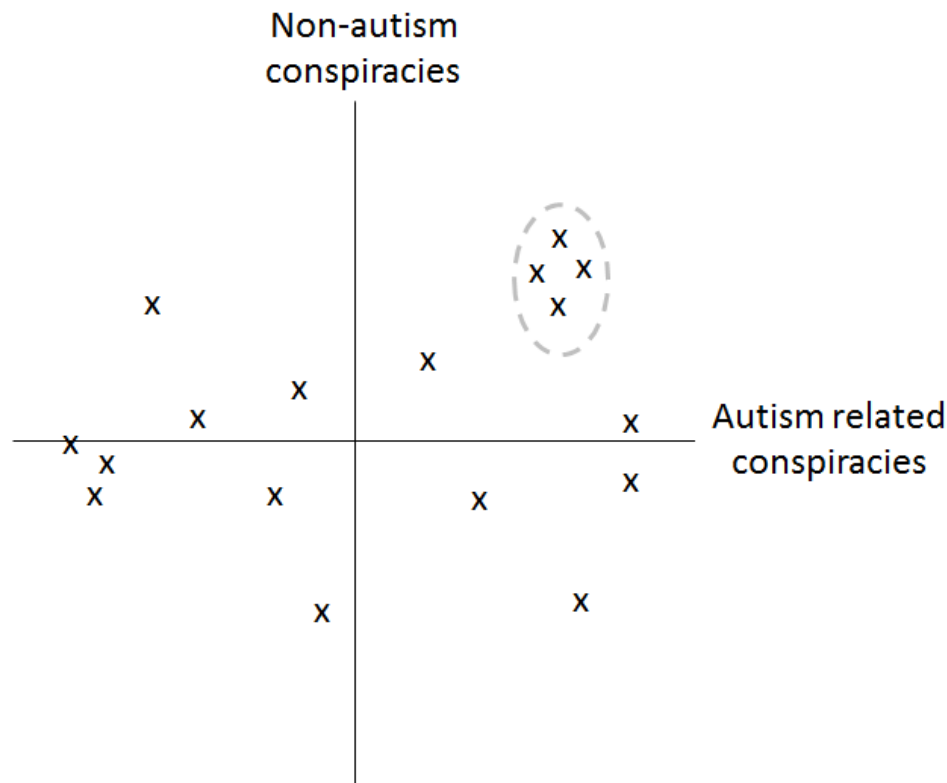
Where the first principal component explains nearly all the variation, the interpretation follows Rommey, Weller, and Batchelder (1986) that the data-collection method effectively measured only one aspect of cultural variation and that individuals are more or less "informed" along this axis. The measure of informed-ness proposed by Romney, Weller, and Batchelder (1986) was to reverse the usual application of PCA and to construct factor loadings from a belief by individual matrix (Table 3.2). This effectively measures the agreement (factor loading) of each individual with main axis of variation in the data set, in contrast to the typical psychometric approach to PCA that measures the level of agreement (factor loading) of each variable with the main axis.

Table 3.2. Belief by Individual Table Used for Cultural Consensus Analysis

	Individual 1	Individual 2	Individual 3
Vaccines cause autism	1	0	1
Vaccines do not stop diseases	1	0	0
The FDA conceals data on vaccines and autism	1	0	0
Vaccines cause ADHD	0	1	1
Vaccines are safe	0	1	0
Vaccines prevent flu	0	1	0
Vaccines kill bacteria	0	1	1

Another way to think about this procedure is that it is similar to a cluster analysis but with a continuous rather than nominal outcome variable. In most cluster analyses, the procedure attempts to summarize variation across all the input variables by putting individuals into groups that maximize variation between the groups and minimize variation within the groups (Kaufman and Rousseau, 1990). By reversing the usual application of PCA, CCA constructs a single continuous measure that, similar to cluster analysis, tries to put individuals close together who share similar beliefs, with higher factor loadings reflecting individuals who are closest to the general pattern in the whole data set.

One could also use the principal component scores for each individual, but an advantage of the classic CCA approach is that it creates factor loadings going up to 1 that intuitively indicate the degree to which an individual is "knowledgeable" about that cultural domain. Another difference in the classic CCA approach compared with using the principal component scores on the first principal component (PC1) is that the classic CCA factor loadings of individuals implicitly weight each variable equally in its contribution for the loading of each individual. This is not true for the principal component scores of individuals on PC1, as individuals will receive proportionately higher or lower scores for variables that load more strongly on the PC1. Variables that do not load on PC1 do not contribute to the scores of individuals on that principal component, but all variables contribute equally to the factor loadings on the transposed CCA technique. This is intuitive if one considers that all individuals contribute equally to the variable

loadings in the typical PCA technique performed in psychometrics. Whether this difference is an advantage or not is up to interpretation. If, as assumed by Romney, Weller, and Batchelder (1986), all the variables equally reflect a single underlying cultural domain, then the classic CCA approach makes good sense. If the researcher is less certain that the initial qualitative work strongly justifies this assumption, then using principal component scores may be advisable as a more data-driven approach.

We note that some readers may rightly be concerned that the number of individuals may exceed the number of variables, thus preventing the form of analysis that extracts loadings of individuals because some statistical implementations of PCA require there to be more rows than columns in the original data (e.g., the princomp function in R). When there are more individuals than variables, the classic CCA on the matrix transpose will result in a variable-by-individual matrix with more columns than rows in cases in which the number of sampled individuals exceeds the number of variables measured (i.e., the dimension). In typical anthropological data sets, it is actually the number of variables (e.g., survey questions, coded themes) that exceed the number of respondents such that the classic CCA approach of working on the matrix transpose conforms to the usual psychometric rule of having a matrix with more rows than columns.

A more important point, however, is that PCA does not require more individuals than variables when it is calculated with singular value decomposition (SVD) as opposed to Eigenvector decomposition (Shen et al., 2016; Yata and Aoshima, 2010) for implementation (see the prcomp function in R). These statistical procedures in fact produce equivalent results, although SVD is more numerically accurate for some data and so preferable on those grounds. At a maximum, as many eigenvectors are interpretable only as there are data points, and the entire purpose of PCA usually is to interpret only a handful of the full set of eigenvectors. With slight statistical adjustments, PCA is being used effectively for new data types in which the number of variables measured routinely is more than an order of magnitude greater than the number of data points. Such data include acoustic and genetic data in which a researcher can easily measure thousands of variables on mere hundreds of individuals (Fokoue and Ma, 2013; Paschou et al., 2007). In the case of most cultural data, the number of variables often approximates the sample size or is slightly larger, and under such conditions PCA is shown to be generally statistically consistent (Shen et al., 2016).

It is often useful to compare the results of a CCA with other variables of interest that did not go into CCA. This can be done either to find causal explanations for individuals espousing the cultural beliefs they have or to assess how these cultural beliefs cause other decisions. Of course, the usual concerns about correlation and causation apply. For example, we might be interested in whether knowing a child with autism affects where someone lies on the autism vaccine-conspiracy axis. This could be tested through a regression model, with individuals' scores on the PC1 of our vaccine beliefs analysis as the outcome variable. It would seem fairly certain that any statistical association results from knowing a child with autism increasing one's adoption of autism vaccine conspiracies, since we likely can assume that adopting autism vaccine

conspiracies does not cause children to have autism. We might also be interested in how a person's position in the cultural space affects other decisions. For example, if we had subsequent information about whether the sampled individuals vaccinated their children, then we could use regression models to test whether an individual's position on the cultural axes is associated with the vaccination decision. Perhaps some of the association is causal, but the vaccination decision and the cultural measure for an individual might both be caused by a third unmeasured variable.

Empirical Examples of Cultural Consensus Analysis

This section provides examples of CCA focused on different topics. Each example uses different methods to provide analytic leverage on the content, form, and distribution of cultural beliefs and behaviors within a population.

Exploratory Methods for Cultural Domain Analysis: Culture of HIV Beliefs in the Deep South

It is possible to use existing scales with a CCA, but there is no guarantee that these scales are measuring constructs that are locally meaningful to the population. To ground cultural consensus items in local understandings, a variety of techniques (often referred to collectively as *cultural domain analysis*) can be used to allow ethnographic informants to map the cultural domain of study (Borgatti, 1999). This section illustrates one such technique—freelisting (Ryan, Nolan, and Yoder, 2000)—and shows how CCA analytic techniques can be used to help make sense of freelisting results.

HIV in the American South

A new human immunodeficiency virus (HIV) epidemic has emerged in the American South. Not only are prevalence and new infection rates high, survival rates are also lower than in the rest of the United States, which raises concerns not only about risk of transmission, but also puts a spotlight on the treatment environment.

Similar to the way in which Nisbett and others (Cohen and Nisbett, 1994; Cohen, Vandello, and Rantilla, 1998; Grosjean, 2014) theorize that a "culture of honor" (pressure for men to save face to preserve social reputation) is partially responsible for higher rates of violence in the American South, culture is often cited as a reason for high HIV diagnosis rates and lower survival rates in the region (Lichtenstein, 2005; Lichtenstein, Hook, and Sharma, 2005; Elmore, 2006). The default explanatory frame is to assume that the entire region has particular cultural characteristics, including high levels of religiosity, enforcement of traditional gender roles, and stigma around HIV. However, the degree to which (1) the American South espouses these characteristics in a unique way or to greater degree than other regions and (2) these characteristics are to blame for higher HIV morbidity and mortality are empirical questions.

One place to look for evidence of culture at work is by analyzing local variability. Using AIDSVu (undated), an interactive online map illustrating the prevalence of HIV in the United States, we began to look at HIV prevalence in specific cities and counties in Mississippi. We identified two comparable counties—Jackson and nearby Hattiesburg—that have similar population density and similar demographic profiles, but significantly disproportionate HIV rates. Despite having lower poverty and a lower percentage of uninsured residents, Jackson County's HIV rate for both blacks and whites is five times the HIV rate in Hattiesburg (439 versus 71 per 100,000).

We used two methods to collect pilot data from Hattiesburg and Jackson on cultural/explanatory models of HIV/acquired immune deficiency syndrome (AIDS). The first data set was collected using a convenience sample of 25 waiting-room patients from a clinic in Jackson and 26 undergraduates from a state school in Hattiesburg. We also used freelists, a standard technique for exploring a domain where you have some idea about dimensions, but little idea regarding the local cultural perceptions of the content for each dimension. The freelist questionnaire covered the following areas:

- *Cause*: "There are many ideas and opinions about how a person gets HIV. Please list as many behaviors, actions, situations, or other things that lead a person to get HIV."
- *Prevention*: "Now, how about ways a person can avoid getting HIV? Please list as many actions, behaviors, or other things that someone can do to avoid getting HIV."
- *Diagnosis*: "We've talked about what HIV/AIDS is and how you can get it or avoid getting it. How does someone know they have it? Please list all the ways someone could tell they have HIV/AIDS, including signs and symptoms."
- *Disease course*: "After someone gets HIV, what happens to them? Please list all of the outcomes and changes that might occur."
- *Treatment*: "Let's say someone knows they have HIV. What can they do now? Please list all of the actions, procedures, or other things someone can do to treat HIV or manage the effects of it."

Using such tools as Anthropac (Borgatti, 1996) and UCINET (Borgatti, Everett, and Freeman, 2002), we examined the frequency of responses within and across questions to group differences and similarities into categories. We did this iteratively in teams until we reached a manageable level of data reduction, still preserving individual variability but combining items that are simply different wordings of the same response (see Figure 3.2).

Figure 3.2. Screen Plot of Responses to the "Cause" or to the Question "How Do You Get HIV?"
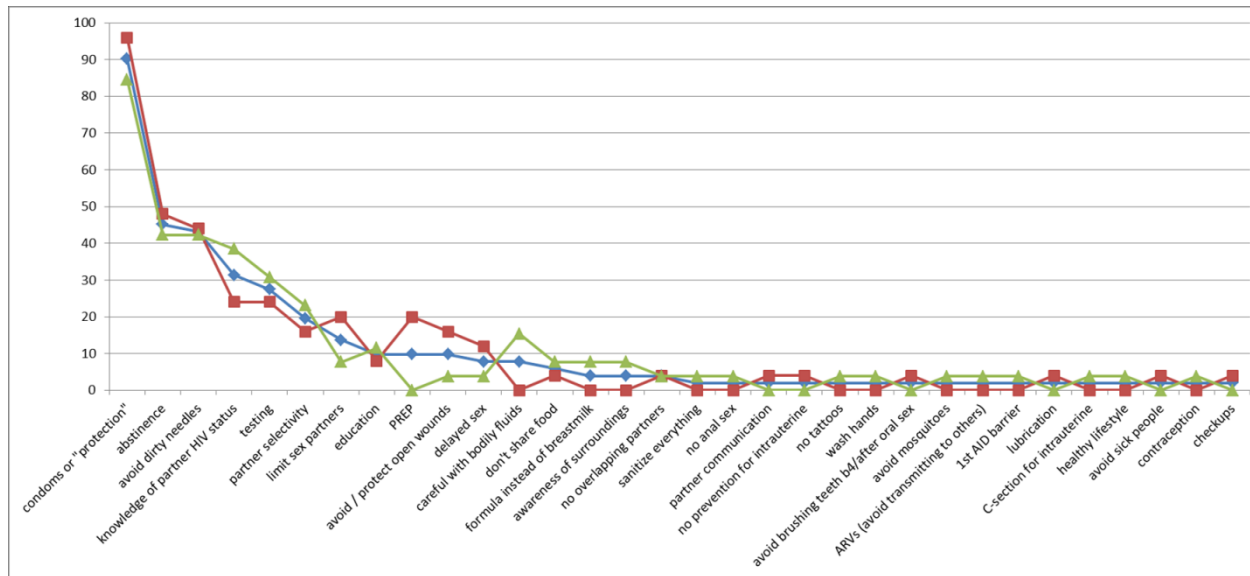
NOTE: Red line represents Jackson respondents and green line Hattiesburg; blue line is the average of the two.

We then performed CCA on all items mentioned by 10 percent or more of respondents (five or more out of 51), creating a variable for each of these items indicating whether each respondent mentioned or did not mention the item. C is not an ideal fit for freelist data, as CCA assumes that all respondents have had a chance to respond to each item. However, CCA with freelist data can give a rough sense of items that are "top of mind" within and across groups—in this case, within and across groups of participants from Jackson and Hattiesburg.

To conduct this analysis, we transposed the matrix of items by respondents into respondents-by-items matrix, conducted a PCA, and examined the Eigenvector value ratios of the first to second principal components. Although CCA analysis "rules of thumb" were not met for a single culture involving the full set of 51 respondents (ratio of Eigenvalue between first and second factors = 2.6), this rule of thumb was met for Hattiesburg respondents separately (Eigenvalue ratio = 3.1) and was almost met for Jackson respondents separately (Eigenvalue ratio = 2.9).

Next, we examined the items endorsed more frequently by individuals with high positive principal component scores for the combined sample and for the Jackson and Hattiesburg subsamples independently. Our combined freelist/CCA analysis indicated a core set of shared items that include

- sex as transmission pathway or avoiding sexual contact as prevention
- dirty needles and open wounds as transmission pathways
- weight change as expected sign of HIV
- professional medical care and adherence to HIV medication regime as essential for successful treatment.

32

We then examined the frequency with which individuals with high principal component scores endorsed certain items for the Jackson and Hattiesburg samples separately. Generally, "culturally knowledgeable" Jackson respondents were more focused on the social and mental health effects of HIV, such as social isolation, depression, and needing counseling following an HIV diagnosis. Meanwhile, Hattiesburg respondents with high principal component scores were more focused on the risk of transmitting HIV to sex partners (including intrauterine transmission) and the importance of the formal involvement of the medical system for diagnosis. While it is too early to make much of these differences or to try to tie them to differential rates of infection, this analysis provides a starting point for further explorations of different cultural perspectives on HIV in hotspots such as the South.

