Chapter 1. Introduction: Bayesian Reasoning for Qualitative Research

Tasha Fairfield and A.E. Charman
Social Inquiry and Bayesian Inference: Rethinking Qualitative Research
Forthcoming, Cambridge University Press

The way we intuitively approach qualitative case research is similar to how we read detective novels or murder mysteries. We want to explain what happened—who killed Samuel Ratchett on the Orient Express, or how democracy emerged in South Africa. We consider different hypotheses along the way, drawing on our own ingenuity and the literature we have read—whether other Agatha Christie mysteries, or theories of regime change. As we gather evidence and discover new clues, we continually revise our assessment about which hypothesis provides the best explanation. Bayesianism provides a natural framework to govern how we should adjust our degree of belief in the truth of a hypothesis—e.g., *A lone gangster slipped onboard the train and killed Samuel Ratchett as revenge for being swindled*, or *Mobilization from below drove democratization in South Africa by altering economic elites' regime preferences* (Wood 2001)—given our previous knowledge and the new information that we learn during our research.

To illustrate how Bayesian updating works intuitively, consider the following more extended example. During the Latin American debt crisis in the 1980s, an intriguing phenomenon arose in which new presidents who had explicitly promised voters that they would *not* implement austerity measures nevertheless imposed harsh neoliberal reforms after taking office. Stokes (2001) entertains two alternative explanations for why these presidents violated their policy mandates. The first hypothesis proposes that presidents sought to represent the people's best interests and realized that austerity was the only way to fix the economy, despite voters' trepidations. The second hypothesis proposes instead that presidents were motived by opportunities for their own private financial gain, for example, kickbacks and bribes from business. Now consider the case of Venezuela's President Pérez, one of Latin America's policy-mandate violators—why did he opt to impose "neoliberalism by surprise"? Given whatever relevant information we know, we might strongly favor one of Stokes' (2001) hypotheses over the other, we might weakly favor one of them, or we might simply be indifferent. In Bayesian terminology, our initial view about the plausibility of a hypothesis is called the *prior probability*.

Now suppose we learn the following information. According to multiple sources, Pérez learned how bad Venezuela's economic situation was only after taking office. And advisors reported that Pérez was paying close attention to developments in Peru, where he saw that the heterodox economic stabilization policies that his friend President García had implemented were not working (Stokes 2001:68-69). In light of this new information,[1] we might change our view about which hypothesis—representation, or rent seeking—is more plausible. The Bayesian language for this revised view, which takes into account both our previous knowledge and the new evidence, is the *posterior probability.*

Figure 1 displays the results of this intuitive Bayesian updating exercise as conducted with roughly 80 participants at the Syracuse Institute for Qualitative and Multi-Method Research. Not surprisingly, given that the students came from a wide range of backgrounds and subfields, prior views were distributed across the spectrum, from strongly favoring the representation hypothesis to strongly favoring the rent-seeking hypothesis. But after considering the Venezuelan case evidence, beliefs tended to shift in favor of the representation hypothesis, with most participants converging on a weak preference for that explanation over the rent-seeking alternative.

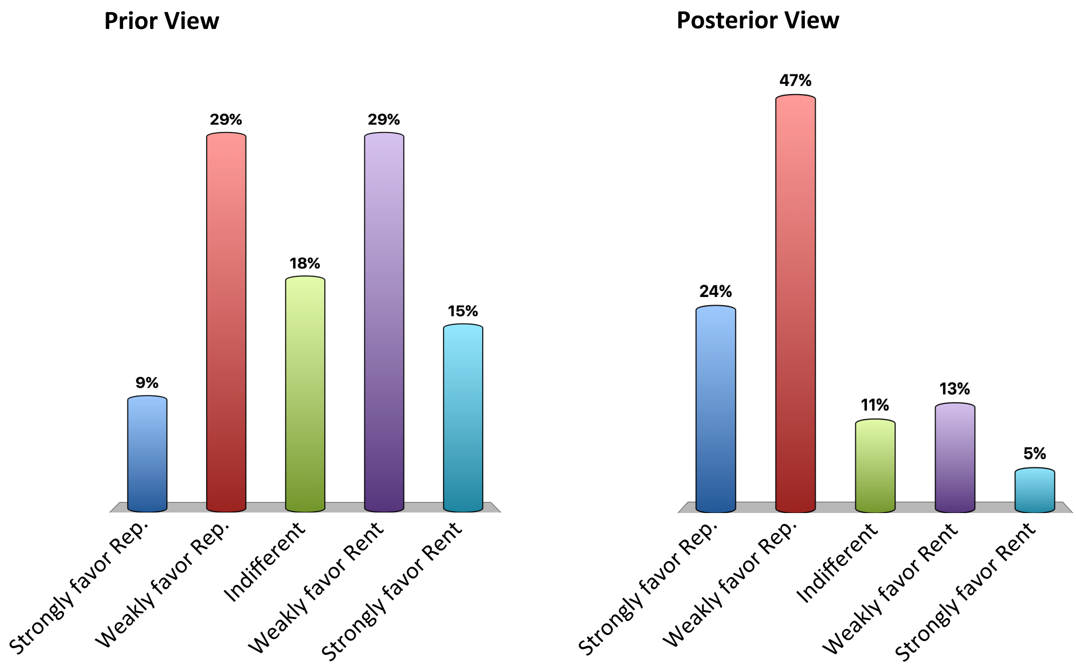

FIG. 1 **Intuitive Bayesian Updating.** Representation vs. Rent-seeking Hypotheses, drawing on Stokes (2001).

[1] We assume for the sake of illustration that readers are not already familiar with these particulars of the Venezuelan case. If these details are already known, they should inform the prior probability of the hypotheses, and coming across this information again would not change our initial views.

This book aims to improve the way we evaluate the import of qualitative evidence and conduct case study research by drawing on insights from logical Bayesianism, an inferential framework originally developed for the natural sciences. Logical Bayesianism conceptualizes probability as the rational degree of belief that we should hold in the truth of a proposition—e.g., a causal hypothesis—given the information we possess, which is inevitably limited. Bayesian probability, and Bayes