Culture of Childbearing Motivation Across Three Countries

We conducted a CCA of childbearing motivation, using questionnaire data collected from respondents in Honduras, the United States, and Iran. The items come from the Childbearing Questionnaire (CBQ), which was developed in English based on a longitudinal study of fertility intentions of more than 400 married couples living in the Silicon Valley area of California in the United States in 1989 (Miller, 1995). The couples were predominantly European American and middle class. The CBQ aims to measure overall childbearing motivation by measuring separate dimensions of positive and negative motivation in separate subscales. These scales are designed to measure the respondents' explicit self-rating of the desirability or undesirability of a comprehensive list of positive and negative childbearing experiences. Although the CBQ was developed and tested with a fairly homogenous group of participants, the scale was intended to be used in different populations or cultural groups. The CBQ was later translated and adapted for measuring childbearing motivation in two different languages for use in two different countries in 2000–2001: Farsi for use in Iran (Pezeshki, Zeighami, and Miller, 2005) and Spanish for use in Honduras (Kennedy, 2005; Kennedy, 2002). At the time of data collection in each of these countries, the national fertility rates were experiencing contrasting fertility rates and rates of change: Hondurans was experiencing steady decline but continued high fertility, the United States was experiencing decades-long sustained low fertility, and Iran was experiencing a recent drastic fertility transition from high to low (Figure 3.3).

Figure 3.3. Comparison of Fertility Rates for Honduras, Iran, and the United States
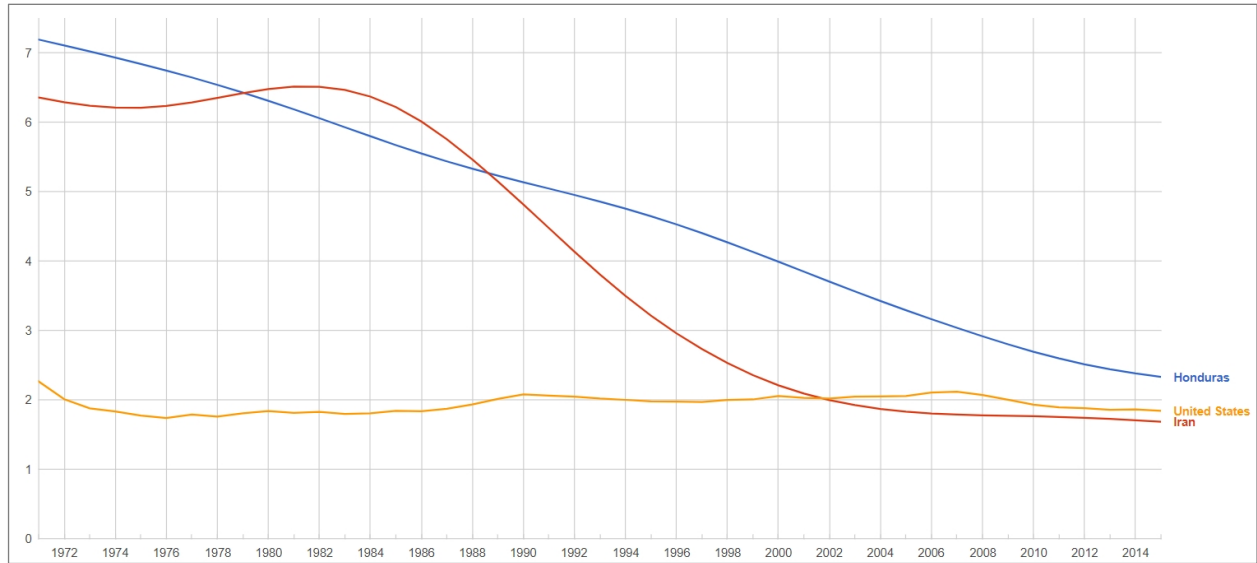
NOTE: Data from the World Bank. Graph constructed using Google Public Data. The three countries show contrasting trajectories of fertility defined as the number of average births per woman during administration of the CBQ: 1989 for respondents in the United States, and 2001 for Honduran and Iranian respondents.

The literature on demographic transitions, in particular the fertility decline literature (Handwerker, 1986), often cites culture as a change agent in triggering reductions in births as a population shifts from valuing large families (i.e., exhibits a "pro-natal" culture [Harris and Ross, 1987; Mason, 2001; Mason and Taj, 1987; Sangree, 1997]) to an anti-natal culture (Harris and Ross, 1987). Given this theory of culture and fertility, a testable hypothesis would be that populations at different points on the fertility-decline trajectory would exhibit different cultural traits. Therefore, use of the CBQ to measure positive and negative childbearing motivation in each of these three countries during these periods provides data necessary to conduct a test of this hypothesis.

For this test, we conducted a CCA of the CBQ data from each of these three samples. The CBQ data collected from the United States, Honduras, and Iranian populations were combined into one data set. The U.S. sample included 810 randomly selected individuals, the Iranian cross-sectional sample included 600 respondents (300 couples visiting the Shiraz Health Clinic in Fars Province in southern Iran for premarital screening), and 200 Honduran respondents sampled cross-sectionally (based on a random-sample stratified on gender and rural and urban areas around the town of Catacamas in the central state of Olancho). The analytic sample included 47 total items, 27 positive aspects of childbearing, and 20 negative aspects of childbearing. For positive items, respondents were asked to rate how desirable each item was and, for negative items, how undesirable. The response options across the three data sets were standardized so that a rating of "not at all" (desirable/undesirable) was given a value of 0, a rating of "somewhat" was given a value of 1, and "very" was given a value of 2.

Table 3.3. Desirability/Undesirability of Childbearing Questionnaire Items for Three Countries for Primary Factor

Item text	Average Rating			Factor Score
	Honduras	Iran	United States	
Holding and cuddling a baby[a]	1.71	1.85	1.58	1.21
Giving my spouse the satisfaction of being a parent[a]	1.54	1.83	1.50	1.12
Sharing child raising with my spouse[a]	1.68	1.91	1.75	1.12
Burdening our family finances with a child[b]	0.71	0.48	0.67	−1.48
Straining our marriage with a baby[b]	0.77	0.31	0.97	−1.58
Being kept from my (having my wife being kept from her) career or job by a baby[b]	0.59	0.37	0.87	−1.64

NOTE: Table includes the items with the three most extreme factor scores on Factor 1.
[a] Positively worded item: respondents asked how "desirable" item was for them.
[b] Negatively worded item: respondents asked how "undesirable" item was for them.

A PCA was conducted on the full CBQ data using SAS's "proc factor" procedure with no rotation. The data were first transposed to produce a variable-by-informant matrix rather than an informant-by-variable matrix (Handwerker, 2001, p. 190). This allows for the creation of an informant-by-informant similarity matrix (Romney, Weller, and Batchelder, 1986, p. 322; Handwerker, 2002, p. 113). Figure 3.4 presents a scatter plot of the informants' loadings on the first two factors of all of the CBQ data. The graph is a representation of the similarities and dissimilarities among the respondents. Each point on the graph represents an individual that responded to the questionnaire. Factor 1, the horizontal axis, represents the broadest range of

commonalities among the respondents (Handwerker, 2001, p. 190). Factor 2, the vertical axis, represents the second-broadest range of commonalities among the respondents. The closer the points are to one another represents the closer the totality of the individuals' CBQ responses are, at least with regard to the first two factors. Table 3.4 provides the item text and factor scores for the three items with the most extreme positive and negative factor scores on the primary PCA factor, Factor 1. These items are identified as coming from either the positive childbearing motivation or negative childbearing motivation subscales. The table also provides the average item rating for respondents in each of the countries. Table 3.4 provides the same information (item text, positive/negative designation, factor scores, average scores) for the secondary PCA factor, Factor 2.

Table 3.4 Desirability/Undesirability of Childbearing Questionnaire Items for Three Countries for Secondary Factor

	Average Rating			Factor Score
Item text	Honduras	Iran	United States	
Having an unhappy and poorly adjusted child[a]	1.11	1.18	1.63	2.03
Guiding and teaching my child[a]	1.56	1.96	1.80	1.89
Having a baby who is a burden to my spouse[b]	0.79	0.23	1.50	1.86
Feeling more complete as a woman/man through my baby[1]	1.37	1.78	0.88	−1.15
Spending time and energy involved in childcare[b]	0.91	0.24	0.52	−1.54
Fulfilling my religious feelings about family life[a]	1.54	1.66	0.61	−2.03

NOTE: Table includes the items with the three most extreme scores on Factor 2.
[1] Positively worded item: Respondents asked how "desirable" item was for them.
[2] Negatively worded item: Respondents asked how "undesirable" item was for them.

The results of this CCA using PCA show that these three groups do not represent a single cultural group. None of the rules of thumb support cultural consensus for the combined sample. The eigenvalue of the first component is less than three times the second (2.55), and the first

component explains only 25 percent of the variance (less than the 50 percent threshold). Figure 3.4 provides a visual contrast between the groups: The U.S. sample is spread throughout the Component 1 dimension and scores high on Component 2. In contrast, the Honduran sample clusters tightly on the high end of Component 1, with some random clustering above and below the 0 axis for Component 2. The Iranian sample appears midway between the contrasting Honduran and U.S. sample distribution. These contrasts meet our expectation of cultural differences between the respondents from these three different populations. However, there was not necessarily a sharp contrast between populations on the main factor. We labeled the main component "pro-natality" versus "anti-natality" after examining the factor loadings of the items. Each of the items that loaded positively on Component 1 was a positive item, and each of the negative items loaded negatively. Therefore, this component is a clear contrast between endorsing positive and negative aspects of having children. The Honduran sample clearly clusters toward the pro-natality pole of the first component, which met our expectation of a population with high fertility. However, although the U.S. sample was drawn during a time of sustained low fertility, the sample does not cluster toward the anti-natality pole of the first component. The sample includes a distribution of scores across the full range of Component 1 but with a higher concentration toward the prenatally pole. Therefore, although the U.S. sample clearly exhibits more anti-natality elements, it cannot be characterized as a fully anti-natal culture. When analyzing the Honduran sample alone, there is evidence of one shared culture of pro-natality among Hondurans: The ratio of the first to second component eigenvalue is 8.77, and the variance explained is 61 percent. Therefore, we conclude that there is evidence of pro-natality culture associated with high fertility. However, we do not find evidence of a consensus around anti-natality associated with low fertility. The cultural pattern of the low-fertility population appears to exhibit disagreement about the positive and negative aspects of childbearing.

Examining the second component factor scores is useful to better understand the cultural variation and the possible drivers of overall disagreement. Component 2 explains roughly 10 percent of the variation in the agreement for the three-country sample. After examining the items exhibiting extreme values on Component 2, we describe this contrasting component as "child development concerns" versus "parental self-fulfillment." Unlike Component 1, which had a strong contrast between positive and negative childbearing motivation items, Component 2 had a mixture of positive and negatively worded items contrasting on both poles of the component. The three items scoring highest on Component 2 are all related to the benefits and problems associated with investing in the development of children. The items that score most extremely negatively on Component 2 are each related to the benefits and costs associated to becoming a parent. The most extreme item relates to religious fulfillment that is brought through parenthood. The U.S. sample shows a sharp contrast between the other two samples on this item, demonstrating an important national difference on this secondary dimension of cultural variation of childbearing motivation. The U.S. sample also stands out in its high undesirable rating of the

concerns about a child being unhappy or poorly adjusted and the item about a child being a burden as well as its lower desirability of the item about gender fulfillment of parenthood.

Figure 3.4. PCA Component 1 Versus Component 2 Scatter Plot
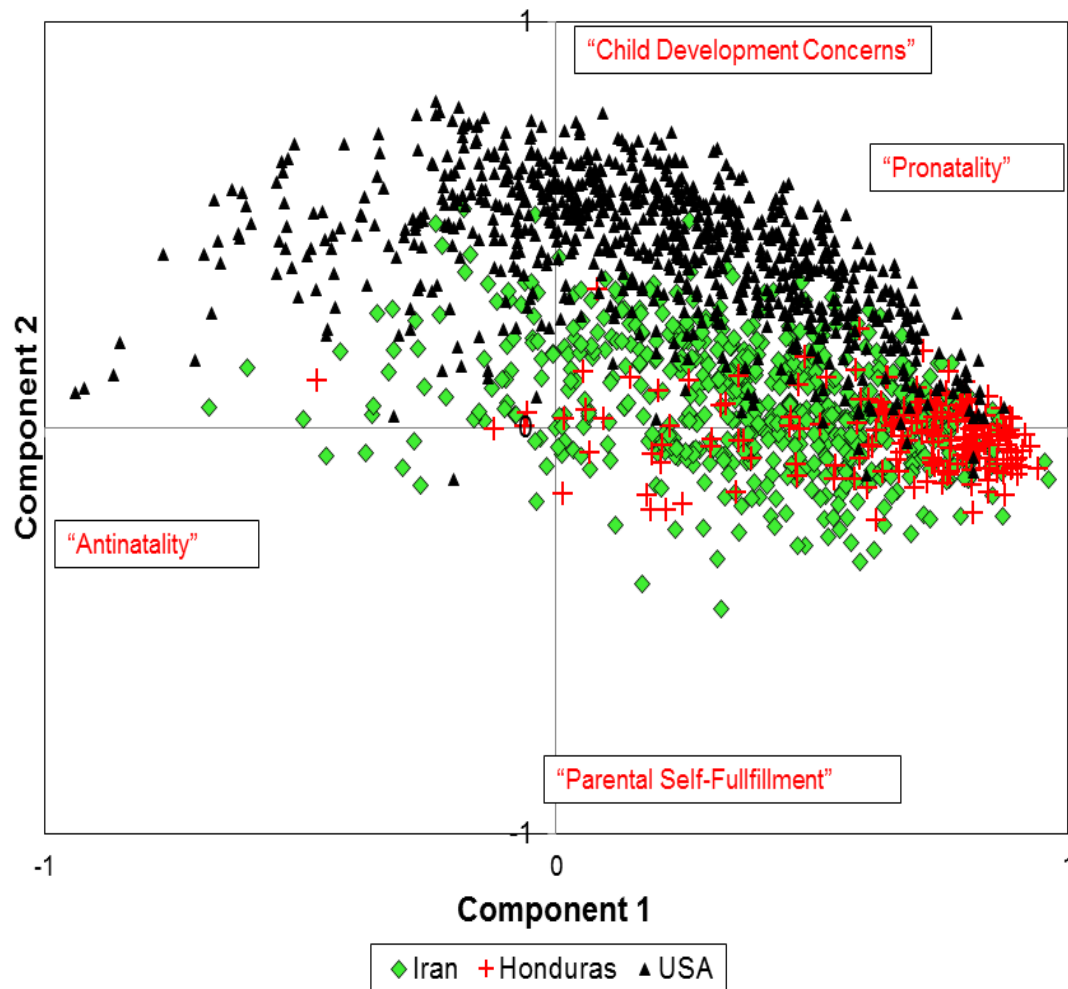
NOTE: The PC1, which represents agreement with the overall consensus on childbearing motivation questions, is depicted on the x-axis. The second component, which explains less variance and represents child development concerns versus parental self-fulfillment, is depicted on the y-axis.

Culture of Masculinity and Sexual Risk Among Homeless Men

Measuring culture using a latent variable approach enables researchers to address the question "is there *a* culture?" For example, we applied this approach in a study of HIV risk and masculinity among heterosexual homeless men. The literature on heterosexual HIV risk suggested that adherence to traditional gender roles was a predictor of heterosexual HIV risk behavior, especially among economically marginalized men. We developed and tested a

consensus-analysis approach to test this hypothesis with heterosexual homeless men in Los Angeles.

To develop the consensus model, we integrated qualitative and quantitative steps. First, we conducted exploratory interviews with homeless men about their views of gender and masculinity. Following the exploratory interviews, we conducted qualitative analysis to identify themes and constructed structured interview items that illustrated these themes. We then produced a structured questionnaire that we administered with a larger sample. Items in this structured questionnaire were often verbatim or close to verbatim statements from the exploratory interviews. Next, we conducted CCA on a data set from the larger sample of homeless men and tested for the existence of a culture of masculinity among them.

As shown in Figure 3.5, the points cluster toward the maximum value of Component 1 with the appearance of random scatter for Component 2. These and other rules of thumb suggest one (single) culture of masculinity among these men, rather than multiple centers of cultural consensus. In other words, most men agree with statements endorsing responsibility and equality in relationships and describing how difficult relationships can be when homeless. And most men disagree with statements endorsing "traditional" masculinity, including double standards for men versus women or endorsement of patriarchal beliefs. Meanwhile, men were more scattered than clustered in terms of their component scores on principal component 2.

Figure 3.5. Cultural Consensus Analysis Results Using Principal Component Analysis for an Informal Consensus Approach
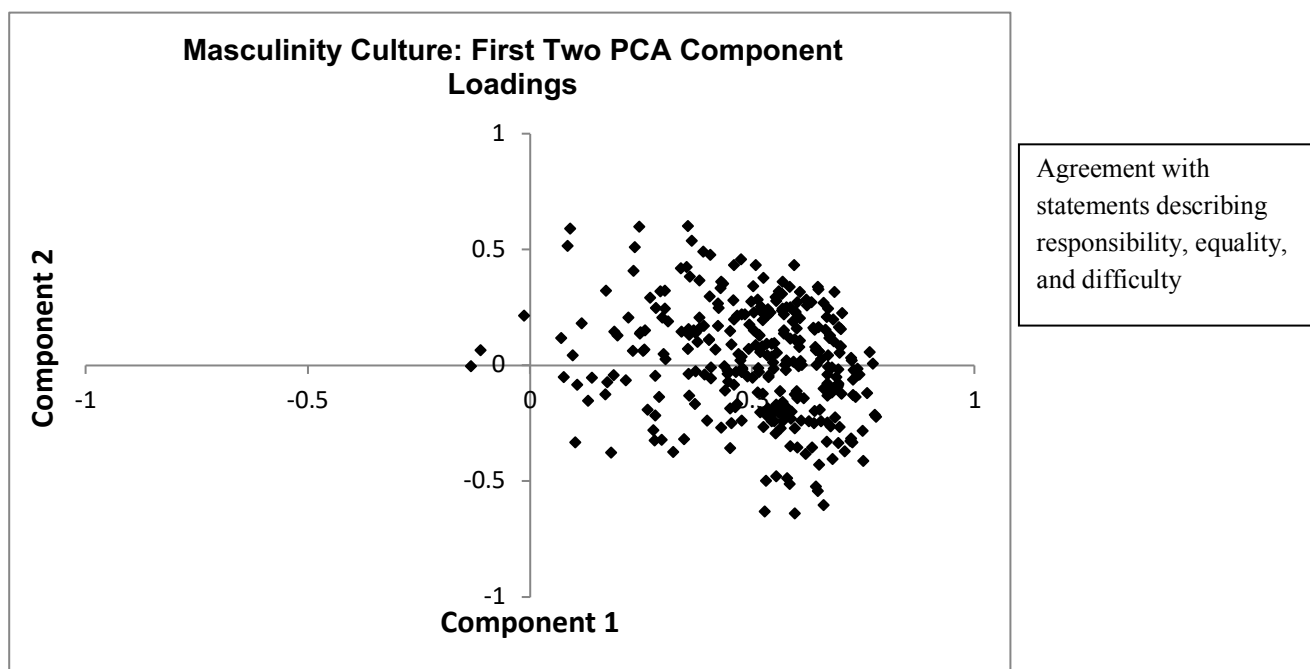
NOTE: The figure is a scatter plot of the first and second component loadings for each of the 305 men in the sample. Each point represents a respondent. Component 1 is labeled to show where individuals are located who tend to agree with different types of items.

Examination of the statement scores driving the components along with qualitative data provided insight into interpreting the first component. Statements that had high scores for Component 1 included the following:

- "A man's number 1 responsibility is to protect and provide for his family."
- "Men and women should share decisions equally."

Statements that had low scores for Component 1 included the following:

- "If a man pays for sex, he should not have to use a condom."
- "Men who have a lot of sex with different women should be admired."

Component 1 thus appears represent a range of beliefs between high agreement with gender ideals (on the right side of Figure 3.4), such as equality between men and between women and men and traditional masculinity (on the left side of Figure 3.4). Component 1 is also driven by ideals of men needing to display responsibility in their relationships with women as well as by statements about the challenges homeless men face when trying to pursue relationships on the street; for example, "It's hard to have serious relationships on the street, because there is no privacy or place to be alone together." Thus, we transformed individual scores on Component 1 into a measure of agreement with traditional masculinity ideals.

We then tested whether the resulting masculinity measure corresponded to HIV risk behavior with female partners (see Table 3.5).

Table 3.5. Multilevel Predictors of Unprotected Sex Among Heterosexual Homeless Men

	Significant Variables	OR (95% Confidence Interval) P-value
Individual	• Overall mental health • Condom efficacy • Number of male sex partners • CCA masculinity measure	0.98 (0.95, 1.0) [*] 0.31 (0.13, 0.77) [*] 2.10 (1.09, 4.07) [*] *Ns*
Relationship	• Talked with about HIV • Talked about condoms • Met on street • Frequency of contact • Emotional support • Relationship commitment	0.10 (0.04, 0.28) [**] 0.05 (0.02, 0.13) [**] 2.36 (1.01, 5.49) [*] 1.59 (1.06, 2.39) [**] 4.48 (1.68, 11.92) [**] 2.03 (1.29, 3.2) [**]
Social network	• Closeness centrality	1.02 (1.0, 1.04) [**]

SOURCE: Kennedy et al., 2013.
NOTE: * $p < 0.05$, ** $p < 0.01$

Controlling for other factors, we found individual, relationship, and social network factors that predicted unprotected sex with particular partners. However, the measure of masculinity that emerged from CCA did not predict HIV risk behavior. It showed some promise of significance in bivariate test but was not significant when controlling for other factors. Therefore, this test suggests that other factors besides adherence to traditional masculinity beliefs were more important in HIV risk behavior among homeless men.

Cultural Consensus Analysis, Qualitative Methods, and Statistical Methods

This chapter has covered CCA, which is a core quantitative technique for analyzing qualitative data. We note that, prior to applying CCA, qualitative methods should be used to develop a theory of the cultural model and the questions that will be used to measure aspects of this model. Qualitative research strategies such as participant observation, case studies, or repeated unstructured interviewing aimed at eliciting cognitive domains fulfill a central role in the ethnographic process. This being said, to determine the amount by which informants differ from one another on complex variables with subtle differences, ethnographers often need more precision than qualitative methods alone can provide. Developing quantitative measurements of concepts developed qualitatively allows for precise comparisons of variables and testing of theories using multivariate statistical techniques. Qualitative interpretation of data (either quantitative or qualitative) implies quantitative comparisons. Operationalization of conceptual variables and analysis using multivariate statistical techniques makes these comparisons explicit (Handwerker, 2001, pp. 11–21).

The nature of cultural data implies that responses given by one informant are, by definition, spatially and temporally autocorrelated with the answers given by other informants (Handwerker and Wozniak, 1997). Thus, multivariate statistical methods can be more appropriate for analyzing culture than are traditional statistical modeling techniques such as multiple regression. Although nonindependence is considered a problem needing to be corrected in regression analysis, the goal of statistical analyses of cultural data is to understand, define, and measure this correlation (Handwerker, 2001, p. 19). Because multivariate statistical techniques can be used to analyze variance in informants rather than variables, they provide the means to measure this autocorrelation.

The socially constructed nature of cultural data also causes another problem with classical statistical methods because of sample size requirements. To determine subtle cultural differences that may lead to cultural boundaries, the qualitative techniques that allow researchers to uncover cultural meanings are often too taxing and labor-intensive to undertake on large sample sizes. Multivariate techniques such as factor analysis and PCA can determine whether data sets have structure or random distribution using much smaller sample sizes than traditional demographic or psychometric statistical techniques. For example, Weller (2007, p. 354) shows how interviews

with cultural "experts" who have a high degree of knowledge about core shared cultural domains can produce evidence of cultural consensus in sample sizes of ten or fewer.

To determine external validity (i.e., the ability to generalize from the analysis of results on one sample), using structured interview questions allows for greater comparability than traditional qualitative ethnographic techniques used in isolation. Instruments developed for a community study can be exported and used in other communities to test hypotheses. Structured interviewing also allows for incorporation of standard large-scale demographic instrument questions, making local data comparable with national data as well as development of national surveys from proven instruments produced at the community level. Conversely, standardized questions developed in local settings can be used in subsequent national or regional large-sample surveys to test the prevalence of local findings in multiple settings.

This method has important ramifications for the development of an anthropological demography. Once an operational definition is given to a culturally based construct variable, empirical evidence can determine how well this construct holds together. First, this method can be used to justify the claim that a theoretical construct has construct validity for one population. This method can then be used to evaluate the consistency of findings with future research that uses the same operational definitions of theoretical concepts for other populations. Thus, this method can be the key to linking traditional anthropological local studies with traditional demographic studies of large national or regional populations.

4. Using Cultural Measurements in Statistical Models with Other Variables

Once we have results from a CCA or a related method, we then can ask what explains the variation in individuals' positions in the inferred cultural space or how those cultural positions predict other variables of interest. An individual's cultural positions might have been measured through some type of cluster analysis that assigned them to a nominal category or through a CCA that assigned factor loadings for each individual on a particular latent cultural construct (see Chapter 3). Recall that one goal of CCA can be to construct a set of axes of cultural variation from the original measured traits. This process makes most sense when the hypothesis is that our direct observations are in fact correlated measures of one or more underlying cultural constructs. For example, we may have asked a number of survey questions about race and ethnicity, with the hypothesis that the answers to those questions reflect a single underlying cultural construct of ethnocentrism. The variation in the questions that does not load on PC1 or factor 1 is, in this case, interpreted as measurement error. If, instead, we think there may be multiple qualitatively distinct subcultures in our data, such that the measured individuals do not all lie on a single continuum, then we should explore a PCA or cluster analysis that could first suitably group the variables or data points. If sample size is sufficient, we then might apply CCA within each empirically determined subculture.

Modeling Cultural Measures Together with Attribute Data

Once we have such a result, we can use the cultural measures to predict other outcomes of interest, such as health traits or security risks. In these types of analyses, the cultural measurements can be used in the same manner as any other variable in a statistical model.

We also might try to explain what causes individuals to be positioned as they are in the cultural space. One approach to this problem would be through more or less standard linear regression of the consensus results onto a number of attributes of the individuals. These attributes could be interpreted as (1) features that make individuals more receptive to some cultural information, (2) proxy measures of individuals' social exposures, or (3) causal variables that are processes unrelated to social learning that nevertheless influence the results of CCA. If a measured attribute of individuals is significantly associated with the cultural measure, then any three of these interpretations may apply.

The reason for this threefold interpretation becomes clear if we recall that our definition for culture is socially learned behavior rather than the sum of all things people do. Thus, for any trait measured through CCA, there may be some important causal contributors to the variation in the trait that are not related to social learning. Consider the hypothetical example that we had

measured how close individuals were to the culture of the National Association for Stock Car Auto Racing (NASCAR). This is not merely a matter of how much individuals like NASCAR or even follow it, but how much they agree with the NASCAR-related beliefs, behaviors, and attitudes of the model NASCAR fan. Individuals who are close to the model NASCAR fan in their cultural measure may also be similar in a number of socioeconomic attributes, such as their income, their state of residence, whether or not they live in the American South, whether they live in an urban or rural area, and their profession. What should we make of such attribute correlations with our cultural measures of NASCAR?

Interpretation 1 would argue that these attributes make NASCAR more appealing or easier to learn. Thus, all Americans might be equally exposed to NASCAR, but some attributes of individuals make NASCAR a more preferred entertainment. For example, rural Americans may have fewer entertainment options than urban Americans, which might make NASCAR relatively more appealing if it is nearby. Perhaps the cost of tickets or mode of domicile (e.g., trailers, tailgating) make NASCAR more appealing to lower-income Americans than opera performed at a major opera house within the heart of an expensive urban center (e.g., the Metropolitan Opera in Manhattan). Thus, under Interpretation 1, the attributes of individuals alter their receptivity to a socially learned trait (such as NASCAR fandom) potentially because they reflect variable incentives that make the socially learned trait more or less economically "rational" for some groups of people compared with others.

Interpretation 2 proposes that rather than altering receptivity, the correlated attributes are not really causal but instead are proxies for how much exposure individuals have to NASCAR fanship. Culture is socially learned information, and perhaps income, region, and other attributes are good proxy measures for how much social exposure individuals have to NASCAR. We often think of these exposures as occurring through social networks of family and friends that surround a person. Perhaps many residents on the Upper West Side of Manhattan would actually enjoy NASCAR, but their social networks are depriving them of the exposures needed to socially learn NASCAR fanship. Maybe people living in the rural South would actually love opera, but they do not know anyone who would expose them to it.

Interpretations 1 and 2 are not mutually exclusive. Any single attribute (e.g., household income) could correlate with the cultural measure both because it alters individuals' receptivity to that cultural construct and because it is a proxy for social exposure. The only way to disentangle these is to measure the social exposures of individuals more directly, which usually is done through some form of social network analysis (see the next section).

The third interpretation also is not mutually exclusive. Interpretation 3 is that the attribute influences the cultural measure through some mechanism that does not involve social learning. In this case, some amount of variation in the cultural measure that is explained by the attribute would actually be noncultural in origin. Genetic variation is one such potential contributor. Could there be a gene for NASCAR? Despite what fans may say, we think this is unlikely. In some cases where mixed methods have been used to measure a cultural outcome, however, the

conclusion is not as clear. Consider, for example, our discussion in Chapter 3 about variation in sexual risk taking. While social learning clearly plays a role in sexual risk taking, there is also known genetic variation that alters individuals' propensities to take sexual risks (Garcia et al., 2010; Harden, 2014).

Modeling Cultural Measures on Networks

The threefold complexity of interpreting regressions of CCA results onto attribute data strongly motivate more-direct measures of social exposure through analysis of social networks. After all, our hypothesis is that CCA results arise from a process of socially learned information. Therefore, the network that governs the social transmission of this cultural trait should be correlated with the variations among individuals in the trait. Individuals who are close together in the network (i.e., directly tied or one degree separated) should have more similar cultural measures than do individuals who are separated by more intermediary nodes in the network structure.

If the social networks change over time, and often they do, then any correlation between the network and the cultural measures may be caused by homophily instead of social influence (Shalizi and Thomas, 2011). Social influence is any process by which connected individuals become more similar over time; for example, a new belief may spread through a network across the personal connections of individuals. Cultural diffusion, the spread of beliefs or behaviors by social learning, is one form of social influence.

Homophily is any process by which individuals with similar traits end up being more likely to form network ties with one another. Homophily can be caused by people preferring to connect with similar other individuals, but also by other processes such as similar people inhabiting the same neighborhoods or workplaces and therefore being more likely to form friendships, even though they do not have a particular social preference for similarity. The latter process is sometimes called baseline homophily. In the United States, there is demonstrable race homophily that is both baseline and preference-induced (McPherson, Smith-Lovin, and Cook, 2001; Wimmer and Lewis, 2010).

How can we sort out homophily (sorting into similar groups) from genuine social influence (members of the network changing each other's behavior)? There is no good analytic solution to this problem (Shalizi and Thomas, 2011); experiments potentially can separate the two but are often intractable for social science research. Instrumental variables have been proposed as an analytic solution, but again, in social science, it often is difficult to find a completely exogenous instrumental variable.

Often, however, we can use other kinds of evidence to exclude one process or another. For example, in the case of the correlation of race on networks (i.e., individuals of the same race are more likely to have ties than are individuals with difference races), we can conclude that the formation of the network ties is likely much more flexible than is the ability of individuals to

pass as a member of another race, even though the latter has happened to a degree. Thus, we can conclude reasonably that correlations of race on networks are likely because of homophily (individuals tending to form more same-race social ties) rather than influence (individuals changing their race to match the race of their social connections).

Often, we can make a similar argument for the results from CCA. For example, if our explanatory network constitutes ties of close friendships among the individuals, then a correlation with cultural traits that change over time scales of ten years or less mostly likely is driven by social influence. This is because adult friendship networks are highly stable and change relatively slowly over time (O'Malley and Christakis, 2011).

Last, it should be noted that often an important research question is to determine which of several plausible networks are most related to the cultural measures somewhat irrespective of the exact directions of causality. For example, we might ask whether the international spread of positive or negative attitudes about the United States occurs more through social media, as measured by the Twitter network, or through face-to-face interactions as measured by the network of international plane travel. Suppose the results of such a study strongly favored Twitter network data over the plane travel network because the Twitter network showed a strong correlation with a CCA of attitudes toward the United States. This would not fully resolve whether the correlation with Twitter is caused by influence or homophily, but it would resolve that social media has some kind of more important relationship to U.S. attitudes than do face-to-face interactions.

Network Science Statistical Techniques

There are two main statistical techniques to assess how cultural measures such as the results of CCA map onto hypothesized transmission networks. These techniques are (1) autoregression and (2) dyadic regression implemented through either matrix permutation or random effects modeling. If a researcher is not using network data and only regressing consensus results onto attribute data, then we recommend consulting any number of linear regression textbooks. When using attribute data only, there is nothing particularly unique about the models for consensus results. Network statistical models for cultural outcomes are more specialized. We lay out the rationale for these models and provide a worked example.

Autoregression

This statistical technique models the cultural measure as a dependent variable in a regression model in which the residual terms are autocorrelated along network connections. If there are no independent predictor variables other than an intercept, then the model is a pure autoregression in which each individual's deviation from the mean is predicted by the average deviation of their social connections. The intercept in such a model reflects the mean of the dependent variable,

which, in our case, is a cultural measure, possibly the principal component scores or factor loadings from CCA.

One advantage of the autoregression approach is that it is straightforward to combine traditional independent variables while simultaneously testing for an effect of a hypothesized social network. Take our hypothetical example, where we wanted to test which network most predicts how individuals vary in their attitudes about the United States. We might want to test whether an online network from Twitter or a plane travel network is more predictive of attitudinal variation after first controlling for whether or not each subject has travelled to the United States or has a close friend who has travelled to the United States. Our hypothesis is that travel to the United States by self or friends improves attitudes toward the country. In an autoregression model, it is easy to incorporate these variables as traditional binary predictors. The effect of the network matrices is then estimated after controlling for the effects of the independent variables.

Incorporating regular independent variables in a single model with network variables is substantially more difficult with the other two techniques discussed in the next section because these techniques require independent predictors to be converted into pairwise distances. Such a procedure reduces the information present in the independent variable. It can lead to errors of statistical inference because usually we are hypothesizing that higher or lower values of the independent variables cause higher or lower values of the dependent variable, rather than simply that similar values of the independent variables cause similar values of the dependent variable. Thus, in theory, autoregression is the most elegant solution to network autocorrelation models because regular multiple regression can be incorporated within it.

The difficulty, however, with autoregression models for network data is that actual implementations of them have not fully resolved statistical errors that include biased estimates for effect sizes and p-values (Dittrich, Leenders, and Mulder, 2017; Karimov and Matthews, 2017; Matthews et al., 2016; Mizruchi and Neuman, 2008; Neuman and Mizruchi, 2010). In the case of simplified networks that branch, such as phylogenetic trees, the math has been validated numerous times (Felsenstein, 1985; Grafen, 1989; Nunn, 2011; Rohlf, 2006). Phylogenetic trees can be considered a kind of network in which each measured individual is connected in a simple bifurcating structure of internal nodes. Thus, at present, the statistically most sound approaches for more complex nonbifurcating network data would be the slightly less elegant random effects and permutation techniques described in the next section. For a lengthier discussion of the implementation of autoregression models and their history in cross-cultural research, see Chapter 6.

Dyadic or Matrix Regression

Another set of techniques to assess how cultural measures pattern on social networks are called either *dyadic regression* or *matrix regressions* in the literature. The basic procedure is to construct pairwise distances out of the dependent cultural measure variable and then regress

those distances on one or more matrices of network connections. As long as we are only evaluating the relationship of the dependent variable to network variables, then this dyadic regression procedure likely produces acceptable false positive rates (Harmon and Glor, 2010). The trouble with conversions to distances comes when we try to combine traditional independent variables in the same analysis. If that type of analysis is needed, the autoregression approach would be preferred.

Dyadic regression can be useful particularly when we are trying to evaluate which of several alternative networks is most associated with the cultural measure. There is at least some literature that establishes that these techniques produce acceptable false positive rates in the context of multiple collinear network predictor variables (de Nooy, 2011; Dekker, Krackhardt, and Snijders, 2007; Karimov and Matthews, 2017; O'Malley and Christakis, 2011). The dyadic regression creates pairwise distances from the dependent variable. There will be $N*(N–1)/2$ distances for a variable, where N is the number of individuals studied. In addition to the change in sample size compared with the original cultural measure, the distances are nonindependent because each individual's cultural measures are used repeatedly to calculate distances to the various other individuals. This repetition of individuals throughout the data is similar to repeated measures in study designs common in physiological and biomedical studies, in which individuals are measured on the same variable over multiple occasions.

For years, the most common implementation of the dyadic or matrix regression was the Mantel test. The classic Mantel test attempted to deal with the nonindependent and repeated measurement of data points by permuting the rows and columns of the dependent distances. The permutations generated a null distribution that was used for significance testing.

Unfortunately, the Mantel permutation does not fully correct for the nonindependent structure of dyadic distances (Dekker, Krackhardt, and Snijders, 2007). In particular, the Mantel permutation is susceptible to inflated false positive rates when the multiple networks being regressed on are themselves somewhat correlated. Although this condition, called *multicollinearity*, is technically a violation of standard regression models, standard regression models are highly robust to even moderate levels of multicollinearity. Dekker, Krackhardt, and Snijders (2007) showed that the Mantel permutation is not similarly robust. A different procedure of permuting the residuals of each independent predictor (the networks in our case) does appear to be robust to a collinear and spurious network being included in the model. This method can be implemented in the R package *sna*, with the functions *netlm* and *netlogit*.

Recall that the repetition of the original cultural measure in the pairwise distances is conceptually similar to a repeated measures design. This observation has led to a wholly different approach to dyadic regression, which uses random effects mixed models to correct for the repetition of individuals' identities. Both O'Malley and Christakis (2011) and de Nooy (2011) independently developed this approach in the context of regressing the formation of network ties on multiple network matrices and distances from variables. They offered some mathematical proofs for the validity of the approach and empirical results, although no

simulation work has been conducted. Their technique essentially includes two crossed random effects for the identity of each member of a social relationship. These random effects allow slightly different intercepts to be applied to each individual.

One advantage of the random effects method for dyadic regression is that it allows fitting these models in an information theoretic framework in which maximum likelihood or Bayesian criteria can be used to decide between models. This is advantageous, particularly when considering policy objectives that often focus not so much on whether this or that network is a significant statistical predictor, but rather what combination of networks best predicts the cultural measure. It is difficult to explore the various combinations of network predictors (the "model space") if we are only using frequentist statistics that focus on the p-value of each network, which itself changes with each model formulation. Using an information theoretic approach— such as likelihood, the Akaike Information Criterion, the Bayesian Information Criterion, or another related method—has advantages because we can compare the predictive utility of models considered as wholes and against other such whole models.

*An Example of Simulating and Analyzing Cultural Data on Network Variables*

The following is a worked example of the network regression techniques discussed in this chapter. The basis is work by Matthews et al. (2013) that examined how theological and liturgical traits were distributed among extremist Christian religious groups called the Anabaptists, sects that originated during the Protestant Reformation in Europe. All of them rejected secular governments and supported the establishment of Christian theocracies. They are an interesting study system because about half of them advocated using violence to establish such theocracies, even taking over the city of Muenster by force at one point. The other half were radical pacifists against any use of warfare. The violent sects were brutally suppressed by the kings, Catholics, and Lutherans, and while the pacifist ones were suppressed as well, they left biological and cultural descendants primarily in the United States as today's Amish and Mennonites.

Matthews et al. (2013) assessed how advocating violence and 44 other cultural characteristics were distributed among these congregations. Specifically, they tested whether the traits were similar across congregations that shared personal connections of leaders (based on records of letters, see Figure 4.1) or shared more recent history as congregations (Figure 4.2). The former produced a network model while the latter a congregational tree model. Matthews et al. (2013) used the random effects technique for this study on dyadic measures, but no study has examined the relative statistical performance of the random effects technique, the permutation technique, and the autoregression technique.

We recently simulated data sets of 45 binary characteristics evolving on the Anabaptist congregational tree or diffusing along the network of connects between group leaders (Karimov and Matthews, 2017). The nodes of the tree and network were 11 congregations of Anabaptists from the Reformation era (Figures 4.1 and 4.2). The simulation methods proceeded through a

series of time steps and essentially are simple forms of ABMs. Since we know how each of these data sets were generated, we can apply each of the statistical techniques and assess the power and type 1 error. Statistical power is the proportion of times we reject the null hypothesis when it is in fact false. So the power to detect inheritance on the congregational tree is the percentage of 1,000 simulated data sets in which we find a significant result for the tree when the data were generated on the tree model. Likewise, for the power to detect network effects, a Type 1 error is the proportion of times one rejects the null hypothesis when it is true. In this exercise, we assessed type 1 error as the proportion of times a tree or network model was significant when the data were not generated from that model.

Figure 4.1. Anabaptist Congregational Network Based on Personal Correspondences of Leaders

SOURCE: Redrawn from Matthews et al. (2013).

50
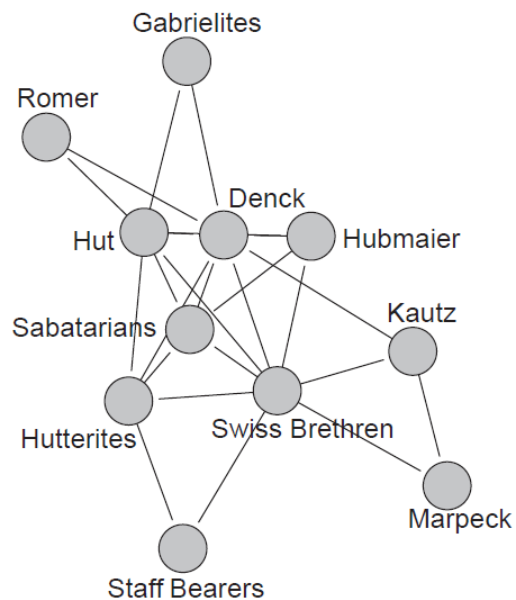
Figure 4.2. Anabaptist Congregational Tree Based on the History of Schisms Among Congregations
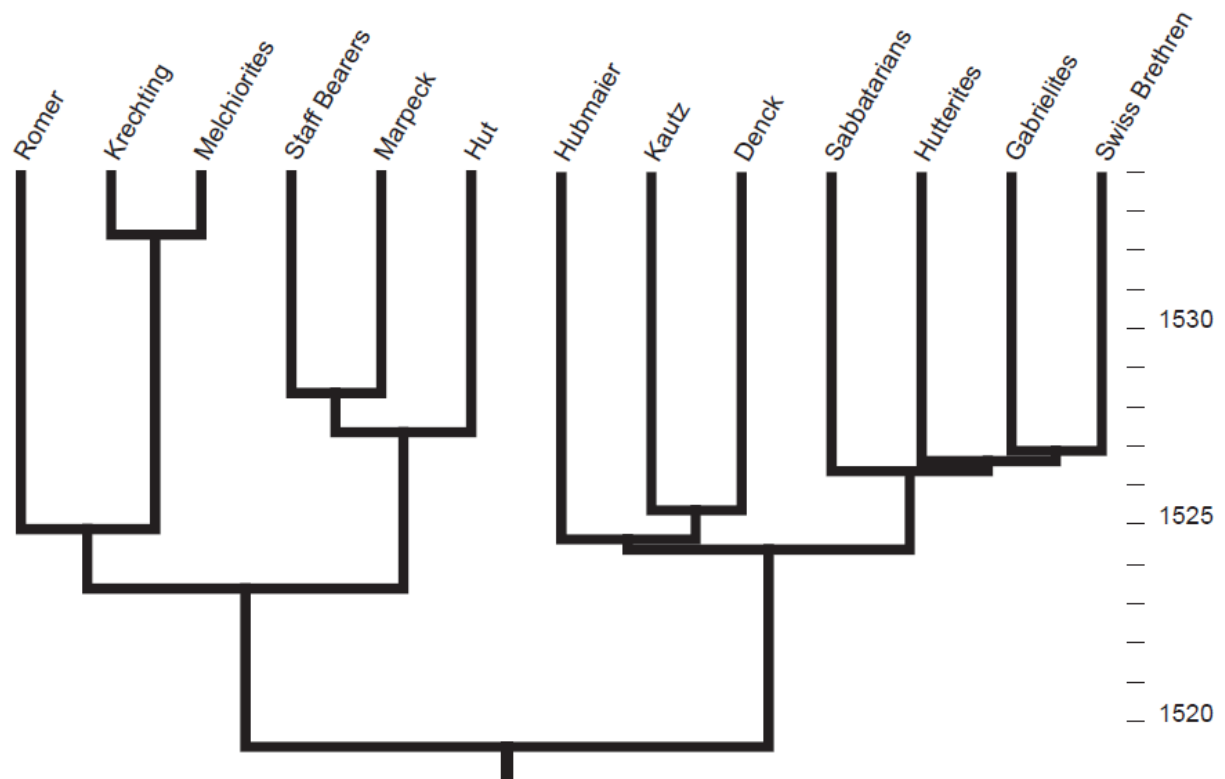
SOURCE: Redrawn from Matthews et al. (2013).

Simulated data are useful because they give us a known known: We know whether the traits were inherited on a tree or diffused on a network. This enables us to evaluate directly whether the statistical procedures for this problem are performing as claimed. We proceeded with standard implementations of autoregression and dyadic regression with permutations or random effects.

When we simulated data on the Anabaptist tree and network, we found that the permutation and random effects approaches performed much better than did autoregression as implemented through *lnam* function in the R package *sna*.

Table 4.1. Performance of Statistical Methods on Simulated Data

Statistical Method	Tree Sims Power (True Positive)	Network Sims Power Ttrue Positive)	Tree Sims False Positive Network	Network Sims False Positive Tree
Performance goal	1	1	0.05 or less	0.05 or less
Continuous character				
Lnam	0.255	0.378	0.212	0.334
Permutation	0.336	0.35	0.049	0.047
Random effects	0.365	0.389	0.093	0.041
Discrete characters				
Lnam	0.719	0.167	0.322	0.310
Permutation	0.978	1	0.041	0.17
Random Effects	0.986	1	0.098	0.051

SOURCE: Redrawn from Karimov and Matthews (2017).

We also performed a fuller set of simulated inheritance and diffusion where the networks and trees were simulated as well. We simulated networks and trees of 20, 50, and 100 nodes. In these simulations, we confirmed that linear regression without accounting for nonindependence performed very poorly and with substantially elevated false positive rates. Across all simulated conditions, the random effects method showed the best performance in terms of statistical power and had statistically acceptable false positive rates under conditions in which both *lnam* and permutations exhibited elevated false positives (Karimov and Matthews, 2017).

It may at first appear to be a contradiction to claim the random effects model both increased power and reduced false positives, since at times these are viewed as trade-offs. Prior simulations, however, have shown that methods to correct for nonindependence increase statistical power because they make the model more accurately reflect the data structure (Rohlf, 2006).

Summary

The random effects technique shows the most promise because it reduced type 1 error to statistically acceptable levels across nearly all conditions while maintaining high power to detect real effects. Thus, our work recommends the random effects technique as the best currently implemented option for determining which among several trees or networks is most correlated to a cultural outcome such as a set of CCA scores.

Another more general lesson from this exercise is that small sample sizes do not necessarily prohibit or invalidate the proper application of statistical techniques such as regression. When we fit models that were appropriate to how the data were generated, statistical power was high and type 1 error rates manageable even for data with only 11 or 20 data points. We made a theoretical case for this at the end of Chapter 3, and this exercise verifies this notion at least in simulated data.

We think the perception that large samples are needed for statistical techniques stems in part from the application of statistics to data with large amounts of *measurement error* or inappropriate models. Models can be inappropriate for many reasons, but some common ones include those (1) lacking important confounding variables, (2) that have inappropriate control variables that are not actually causally related to the outcome, and (3) that do not reflect the data-generative process.

ABMs similar to the ones from this exercise are also an excellent way to test model performance under different hypothesized processes that generate observed data. The ABMs in this exercise are relatively simple and easy to run for someone proficient in R. We believe the routine application of this approach—thinking through the range of plausible statistical models and testing them systematically with ABMs—will open up great opportunities for statistically investigating cultural phenomenon even when sample sizes are small.

5. Dealing with Cultural Data-Point Nonindependence—Galton's Problem

Within the cultural phylogenetic toolkit, the methods most directly applicable to policy research center on Galton's problem, a statistical issue first identified in the context of cross-cultural research. The cultural and biological phylogenetics communities have been substantively involved in the statistical solution to Galton's problem, but the issue applies broadly to many cultural systems even without deep temporal durability.

The statistical issue of data-point nonindependence has played a prominent role in the analysis of cultural data. The central issue is that standard statistical techniques, such as t-tests, regression, and ANOVA, all rely on the assumption that individual data points are completely independent measures of the relationships among the dependent and independent variables. This assumption is violated frequently in cultural data because at its core, culture is a social phenomenon that arises from groups of interacting individuals who, through that interaction, come to share beliefs and practices. Cultural traits are nonindependent at a fundamental level because culture is epidemiological in the sense that it is transmitted among agents who learn socially from one another. Thus, one person's socially learned beliefs and practices cannot be a fully independent measurement from the beliefs and practices of the person from whom they learned.

The issue of data-point nonindependence extends to emergent levels beyond pairs of individuals to include larger groupings because social life creates clusters of culturally interrelated individuals. Even simple social-learning mechanisms working within such subgroups cause the emergence of discontinuities in socially learned behavior (Matthews, 2009). We colloquially recognize these discontinuities as *cultures*, and these cultures are connected through higher-order patterns of differential relationships and behavioral similarities (Mace et al., 1994). The main cultural processes discussed at the level of groups or societies are diffusion, which is the spread of beliefs and behaviors among contemporaneous groups, and inheritance, which is the vertical transmission through time of beliefs and practices from parent groups to child groups. Both processes create nonindependence in cultural data. Diffusion processes usually are modeled statistically through some form of network analysis, while inheritance processes are modeled through phylogenetic analysis. Both of those analytic methods share much common ground statistically, as they do with spatial autocorrelation analysis.

The issue of the frequent nonindependence of cultural data goes back to the beginnings of anthropology. In 1889, Edward Tylor gave a presentation to the Royal Anthropological Institute that detailed his correlational analysis of marriage and property inheritance patterns across hundreds of societies. During the presentation, Francis Galton questioned Tylor about how he could account for the fact that various groups of people were differentially related to one another

by varying degrees of shared inherited or diffused beliefs and behaviors (e.g., all the Scandinavian countries share a more common heritage with one another than they do with other European countries, Pacific Northwest Native Americans share much cultural heritage, and so forth). Galton argued that the fact cultures were interrelated in this way meant the data were statistically nonindependent and therefore inappropriate for correlation analysis, which assumed independence of data points (Naroll, 1961).

This is why the issue of statistical nonindependence of societies in cross-cultural analysis came to be known as Galton's problem. Unfortunately, Galton had no solution for Tylor. Among other factors, Galton's problem led some contemporary and subsequent anthropologists to write off the entire practice of cross-cultural analysis. For example, Frans Boas, who had a tremendous influence on 20th-century anthropology, considered Galton's problem to be fatal to the cross-cultural research enterprise (Naroll, 1961).

The problem of statistical nonindependence is not unique to cross-cultural research, and it can apply equally at many levels of analysis. For example, studies of the IQ of individuals could exhibit nonindependence if the individuals share varying levels of genetic relationship or different exposures to unmeasured variables that only some of the sampled individuals have in common. Cities and nations also can exhibit statistical nonindependence as a result of spatial proximities (Corrado and Fingleton, 2012).

Nonindependence likely gets more attention in cross-cultural analyses because the very processes of cultural inheritance and diffusion should produce nonindependence. The same is true of inheritance in biological systems. Thus, it is not surprising that biologists and anthropologists have thought a lot about how to correct for nonindependence (Nunn, 2011).

As a statistical issue, however, it does not matter whether data-point nonindependence is caused by inheritance, diffusion, unmeasured confounding variables, or the tendency of individuals with similar traits to associate with one another (i.e., homophily) (Dow et al., 1984; Felsenstein, 1985; Grafen, 1989; Leenders, 2002; O'Malley and Marsden, 2008; Pagel, 1999). Some of these processes, such as inheritance and diffusion, imply that the connections measured among the data points are truly causal in some sense. Some, such as unmeasured confounding, would mean that connections among the data points are not causal but mere useful proxies of unmeasured variables. Some processes, such as simultaneous diffusion and homophily together, cannot be fully separated in a causal analysis (Shalizi and Thomas, 2011). Regardless of the causal question, data-point nonindependence produces the same problems in standard statistical models, and it therefore deserves serious attention, even if only as a statistical matter.

Corrado and Fingleton framed this issue well within the economics community by stating that "[s]patial econometrics has been criticized by some economists because some model specifications have been driven by data-analytic considerations rather than having a firm foundation in economic theory" (2012). That is a fair characterization, but is it a criticism or a complement? Indeed, we want to test theory, but any scientific endeavor involves confronting theory with evidence. Thus, the theory cannot completely drive model construction, as this

would be an overly strong prior. Sometimes data-analytic considerations (i.e., accepting the real complexities of real data) can rightly drive components of a model even if that renders our explanations less elegant, clear, or theoretically coherent. Elegance, clarity, and coherence are all appropriate arbiters of artistic and philosophical investigation, but science is a confrontation of ideas with an external reality that, at times, is organized inconveniently for our cognitive and statistical models that are the historical products their own cultural evolution.

How Nonindependence Affects Models

The main context for dealing with data-point nonindependence has been regression modeling, especially since ANOVA, t-tests, and similar parametric statistical procedures can be seen as special cases of the regression. Regression models assess the relationship of a dependent variable on each independent predictor variable after accounting for the effects of all the other independent variables in the model. One important assumption of the regression is that the residual (unexplained) variation in the dependent variable is statistically independent. A plot of the residuals should look like white noise.

Nonindependent data violate this assumption that the residuals are like white noise. Instead, the data are clustered with high or low residuals based on some characteristic or identity.

There appears to be a relationship between the frequency of a gene in populations and the distance of the genes from the equator in Figure 5.1, but there is also clear clustering around the regression line based on continent. This is data-point nonindependence. There also appears to be a bit of heteroskedasticity (uneven variance) because the spread around the regression line looks larger as the latitude degrees increase. This is a different statistical problem from nonindependence; we will not address this.

Figure 5.1. Example of Statistical Nonindependence in Population Genetic Data
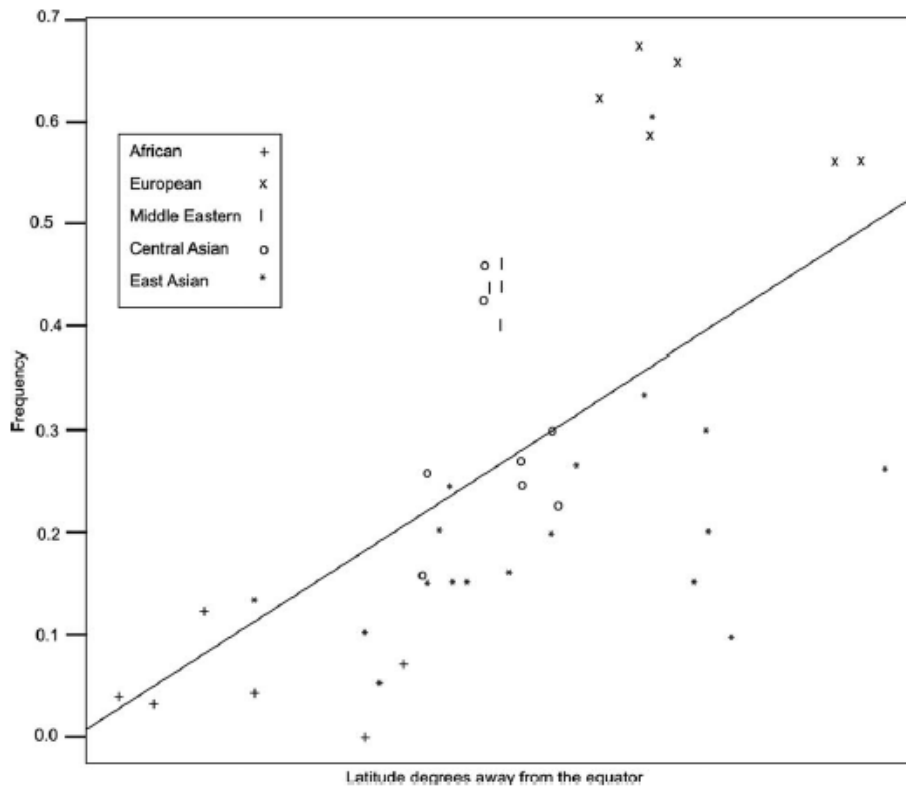
SOURCE: Redrawn from Coop et al. (2010).

So, what does nonindependence do to your statistical model? Essentially, it inflates the variance in the estimate of the slope coefficient outside the bounds of what the standard regression model estimates. This means that using standard regression on nonindependent data will result in a greater-than-expected number of false positive results, while also resulting in a greater-than-expected false negative result rate. In other words, if a standard regression is set at a 5 percent significance level, you expect that 1/20 results are actually false positive findings. If the data are nonindependent, however, the actual rate of false positives will be greater than 1/20. Suppose the power for your standard regression is estimated at 80 percent chance of rejecting a null when it is in fact false. If the data are nonindependent, the slope might be incorrectly estimated as too small rather than too large because nonindependence increases the variance of the estimate rather than biasing it, meaning your actual power is something less than 80 percent (Rohlf, 2006; Dow et al., 1984; Ward and Gleditsch, 2008).

Both effects will occur more frequently in small or moderately sized data sets (dozens to hundreds of data points). The change in the slope estimate after correcting for the nonindependence can result in a dramatically higher or lower slope estimate and a significant or nonsignificant finding (Rohlf, 2006). This is true because nonindependence does not bias the slope estimate itself, *but it does cause it to be randomly too high or too low* (Kissling and Carl,

2008; Rohlf, 2006). Here is an empirical example of this exact phenomenon in action ("PGLS" refers to a phylogenetic generalized least squares model, which is one method used in biological data to correct for nonindependence):

> Our PGLS example analysis shows a trend toward species with relatively larger brains having greater inhibitory control ($\beta = 66.93$, p = 0.06), indicating a possible functional link between these traits. In this case, we allowed λ to be estimated at its maximum likelihood value, which was 1.02 and therefore indicates that related species show similar deviations in inhibitory control relative to the expected value based on relative brain size (i.e., λ estimate is for the residuals of the model). To examine how including phylogeny in the analysis affects results, we conducted the same analysis with λ fixed to 0 [*This reduces the model to a standard regression that assumes independence*]. This analysis produced a much weaker association between brain size and inhibitory control ($\beta = 19.4$, p = 0.49).
>
> This difference in outcomes reflects an often underappreciated aspect of phylogenetically informed research: analyses incorporating phylogenetic information increase statistical power to detect real relationships, while reducing the probability of erroneously inferring significance when no association exists. (MacLean et al., 2012)

Policy-Relevant Applications of Modeling Nonindependence

Statistical models for nonindependence have played a significant role in several recent policy-relevant empirical studies. For example, Atkinson et al. (2016) examined whether more ecofriendly cultural beliefs or technologies were more important to the rate of deforestation on Austronesian islands. Because both cultural beliefs and technologies are socially learned phenomenon (i.e., culture) that are inherited from prior generations, the researchers used phylogenetic autoregression modeling to correct for data-point nonindependence. The historical divergences of each group's language from the others was used as a proxy measure for cultural sharing. This was reasonable because many prior anthropological studies have shown that a multitude of cultural traits correlate across recency of historical language divergence, at least for traditional societies. Atkinson et al.'s final analysis showed that technological capacity was much more determinative of deforestation than were cultural beliefs, which only mitigated environmental impacts to a small degree.

Building on the observation that language divergence predicts many cultural traits in traditional societies, Matthews et al. (2016) tested whether it similarly predicted political and economic behaviors of modern nation-states. This seemed plausible, since countries influence each other regarding their politics and economies, and culturally more similar countries may be of greater influence than are more culturally distant ones, either because countries want to differentiate from culturally similar countries or because they more readily adopt innovations from them. Consistent with this hypothesis, the study found that changes in the autocratic versus democratic character of a country and decisions to default on national debt frequently correlated

either positively or negatively across ties of language divergence over the past 100 years. Further research may be able to refine this model to render it accurate enough to better diversify investments in foreign nations by business or governments.

An example from policy research where data-point nonindependence was not considered fully is the work on "contagion" of obesity by Christakis and Fowler (2007). They examined changes in individuals' obesity status over a 30-year period in the Framingham Heart Study population. They found that when a person's friend became obese, the person themselves were more likely to become obese. These results were questioned for a number of statistical issues, one of which was inadequate control for the nonindependent nature of the data (Lyons, 2011). Subsequent statistical modeling has shown that the original Christakis and Fowler (2007) findings for obesity are substantially correct (i.e., most of the apparent spread of obesity was in fact due to some form of social influence) (VanderWeele, 2011; VanderWeele, Ogburn, and Tchetgen, 2014). Significant questions could have been avoided, however, if the original research had used statistical models that suitably corrected for data nonindependence.

Methods to Correct for Nonindependence

A fair number of methods have been proposed to deal with data-point nonindependence, due to a long and contentious history of researchers seeking a broadly applicable solution. Galton's problem renders indistinguishable whether traits are correlated because of some kind of functional relationship, such as maximum utility decisions under scarcity, or because both traits are simply inherited cultural features groups. This was the essence of the discussion between Tylor and Galton in 1889, with Tylor proposing that the marriage and property inheritance rules were functionally linked (a very reasonable hypothesis under any economic theory), and Galton arguing that these traits could be functionally unlinked and simply coinherited though the same pathways of social learning. Because Galton's problem renders functional and inheritance relationships indistinguishable in a comparative cultural analysis, Naroll (1961) stated, "Galton's problem is one of the most important problems of scientific method facing social scientists today." This remained true essentially until modern computers enabled the routine use of maximum likelihood estimation techniques.

Many empirical and simulation studies have been conducted in spatial statistics, network statistics, and biological and cultural phylogenetics to test the various methods proposed to deal with nonindependence. Published work supports that, at least when the relevant social learning pathways can be represented as treelike branching structures, the autoregressive model implemented through either maximum likelihood or Bayesian computation will produce appropriate statistical properties (Matthews and Butler, 2011; Páez, Scott, and Volz, 2008; Rohlf, 2006). Some autoregressive models cannot be validly solved through least squares and related methods (O'Malley and Marsden, 2008), which inhibited a full solution to data-point

nonindependence until the 1990s, when standard hardware-computing capabilities were sufficient for routine use of maximum likelihood estimation.

The recommendation of autoregression in this case appears to contrast with the findings presented in Chapter 4 of this manual, but the contexts are slightly different. In Chapter 4, we assessed dealing with nonindependence as it manifests when the researcher is selecting among multiple networks or trees to determine which set of proposed social learning pathways best predict the trait data. Here we are dealing with Galton's original identification of cultural nonindependence as it manifests when regressing an outcome trait against one or more predictor traits. In the former case of selecting among multiple hypothesized cultural networks or trees, we found that the standard implementation of network autoregression in R failed badly (*lnam* function in the R package *sna*). At present however, we are dealing with the ability of the autoregressive model to correct trait-trait correlations when the correct nonindependence measure is known and suitably accurately measured. For that use case, several studies point to the autoregressive model performing in a statistically correct manner (Matthews and Butler, 2011; Páez, Scott, and Volz, 2008; Rohlf, 2006), although none of those use the *lnam* function.

The autoregressive model is highly flexible because other common statistical procedures such as t-tests and ANOVA can be formulated as special cases of the autoregressive model. The autoregressive model is written as:

$$y = X\beta + \varepsilon \qquad \varepsilon = \lambda w \varepsilon + v$$

where
    y is the dependent variable
    X is the matrix of independent (predictor) variables
    β is the vector of typical regression coefficients.

As in a standard regression, ε is the residual variation that remains unexplained by the independent variables. It is this unexplained variance that standard statistical procedures (whether regression, t-tests, ANOVA) assume to exhibit data-point independence. In the autoregressive model, this variation is partitioned into a fully independent part (v) and a dependent part ($\lambda w \varepsilon$). The term w is a matrix of expected correlations among the data points and is typically measured from spatial distances, a social network graph, or a phylogenetic tree. The researcher's knowledge of the study system should drive the specification of w, and if several w are plausible, then model selection is usually performed to select among them. The term λ is a coefficient that multiplies against w, and it reflects the amount of autocorrelation in the residuals.

Upon inspecting the equation, it is noted that the autocorrelation parameter applies only to the residual variation after the variance explained by the independent variables is already accounted. Thus, autocorrelation models do not cause the network of expected correlations to "compete" in any way with the independent predictors, which are usually the variables of

primary theoretical interest. The purpose of the autoregressive model is to deal with the data-point nonindependence so as to obtain more-correct parameter estimates about the independent variables, which are allowed to have their full effects. A somewhat different specification of the autoregressive model does allow for the network of expected correlations to compete as an explanatory variable itself. This model, however, is unnecessary to account for the issue of nonindependence. It also has some fundamental specification problems that warrant caution in its use or at least modifying the interpretation of its output (Kissling and Carl, 2008; Ward and Gleditsch, 2002).

There are other methods to correct for nonindependence that deserve mention. One is the use of cultural areas such as those from the Standard Cross Cultural Sample (Murdock and White, 1969). This method attempts to correct for nonindependence by selecting only a few societies from each of several broader cultural zones. The method may work for some dependent variables if the sampled societies are really so distant and the variable changes fast enough that any vestiges of past inheritance or diffusion are erased. If the variable changes slowly, however, the cultural areas approach will only reduce sample size while still exhibiting nonindependence (Mace et al., 1994).

Another model that may be useful when the autoregressive is inapplicable is the random effect model, which is also called *hierarchical model* or *mixed model* (these terms can refer to slightly different specifications but are often used interchangeably). These models allow for the same regression of a dependent variable on independent variables as the autoregressive, but instead of specifying a matrix of expected correlations (*w*), they specify one or more groupings of the variables (Corrado and Fingleton, 2012). This is different from how we used the random effects model in Chapter 4 to control for the repetition of individuals' identities when the original data had been converted into pairwise distances. In this case, the data are kept in their original form (not converted into distances), and the random effects are groups of nodes that are culturally related (e.g., Germanic-speaking countries, Romance-language countries). Adding each random effect grouping adds only one parameter to the model, unlike adding grouping variables as fixed effect "factors." This is because random effect models are doing something similar to the autoregressive: They are modeling how the variations in the data are correlated for data points that have the same group membership. Often, a nested series of grouping are used to express a hierarchy of expected similarity among the data points.

When would we use a random effect model rather than an autoregressive model? One situation is when the full social network cannot be measured accurately, so we cannot confidently specify *w*, but we can bucket the data points into a set of nested groups. In this situation, we are saying that we do not know which individuals or groups share a tie or the strength of ties, but we know roughly a rank order of which sets of individuals/groups are much more likely to have ties among them.

Constructing the Expected Autocorrelation Matrix *w*

Whether one uses the autoregressive or random effects models, the correction for nonindependence will only be as good as the specification of *w*: The specification of expected autocorrelations needs to be valid, and it has to reflect the true results of the process that generates autocorrelation. To accurately construct *w*, we must move away from regarding it as a mere "nuisance" parameter in the estimation of the independent variable parameters. Cultural history and social networks are not nuisances; they profoundly are the data-generative context in which functional relationships happen. There are several techniques for constructing *w* that are well-grounded in sociocultural research. Among these are using networks, using cultural distances, and using cultural phylogenies.

Network Methods

One method to construct *w* is to measure a set of network connections among the data points that are likely determinative of the autocorrelation process (Leenders, 2002). This method was first pioneered among spatial autocorrelation analyses and later was picked up by social network theorists. Regardless of the source of data, the method typically follows the same steps. First, the network connections are measured through such modes as surveys, online communication, or spatial distances. If the network ties are weighted rather than binary, then either the weights are retained as informative of the relative amount of expected autocorrelation or a cutoff point is chosen to render the continuously valued tie weights as binary 0/1 ties. The decision to use weights or a cutoff, and where to make the cutoff, usually is based on a mix of social theory, prior literature, and empirical findings. The resulting matrix then has to be normalized to reflect the hypothesized social process. The most commonly used method is row-normalization, which divides the tie value in each cell by the sum of the row. Thus, for a binary matrix, a row with ten ties would be row-normalized so that each cell in the row was equal to 0.1, while a row with five ties would end up with cells equal to 0.2. This normalization reflects the hypothesis that the influence of each connection of an individual is proportional to the number of connections possessed by the receiving individual.

For instance, if one person knows five people, then, on average, each of the people influences the person by 20 percent. A less common alternative would be to column-normalize the matrix, which would assume influence is proportional to the number of outgoing ties. Under that schema, if one of the five people each knows 20 people, then their influence on is calculated as $1/20 = 0.05$. If the person's other four connections each know ten people, then their influences each become 0.1.

The difference between these normalizations is important. Most social processes are thought to proceed through row-normalization. Column-normalization reflects processes where the influence gets "used up" by the act of the provider providing. An example process that should be column normalized might be influence that is purchased with money. As the buyer provides the

money, the resource gets used up. Culture is socially learned information, however, and that information is not used up by its social transfer, so the most common normalization for cultural processes is row normalization

Cultural Distances

Another method to specify w is to construct a set of cultural distances among the data points based on features of the data points that are thought to develop through socially learned information (Dow et al., 1984; Matthews et al., 2016; Spolaore and Wacziarg, 2013). This method is logically linked to the network method because culture is the socially learned information that transmits among us via our network connections. The cultural distances method often has been used when the data points are groups of individuals (e.g., traditional societies, cities, or nations), and the network method has been used when the data points are individuals. This does not have to be the case, however, and one should choose the method that best estimates w given the data constraints of the study system.

One source for cultural distances are linguistic similarities among groups (Dow et al., 1984; Matthews et al., 2016; Spolaore and Wacziarg, 2009). The point of language distances is not usually to capture whether groups speak mutually intelligible languages, but rather to use language as a proxy for the many other unmeasured ways in which groups are similar or different in their "latent" cultural variations. These difficult-to-measure attitudes and beliefs are thought to influence how people negotiate scarcity and balance individual rights versus collective good.

Some attempts have been made to measure cultural attitudes and beliefs directly. The most prominent of these are the nation-level measures of Hofstede and the World Values Survey (Hofstede, 2001; Hofstede, 1984; World Values Survey Association, 2016). The World Values Survey is probably the most comprehensive cross-national values survey and has been conducted over multiple time points. The World Values Survey measures a wealth of views ranging from attitudes toward science and religion to attitudes toward women, fertility, and other ethnicities.

Although at first blush, it may appear obviously superior to measure cultural beliefs more directly, as the World Values Survey does (rather than through such proxies as linguistic divergence), further reflection reveals that the World Values Survey–attribute data carry additional concerns that much of the variation may not be cultural. For example, an extensive literature exists regarding the effects of economic development and women's education on fertility. Thus, much variation in World Values Survey responses about fertility might be driven by individual responses to local economic and educational conditions rather than by socially transmitted beliefs. Because culture is not "everything people do," but the emergent product specifically of social learning, many population-level traits in such surveys as the World Values Survey are probably only partially cultural in nature. We personally think many World Values Survey traits likely have substantial cultural components, but this is a hypothesis amenable to empirical testing. Such proxy measures as past language divergence have a different advantage

in that, although they measure culture less directly, it is difficult to see them as anything other than a measure of a cultural phenomenon.

From any of these data sources, the cultural distances method relies on measuring the pairwise distance among the data points given the language or values data being used. To avoid circularity, w generally should be estimated from separate data than those used as dependent or independent variables in the final regression model. The distance matrix then must be normalized in some manner into a matrix of expected autocorrelations. Row normalization is often used, as in the network method, but since the cultural distances are all on the same scale, there is a case to make for instead dividing by the maximal distance across the entire matrix (Matthews and Butler, 2011). The result of either method are distances between 0 and 1. These distances are then reversed into expected correlations by subtracting the whole matrix from 1.

Cultural Phylogenies

This method is very similar to the cultural distances method, but in this case, the distances among data points effectively come from an inheritance tree model for how the data points were historically interrelated (Atkinson et al., 2008, 2016; Currie et al., 2010). This method typically has been applied to group-level data (but see Dunn and Kirsner, 2011). Most typically, the cultural phylogeny is inferred from linguistic data comprising basic word lists, the most famous being the Swadish list. Basic word lists include commonly used words that we learn as small children, and the cognate patterns among these words have been shown to track well the descent relationships among languages (Nunn, 2011).

Cultural phylogenies also can be recorded directly when a written historical record exists. Matthews (2012) and Matthews et al. (2013) took this approach to constructing phylogenies of religious schisms among Christian denominations. If technological evolution is relevant to the final regression model, then cultural phylogenies often can be inferred robustly from variations among artifacts (Collard, Shennan, and Tehrani, 2006).

Accounting for cultural nonindependence often generates numerous questions from social scientists such as economists and sociologists, who might use cross-cultural data but have not been involved in the ongoing debate about Galton's problem. These questions are understandable since, especially compared with the other topics covered in this manual, correcting for nonindependence intrudes on regression modeling efforts that are conducted commonly across the social sciences. We feel that it is important for provide a response to objections for dealing with nonindependence that seem common based on our collective experience.

Controlling for nonindependence will "control away" or compete with the effects of the independent predictor variables. Especially if one uses the autoregressive model or random effects models as described in this chapter, the analyses will not control away or compete at all with the independent variables because they model the nonindependence of the residuals (Rohlf, 2006). It is the independence of the residual variation that is a statistical assumption of regression modeling.

Controlling for nonindependence makes too many assumptions about which social inheritance process is important, such as assuming it occurs along lines of shared language or religious or spatial proximity. This decision should be driven by the researcher's knowledge of the study system. Both relevant sociocultural theory and prior empirical studies can be brought to bear on this decision because the specification of w is a parametric assumption of the autoregressive model. Note that running a standard regression and assuming independence is also making a specific parametric assumption about w (that it is 0).

Statistical procedures besides regression are not affected by nonindependence. We framed the earlier discussion in terms of regression because much of standard statistics is actually some special case of the regression model (e.g., t-tests). Most other techniques, however, also assume data-point independence, including most "robust" estimators, nonparametric rank-based statistics, nonparametric bootstraps, and Bayesian implementations of regressions.

Cultural nonindependence can be ignored when it is not a variable with theoretical interest for a study. This is a tempting idea, but if the residuals of a model exhibit cultural nonindependence, then researchers should ignore that no more than they would ignore other violations of regression assumptions, such as data not being normally distributed. Researcher interest in independent variables is a good motivation to model data-point nonindependence so that the best estimates of the coefficients can be obtained. Small amounts of nonindependence often wreak havoc on models when they have only dozens or hundreds of data points, and much more havoc than even moderate amounts of heteroscedasticity or multicollinearity, which are statistical issues that often receive a great deal of attention in sociological and economic models.

Nonindependence can be ignored in experimental studies or in ones with instrumental variables. If the researcher knows ahead of time about the process that generates the nonindependence, such as the social network or other structure that specifies w, and then designs the experiment or instrumental around this, then, indeed, nonindependence can be ignored at the statistical analysis stage. This is because the sampling or experimental design effectively controls for the nonindependence. This is actually done in some social network studies in which the researchers sample a large network very sparsely. For example, one might sample 1,000 people out of 1,000,000, and then pick only one friend of each of the sampled 1,000 to study a network phenomenon. The reason this may be done is because it makes it very unlikely that one of the sampled friends is among the first 1,000 people sampled. As soon as the friends sampled start intersecting with the original sample or each other, then data-point nonindependence arises and the researchers would have to model it statistically.

Nonindependence can be ignored when the sample size is large. Sometimes sample sizes might be large enough that nonindependence can be ignored, but this probably occurs rarely in data where cultural variables have been measured. Because nonindependence does not bias or render inconsistent the estimates for the slopes (β), when sample sizes become very large, the standard deviation of the slope estimate approaches zero. Adjusting for nonindependence may

budge it ever so slightly, but even increasing an exceedingly small standard deviation by an order of magnitude will still be an exceedingly small number.

Because of this, the estimate of β will not bounce around when comparing a standard regression with an autoregression for very large data. What is a very large sample size? In our experience, it is possible that autoregression has little effect compared with standard regression in models with greater than several thousand data points. In models with dozens or hundreds of data points (e.g., most cultural data), correcting for autoregression often will change the estimate of the slopes and even can flip the sign to make a positive effect negative or vice versa (Kissling and Carl, 2008; Rohlf, 2006). This is because although the nonindependence of the data does not bias the estimated slope to be high or low, it can greatly increase the variance of the slope estimate. This causes one to incorrectly estimate a p-value that is too low, but it also means the estimated slope may be randomly too high or too low.

This effect of nonindependence on slope estimates is somewhat similar to the effect of omitted variables (Grafen, 1989). Although, in econometrics, omitted variables are usually said to truly bias the slopes, this is true only for a particular model. If we consider a set of 100 correctly specified models that are unrelated to one another, then omit one variable at random from each model, the overall effect on the remaining slope estimates will be unbiased, although for particular models, the slope will have been estimated too high or too low. A difference between the effect of nonindependence from omitted variable bias is that the former is statistically consistent while the latter is not. So, as more data points are added in a model with nonindependence, the slope estimate will be corrected closer to the real slope value, while in the context of an omitted variable, the model will express more certainty about an incorrect slope (i.e., statistical inconsistency). For most comparative cultural data, however, simply adding more data is not an option because the number of data points are constrained inherently—i.e., there are a finite number of languages, religions, countries, and subcultures.

6. Conclusion

Research methods in any mature discipline are not a mere grab-bag of tools, but rather a logically articulated set of tools capable of taking the academic or policy researcher from the raw material of data to the finished analytic product. With this in mind, we first established a theoretical notion of culture. Theory is necessary to define the scope of problem that a toolkit purports to solve. By focusing on culture as socially learned behaviors and beliefs (Chapters 1 and 2), rather than the sum of all things people do, we are able articulate a coherent set of tools to resolve most problems in contemporary analysis of cultural data. The tools described largely are applicable regardless of whether data come from direct observations of behavior, evidence from documents such as tweets or historical records, or interview and survey data. The primary collection and processing varies for each of these data types, and this manual has sought to address the analysis needed for cultural hypotheses after the primary data processing is complete.

By focusing our approach in this manner, in Chapter 3, we described CCA as a means to test whether a single culture is present among a set of sampled individuals. CCA also can be used to assess whether a set of behaviors or beliefs are part of an integrated cultural system or are independently evolving traits, as can the network analysis techniques described in Chapter 4. Chapter 4 also briefly addressed how to incorporate cultural variables with other attribute data such as economic or genetic variables and went into more detail regarding how to identify the social connections most responsible for the diffusion of cultural traits. Finally, in Chapter 5, we discussed how to identify and correct for the statistical dependence that arises from culture being shared and learned behaviors and beliefs.

These tools have developed primarily from cognitive and evolutionary anthropology, which are rather disconnected academic research traditions but that each have developed techniques specifically useful for cultural data. A few tools we have discussed, such as social network analysis and regression models, are not specific to cultural analysis but find a particular use within a broader cultural research project. From our perspective as applied social scientists and policy researchers, these tools can be arranged in a theoretically coherent picture that also solves problems for clients who rely on our research for a comprehensive empirical analysis that provides the best answer available.

More generally, applied researchers not specifically trained in cultural research methods are coming to appreciate that cultural analysis needs to be one component of many comprehensive research projects that deal with human behavior. The cultural research community should respond to this demand with methods that (1) are formal and thereby transparent and reproducible, (2) accommodate the complexity of mixed qualitative-quantitative data, and (3) provide flexible application that goes beyond any particular survey instrument, interview

protocol, or other data-collection tool. We hope that *A Manual for Cultural Analysis* provides a step toward meeting these three aims.

References

"AIDSVu," website, undated. As of April 16, 2018:
  http://www.aidsvu.org

Atkinson, Quentin D., "Phonemic Diversity Supports a Serial Founder Effect Model of
  Language Expansion from Africa," *Science*, Vol. 332, No. 6027, 2011, pp. 346–349.

Atkinson, Quentin D., Ties Coomber, Sam Passmore, Simon J. Greenhill, and Geoff Kushnick,
  "Cultural and Environmental Predictors of Pre-European Deforestation on Pacific Islands,"
  PloS One, Vol. 11, No. 5, 2016, p. e0156340.

Atkinson, Quentin D., Andrew Meade, Chris Venditti, Simon J Greenhill, and Mark Pagel,
  "Languages Evolve in Punctuational Bursts," *Science*, Vol. 319, No. 5863, 2008, pp. 588–
  588.

Bartholomew, David J., Fiona Steele, Irini Moustaki, and Jane I. Galbraith, *The Analysis and
  Interpretation of Multivariate Data for Social Scientists*, Boca Raton, Fla.: Chapman and
  Hall/CRC, 2002.

Bernard, H. Russell, *Research Methods in Anthropology: Qualitative and Quantitative
  Approaches*, 2nd ed., Thousand Oaks, Calif.: Sage, 1994.

Bernard, H. Russell, *Research Methods in Anthropology: Qualitative and Quantitative
  Approaches*, Chapter 1, 6th ed., New York: Rowman and Littlefield, 2018, pp. 1–24.

Bock, John, "Evolutionary Approaches to Population: Implications for Research and Policy,"
  Population and Environment, Vol. 21, No. 2, 1999, pp. 193–222.

Bollen, Kenneth A., "Latent Variables in Psychology and the Social Sciences," *Annual Review of
  Psychology*, Vol. 53, 2002, pp. 605–634.

Borgatti, Stephen P., *ANTHROPAC 4.0 User's Guide*, Natick, Mass.: Analytic Technologies,
  1996.

Borgatti, Stephen P., "Elicitation Techniques for Cultural Domain Analysis," *Enhanced
  Ethnographic Methods*, Vol. 3, 1999, pp. 115–151.

Borgatti, Stephen P., Martin G. Everett, and Linton C. Freeman, "Ucinet for Windows: Software
  for Social Network Analysis," *Connections*, Vol. 15, No. 1–2, 2002.

Boster, James S., "Requiem for the Omniscient Informant: There's Life in the Old Girl Yet," in
  Janet W. D. Doughtery, ed., *Directions in Cognitive Anthropology*, ed. Janet W. D.
  Dougherty, Urbana, Ill.: University of Illinois Press, 1985, pp. 177–197.

Bouckaert, Remco, Philippe Lemey, Michael Dunn, Simon J. Greenhill, Alexander V. Alekseyenko, Alexei J. Drummond, Russell D. Gray, Marc A. Suchard, and Quentin Atkinson, "Mapping the Origins and Expansion of the Indo-European Language Family," *Science*, Vol. 337, No. 6097, August 2012, pp. 957–960.

Boyd, Robert, and Peter J. Richerson, *Culture and the Evolutionary Process*, Chicago, Ill.: University of Chicago Press, 1985.

Brumann, Christoph, "Writing for Culture: Why a Successful Concept Should Not Be Discarded 1," *Current Anthropology*, Vol. 40, No. S1, 1999, pp. S1–S27.

Buchanan, Briggs, and Mark Collard, "Investigating the Peopling of North America Through Cladistic Analyses of Early Paleoindian Projectile Points," *Journal of Anthropological Archaeology*, Vol. 26, No. 3, 2007, pp. 366–393.

Buttelmann, David, Malinda Carpenter, Josep Call, and Michael Tomasello, "Enculturated Chimpanzees Imitate Rationally," *Developmental Science*, Vol. 10, No. 4, 2007, pp. F31–F38.

Byrne, Richard W., "Imitation as Behaviour Parsing," *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, Vol. 358, No. 1431, 2003, pp. 529–536.

Cavalli-Sforza, Luigi Luca, and Marcus W. Feldman, *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton, N.J.: Princeton University Press, 1981.

Chase, Philip G., "Symbolism as Reference and Symbolism as Culture," in Robin Ian MacDonald Dunbar, Chris Knight, and Camilla Power, eds., *The Evolution of Culture: An Interdisciplinary View*, Edinburgh, UK: Edinburgh University Press, 1999, pp. 34–49.

Christakis, Nicholas A., and James H. Fowler, "The Spread of Obesity in a Large Social Network over 32 Years," *New England Journal of Medicine*, Vol. 357, No. 4, July 2007, pp. 370–379.

Cohen, Dov, and Richard E. Nisbett, "Self-Protection and the Culture of Honor: Explaining Southern Violence," *Personality and Social Psychology Bulletin. Special Issue: The Self and the Collective*, Vol. 20, No. 5, October 1994, pp. 551–567.

Cohen, Dov, Joseph Vandello, and Adrian K. Rantilla, "The Sacred and the Social: Cultures of Honor and Violence," in Paul Gilbert and Bernice Andrews, eds., *Series in Affective Science. Shame: Interpersonal Behavior, Psychopathology, and Culture*, New York: Oxford University Press, 1998.

Collard, Mark, Stephen J. Shennan, and Jamshid J. Tehrani, "Branching, Blending, and the Evolution of Cultural Similarities and Differences Among Human Populations," *Evolution and Human Behavior*, Vol. 27, No. 3, May 2006, pp. 169–184.

Coop, Graham, David Witonsky, Anna Di Rienzo, and Jonathan K. Pritchard, "Using Environmental Correlations to Identify Loci Underlying Local Adaptation," *Genetics*, Vol. 185, No. 4, August 2010, pp. 1411–1423.

Corrado, Luisa, and Bernard Fingleton, "Where Is the Economics in Spatial Econometrics?" *Journal of Regional Science*, Vol. 52, No. 2, May 2012, pp. 210–239.

"Corrections and Clarifications," *Science*, Vol. 342, No. 6165, December 2013, p. 1446.

Currie, Thomas E., Simon J. Greenhill, Russell D. Gray, Toshikazu Hasegawa, and Ruth Mace, "Rise and Fall of Political Complexity in Island Southeast Asia and the Pacific," *Nature*, Vol. 467, No. 7317, October 2010, pp. 801–804.

D'Andrade, Roy G., *The Development of Cognitive Anthropology*, Cambridge, UK: Cambridge University Press, 1995.

de Nooy, Wouter, "Networks of Action and Events over Time: A Multilevel Discrete-Time Event History Model for Longitudinal Network Data," *Social Networks*, Vol. 33, No. 1, 2011, pp. 31–40.

de Waal, Frans, *The Ape and the Sushi Master: Cultural Reflections by a Primatologist*, New York: Basic Books, 2001.

Dean, Lewis G., Gill L. Vale, Kevin N. Laland, Emma Flynn, and Rachel L. Kendal, "Human Cumulative Culture: A Comparative Perspective," *Biological Reviews*, Vol. 89, No. 2, May 2014, pp. 284–301.

Dekker, David, David Krackhardt, and Tom A. B. Snijders, "Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions," *Psychometrika*, Vol. 72, No. 4, December 2007, pp. 563–581.

DiMaggio, Paul, "Culture and Cognition," *Annual Review of Sociology*, Vol. 23, No. 1, 1997, pp. 263–287.

DiStefano, Christine, Min Zhu, and Diana Mindrila, "Understanding and Using Factor Scores: Considerations for the Applied Researcher," *Practical Assessment, Research, and Evaluation*, Vol. 14, No. 20, October 2009, pp. 1–11.

Dittrich, Dino, Roger Th.A.J. Leenders, and Joris Mulder, "Bayesian Estimation of the Network Autocorrelation Model," *Social Networks*, Vol. 48, 2017, pp. 213–236.

Dow, Malcolm M., Michael L. Burton, Douglas R. White, and Karl P. Reitz, "Galton's Problem as Network Autocorrelation," *American Ethnologist*, Vol. 11, No. 4, November 1984, pp. 754–770.

Dressler, William W., and James R. Bindon, "The Health Consequences of Cultural Consonance: Cultural Dimensions of Lifestyle, Social Support, and Arterial Blood Pressure in an African American Community," *American Anthropologist*, Vol. 102, No. 2, June 2000, pp. 244–260.

Dunn, John C., and Kim Kirsner, "The Search for HMAS Sydney II: Analysis and Integration of Survivor Reports," *Applied Cognitive Psychology*, Vol. 25, No. 4, July/August 2011, pp. 513–527.

Dunteman, George H., *Principal Components Analysis*, Newbury Park, Calif.: Sage Publications, Inc., 1989.

Durham, William H., *Coevolution: Genes, Culture, and Human Diversity*, Redwood City, Calif.: Stanford University Press, 1991.

Elmore, Kimberlee Ann, "Southern Discomfort: AIDS Stigmatization in Wilmington, North Carolina," *Southeastern Geographer*, Vol. 46, No. 2, 2006, pp. 215–230.

Felsenstein, Joseph, "Phylogenies and the Comparative Method," *The American Naturalist*, Vol. 125, No. 1, January 1985, pp. 1–15.

Fincher, Corey L., Randy Thornhill, Damian R. Murray, and Mark Schaller, "Pathogen Prevalence Predicts Human Cross-Cultural Variability in Individualism/Collectivism," *Proceedings of the Royal Society of London B: Biological Sciences*, Vol. 275, No. 1640, June 7, 2008, pp. 1279–1285.

Fogarty, Laurel, Luke Rendell, and Kevin N. Laland, "Mental Time Travel, Memory and the Social Learning Strategies Tournament," *Learning and Motivation*, Vol. 43, No. 4, November 2012, pp. 241–246.

Fokoue, Ernest, and Zichen Ma, "Modern Multivariate Methods for Accurate Dialect Classification," *RIT Scholar Works*, 2013.

Franz, Mathias, and Luke J. Matthews, "Social Enhancement Can Create Adaptive, Arbitrary and Maladaptive Cultural Traditions," *Proceedings of the Royal Society of London B: Biological Sciences*, Vol. 277, No. 1698, November 2010, p. rspb20100705.

Galef, Bennett G., and Luc-Alain Giraldeau, "Social Influences on Foraging in Vertebrates: Causal Mechanisms and Adaptive Functions," *Animal Behaviour*, Vol. 61, No. 1, January 2001, pp. 3–15.

Garcia, Justin R., James MacKillop, Edward L. Aller, Ann M. Merriwether, David Sloan Wilson, and J. Koji Lum, "Associations Between Dopamine D4 Receptor Gene Variation with Both Infidelity and Sexual Promiscuity," *PloS One*, Vol. 5, No. 11, November 2010, p. e14162.

Grafen, Alan, "The Phylogenetic Regression," *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 326, No. 1233, December 21, 1989, pp. 119–157.

Grosjean, Pauline, "A History of Violence: The Culture of Honor and Homicide in the U.S. South," *Journal of the European Economic Association*, Vol. 12, No. 5, 2014, pp. 1285–1316. As of May 22, 2018:
http://dx.doi.org/10.1111/jeea.12096

Hamilton, William D., "Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics," *Biosocial Anthropology*, Vol. 53, 1975, pp. 133–155.

Handwerker, W. Penn, "The Modern Demographic Transition: An Analysis of Subsistence Choices and Reproductive Consequences," *American Anthropologist*, Vol. 88, 1986, pp. 400–417.

Handwerker, W. Penn, *Quick Ethnography: A Guide to Rapid Multi-Method Research*, Walnut Creek, Calif..: AltaMira Press, 2001.

Handwerker, W. Penn, "The Construct Validity of Cultures: Cultural Diversity, Culture Theory, and a Method for Ethnography," *American Anthropologist*, Vol. 104, No. 1, March 2002, pp. 106–122.

Handwerker, W. Penn, and Stephen P. Borgatti, "Reasoning with Numbers," in H. Russell Bernard, ed., *Handbook of Methods in Cultural Anthropology*, Walnut Creek, Calif.: AltaMira Press, 1998, pp. 549–593.

Handwerker, W. Penn, and Danielle F. Wozniak, "Sampling Strategies for the Collection of Cultural Data: An Extension of Boas's Answer to Galton's Problem 1," *Current Anthropology*, Vol. 38, No. 5, December 1997, pp. 869–875.

Harden, K. Paige, "Genetic Influences on Adolescent Sexual Behavior: Why Genes Matter for Environmentally Oriented Researchers," *Psychological Bulletin*, Vol. 140, No. 2, 2014, p. 434.

Harmon, Luke J., and Richard E. Glor, "Poor Statistical Performance of the Mantel Test in Phylogenetic Comparative Analyses," *Evolution*, Vol. 64, No. 7, July 2010, pp. 2173–2178.

Harris, Marvin, *Cultural Materialism: The Struggle for a Science of Culture*, Walnut Creek, Calif.: AltaMira Press, 2001.

Harris, Marvin, and Eric B. Ross, *Death, Sex, and Fertility: Population Regulation in Preindustrial and Developing Societies*, New York: Columbia University Press, 1987.

Henrich, Joseph, and Robert Boyd, "Why People Punish Defectors: Weak Conformist Transmission Can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas," *Journal of Theoretical Biology*, Vol. 208, No. 1, January 2001, pp. 79–89.

Henrich, Joseph, and Francisco J. Gil-White, "The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission," *Evolution and Human Behavior*, Vol. 22, No. 3, 2002, pp. 165–196.

Henrich, Joseph, and Natalie Henrich, "The Evolution of Cultural Adaptations: Fijian Food Taboos Protect Against Dangerous Marine Toxins," *Proceedings of the Royal Society B: Biological Sciences*, Vol. 277, No. 1701, December 22, 2010, pp. 3715–3724.

Hobbes, Thomas, *Hobbes's Leviathan*, reprinted from the edition of 1651, Oxford, UK: Clarendon Press, 1909.

Hofstede, Geert, *Culture's Consequences: International Differences in Work-Related Values*, Vol. 5, *Cross-Cultural Research and Methodology Series*, Newbury Park, Calif.: Sage Publications, Inc., 1984.

Hofstede, Geert H, *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, Thousand Oaks, Calif.: Sage Publications, Inc., 2001.

Hoppitt, Will, and Kevin N. Laland, "Social Processes Influencing Learning in Animals: A Review of the Evidence," *Advances in the Study of Behavior*, Vol. 38, 2008, pp. 105–165.

Horner, Victoria, and Andrew Whiten, "Causal Knowledge and Imitation/Emulation Switching in Chimpanzees (Pan troglodytes) and Children (Homo sapiens)," *Animal Cognition*, Vol. 8, No. 3, July 2005, pp. 164–181.

Inglehart, Ronald, and Wayne E. Baker, "Modernization, Cultural Change, and the Persistence of Traditional Values," *American Sociological Review*, Vol. 65, No. 1, February 2000, pp. 19–51.

Inglehart, Ronald, and Christian Welzel, "Changing Mass Priorities: The Link Between Modernization and Democracy," *Perspectives on Politics*, Vol. 8, No. 2, June 2010, pp. 551–567.

Karimov, Rouslan, and Luke J. Matthews, "A Simulation Assessment of Methods to Infer Cultural Transmission on Dark Networks," *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, Vol. 14, No. 1, 2017, p. 1548512916679900.

Kaufman, Leonard, and Peter J. Rousseau, *Finding Groups in Data: An Introduction to Cluster Analysis*, Vol. 344, Hoboken, N.J.: John Wiley and Sons, 1990.

Kennedy, David P., *Gender, Culture Change, and Fertility Decline in Honduras: An Investigation in Anthropological Demography*, doctoral dissertation, Gainesville, Fla.: University of Florida, 2002. As of April 16, 2018:
http://etd.fcla.edu/UF/UFE0000551/kennedy_d.pdf

Kennedy, David P., "Scale Adaptation and Ethnography," *Field Methods*, Vol. 17, No. 4, November 1, 2005, pp. 412–431.

Kennedy, David P., Suzanne L. Wenzel, Ryan Brown, Joan S. Tucker, and Daniela Golinelli, "Unprotected Sex Among Heterosexually Active Homeless Men: Results from a Multi-Level Dyadic Analysis," *AIDS Behavior*, Vol. 17, No. 5, June 2013, pp. 1655–1667.

Kissling, W. Daniel, and Gudrun Carl, "Spatial Autocorrelation and the Selection of Simultaneous Autoregressive Models, *Global Ecology and Biogeography*, Vol. 17, No. 1, January 2008, pp. 59–71.

Laland, Kevin N., and William Hoppitt, "Do Animals Have Culture?" *Evolutionary Anthropology: Issues, News, and Reviews*, Vol. 12, No. 3, 2003, pp. 150–159.

Laland, Kevin N., and Jeremy R. Kendal, "What the Models Say About Social Learning," *The Biology of Traditions: Models and Evidence*, in Dorothy M. Fragaszy and Susan Perry, eds., Cambridge, UK: Cambridge University Press, 2003, pp. 33–55.

Laland, Kevin N., and Luke Rendell, "Cultural Memory," *Current Biology*, Vol. 23, No. 17, September 9, 2013, pp. R736–R740.

Langness, Lewis L., *The Study of Culture*, Novato, Calif.: Chandler Sharp Publishers, 1987.

Leenders, Roger Th.A.J., "Modeling Social Influence Through Network Autocorrelation: Constructing the Weight Matrix," *Social Networks*, Vol. 24, No. 1, January 2002, pp. 21–47.

Lichtenstein, Bronwen, "Domestic Violence, Sexual Ownership, and HIV Risk in Women in the American Deep South," *Social Science and Medicine*, Vol. 60, No. 4, 2005, pp. 701–714.

Lichtenstein, Bronwen, Edward W. Hook, and Amit K. Sharma, "Public Tolerance, Private Pain: Stigma and Sexually Transmitted Infections in the American Deep South," *Culture, Health, and Sexuality*, Vol. 7, No. 1, 2005, pp. 43–57.

Lycett, Stephen J., and Noreen von Cramon-Taubadel, "Acheulean Variability and Hominin Dispersals: A Model-Bound Approach," *Journal of Archaeological Science*, Vol. 35, No. 3, 2008, pp. 553–562.

Lyons, Russell, "The Spread of Evidence-Poor Medicine via Flawed Social-Network Analysis," *Statistics, Politics, and Policy*, Vol. 2, No. 1, 2011.

Mace, Ruth, Mark Pagel, John R. Bowen, Biman Kumar Das Gupta, Keith F. Otterbein, Mark Ridley, Thomas Schweizer, and Eckart Voland, "The Comparative Method in Anthropology [and Comments and Reply]," *Current Anthropology*, Vol. 35, No. 5, December 1994, pp. 549–564.

MacLean, Evan L., Luke J. Matthews, Brian A. Hare, Charles L. Nunn, Rindy C. Anderson, Filippo Aureli, Elizabeth M. Brannon, Josep Call, Christine M. Drea, Nathan J. Emery,

Daniel B. M. Haun, Esther Herrmann, Lucia F. Jacobs, Michael L. Platt, Alexandra G. Rosati, Aaron A. Sandel, Kara K. Schroepfer, Amanda M. Seed, Jingzhi Tan, Carel P. van Schaik, and Victoria Wobber, "How Does Cognition Evolve? Phylogenetic Comparative Psychology," *Animal Cognition*, Vol. 15, 2012, pp. 223–238.

Mason, Karen Oppenheim, "Gender and Family Systems in the Fertility Transition," *Population and Development Review*, Vol. 27, Supplement: Global Fertility Transition, 2001, pp. 160–176.

Mason, Karen Oppenheim, and Anju Malhotra Taj, "Differences Between Women's and Men's Reproductive Goals in Developing Countries," *Population and Development Review*, Vol. 13, No. 4, December 1987, pp. 611–638.

Matthews, Luke J., "Intragroup Behavioral Variation in White-Fronted Capuchin Monkeys (Cebus albifrons): Mixed Evidence for Social Learning Inferred from New and Established Analytical Methods," *Behaviour*, Vol. 146, No. 3, March 2009, pp. 295–324.

Matthews, Luke J., "The Recognition Signal Hypothesis for the Adaptive Evolution of Religion," *Human Nature*, Vol. 23, No. 2, June 2012, pp. 218–249.

Matthews, Luke J., and Paul M. Butler, "Novelty-Seeking DRD4 Polymorphisms Are Associated with Human Migration Distance Out-of-Africa After Controlling for Neutral Population Gene Structure," *American Journal of Physical Anthropology*, Vol. 145, No. 3, July 2011, pp. 382–389.

Matthews, Luke J., Jeffrey Edmonds, Wesley J. Wildman, and Charles L. Nunn, "Cultural Inheritance or Cultural Diffusion of Religious Violence? A Quantitative Case Study of the Radical Reformation," *Religion, Brain, and Behavior*, Vol. 3, No. 1, 2013, pp. 3–15.

Matthews, Luke J., Sam Passmore, Paul M. Richard, Russell D. Gray, and Quentin D. Atkinson, "Shared Cultural History as a Predictor of Political and Economic Changes Among Nation States," *PloS One*, Vol. 11, No. 4, 2016, p. e0152979.

McKeown, Bruce, and Dan B. Thomas, *Q Methodology*, 2nd ed., Vol. 66, Newbury Park, Calif.: Sage Publications, Inc., 2013.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, Vol. 27, No. 1, 2001, pp. 415–444.

Miller, Warren B., "Childbearing Motivation and Its Measurement," *Journal of Biosocial Science*, Vol. 27, No. 4, October 1995, pp. 473–487.

Mizruchi, Mark S., and Eric J. Neuman, "The Effect of Density on the Level of Bias in the Network Autocorrelation Model," *Social Networks*, Vol. 30, No. 3, July 2008, pp. 190–200.

Murdock, George P., and Douglas R. White, "Standard Cross-Cultural Sample," *Ethnology*, Vol. 8, No. 4, 1969, pp. 329–369.

Naroll, Raoul, "Two Solutions to Galton's Problem," *Philosophy of Science*, Vol. 28, No. 1, January 1961, pp. 15–39.

Nash, Dennison, *A Little Anthropology*, 3rd ed., London: Pearson College Division, 1998.

Neter, John, Michael H. Kutner, Christopher J. Nachtsheim, and William Wasserman, *Applied Linear Statistical Models*, Vol. 4, Chicago, Ill.: Irwin, 1996.

Neuman, Eric J., and Mark S. Mizruchi, "Structure and Bias in the Network Autocorrelation Model," *Social Networks*, Vol. 32, No. 4, October 2010, pp. 290–300.

Nielsen, Mark, "The Imitative Behaviour of Children and Chimpanzees: A Window on the Transmission of Cultural Traditions," *Revue de primatologie*, Vol. 1, 2009.

Nunn, Charles L., *The Comparative Approach in Evolutionary Anthropology and Biology*, Chicago, Ill.: University of Chicago Press, 2011.

Nunn, Charles L., Peter H. Thrall, Kevin Bartz, Tirthankar Dasgupta, and Christophe Boesch, "Do Transmission Mechanisms or Social Systems Drive Cultural Dynamics in Socially Structured Populations?" *Animal Behaviour*, Vol. 77, No. 6, June 2009, pp. 1515–1524.

O'Malley, A. James, and Nicholas A. Christakis, "Longitudinal Analysis of Large Social Networks: Estimating the Effect of Health Traits on Changes in Friendship Ties," *Statistics in Medicine*, Vol. 30, No. 9, April 2011, pp. 950–964.

O'Malley, A. James, and Peter V. Marsden, "The Analysis of Social Networks," *Health Services and Outcomes Research Methodology*, Vol. 8, No. 4, December 2008, pp. 222–269.

Páez, Antonio, Darren M. Scott, and Erik Volz, "Weight Matrices for Social Influence Analysis: An Investigation of Measurement Errors and Their Effect on Model Identification and Estimation Quality," *Social Networks*, Vol. 30, No. 4, October 2008, pp. 309–317.

Pagel, Mark, "Inferring the Historical Patterns of Biological Evolution," *Nature*, Vol. 401, No. 6756, October 1999, pp. 877–884.

Paschou, Peristera, Elad Ziv, Esteban G. Burchard, Shweta Choudhry, William Rodriguez-Cintron, Michael W. Mahoney, and Petros Drineas, "PCA-Correlated SNPs for Structure Identification in Worldwide Human Populations," *PLoS Genetics*, Vol. 3, No. 9, September 2007, p. e160.

Pezeshki, Mohammad Z., Bahram Zeighami, and Warren B. Miller, "Measuring the Childbearing Motivation of Couples Referred to the Shiraz Health Center for Premarital Examinations," *Journal of Biosocial Science*, Vol. 37, No. 1, January 2005, pp. 37–53.

Potts, Richard, "Variability Selection in Hominid Evolution," *Evolutionary Anthropology: Issues, News, and Reviews*, Vol. 7, No. 3, 1998, pp. 81–96.

Provenza, Frederick D, "Tracking Variable Environments: There Is More Than One Kind of Memory," *Journal of Chemical Ecology*, Vol. 21, No. 7, July 1995, pp. 911–923.

Radcliffe-Brown, Alfred Reginald, "Three Tribes of Western Australia," *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, Vol. 43, 1913, pp. 143–194.

Richerson, Peter J., and Robert Boyd, "Not by Genes Alone: How Culture Transformed Human Evolution," Chicago, Ill.: University of Chicago Press, 2005.

Rohlf, F. James, "A Comment on Phylogenetic Correction," *Evolution*, Vol. 60, No. 7, 2006, pp. 1509–1515.

Romney, A. Kimball, "Cultural Consensus as a Statistical Model," *Current Anthropology*, Vol. 40, Supplement, February 1999, pp. S93–S115.

Romney, A. Kimball, Susan C. Weller, and William H. Batchelder, "Culture as Consensus: A Theory of Culture and Informant Accuracy," *American Anthropologist*, Vol. 88, No. 2, 1986, pp. 313–338.

Ryan, Gery W., Justin M. Nolan, and P. Stanley Yoder, "Successive Free Listing: Using Multiple Free Lists to Generate Explanatory Models," *Field Methods*, Vol. 12, No. 2, 2000, pp. 83–107.

Sahlins, Marshall David, *Tribesmen*, Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Sangree, Walter H., "Pronatalism and the Elderly in Tiriki, Kenya," in Thomas S. Weisner, Candice Bradley, and Philip L. Kilbride, eds., *African Families and the Crisis of Social Change*, Westport, Conn.: Bergin and Garvey, 1997, pp. 184–207.

Service, Elman R., *The Hunters*, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1966.

Shalizi, Cosma Rohilla, and Andrew C. Thomas, "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies," *Sociological Methods, and Research*, Vol. 40, No. 2, 2011, pp. 211–239.

Shen, Dan, Haipeng Shen, Hongtu Zhu, and J. S. Marron, "The Statistics and Mathematics of High Dimension Low Sample Size Asymptotics," *Statistica Sinica*, Vol. 26, No. 4, October 2016, p. 1747.

Shennan, Stephen, and Mark Collard, "Investigating Processes of Cultural Evolution," *The Evolution of Cultural Diversity: Phylogenetic Approach*, in Ruth Mace, Clare J. Holden, and Stephen Shennan, eds., Oxford, UK: Cavendish Publishing, 2005, pp. 133–164.

Spolaore, Enrico, and Romain Wacziarg, "The Diffusion of Development," *The Quarterly Journal of Economics*, Vol. 124, No. 2, 2009, pp. 469–529.

Spolaore, Enrico, and Romain Wacziarg, "How Deep Are the Roots of Economic Development?" *Journal of Economic Literature*, Vol. 51, No. 2, June 2013, pp. 325–369.

Thornton, Alex, and Katherine McAuliffe, "Teaching in Wild Meerkats," *Science*, Vol. 313, No. 5784, July 14, 2006, pp. 227–229.

Tomasello, Michael, "The Human Adaptation for Culture," *Annual Review of Anthropology*, Vol. 28, No. 1, October 1999, pp. 509–529.

Tomasello, Michael, and Josep Call, *Primate Cognition*, Oxford, UK: Oxford University Press, 1997.

Tylor, Edward Burnett, *Primitive Culture: Researches into the Development of Mythology, Philosophy, Religion, Art, and Custom*, Vol. 2, London: John Murray, Albemarle Street, 1871.

VanderWeele, Tyler J., "Sensitivity Analysis for Contagion Effects in Social Networks," *Sociological Methods and Research*, Vol. 40, No. 2, 2011, pp. 240–255.

VanderWeele, Tyler J., Elizabeth L. Ogburn, and Eric J. Tchetgen Tchetgen, "Why and When 'Flawed' Social Network Analyses Still Yield Valid Tests of No Contagion," *Statistics, Politics and Policy*, Vol. 3, No. 1, February 2014, p. 4.

Ward, Michael D., and Kristian Skrede Gleditsch, "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace," *Political Analysis*, Vol. 10, No. 3, Summer 2002, pp. 244–260.

Ward, Michael D., and Kristian Skrede Gleditsch, "An Introduction to Spatial Regression Models in the Social Sciences," *SAGE Publications*, Vol. 155, 2008.

Weller, Susan C., "Structured Interviewing and Questionnaire Construction," in H. Russell Bernard and Clarence C. Gravlee, eds., *Handbook of Methods in Cultural Anthropology*, 2nd ed., New York: Rowman and Littlefield, 1998, pp. 365–409.

Weller, Susan C., "Cultural Consensus Theory: Applications and Frequently Asked Questions," *Field Methods*, Vol. 19, No. 4, 2007, pp. 339–368.

Weller, Susan C., and A. Kimball Romney, *Systematic Data Collection*, Vol. 10, *Qualitative Research Methods*, Newbury Park, Calif.: Sage Publications, 1988.

Welsch, Robert L., John Terrell, and John A. Nadolski, "Language and Culture on the North Coast of New Guinea," *American Anthropologist*, Vol. 94, No. 3, September 1992, pp. 568–600.

Whiten, Andrew, "Imitation of the Sequential Structure of Actions by Chimpanzees (Pan troglodytes)," *Journal of Comparative Psychology*, Vol. 112, No. 3, 1998, p. 270.

Whiten, Andrew, Deborah M. Custance, Juan-Carlos Gomez, Patricia Teixidor, and Kim A. Bard, "Imitative Learning of Artificial Fruit Processing in Children (Homo sapiens) and Chimpanzees (Pan troglodytes)," *Journal of Comparative Psychology*, Vol. 110, No. 1, 1996, p. 3.

Whiten, Andrew, Jane Goodall, William C McGrew, Toshisada Nishida, Vernon Reynolds, Yukimaru Sugiyama, Caroline E. G. Tutin, Richard W. Wrangham, and Christophe Boesch, "Cultures in Chimpanzees," *Natur*e, Vol. 399, No. 6737, June 17, 1999, pp. 682–685.

Whiten, A., and R. Ham, "On the Nature and Evolution of Imitation in the Animal Kingdom: Reappraisal of a Century of Research," *Advances in the Study of Behavior*, Vol. 21, 1992, p. 239.

Wimmer, Andreas, and Kevin Lewis, "Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook," *American Journal of Sociology*, Vol. 116, No. 2, September 2010, pp. 583–642.

World Values Survey Association, "World Values Survey Wave 6 2010–2014 Official Aggregate v. 20150418," Aggregate File Producer: Asep/JDS, Madrid, Spain, 2016. As of April 17, 2018:
http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp

Yata, Kazuyoshi, and Makoto Aoshima, "Effective PCA for High-Dimension, Low-Sample-Size Data with Singular Value Decomposition of Cross Data Matrix," *Journal of Multivariate Analysis*, Vol. 101, No. 9, October 2010, pp. 2060–2077.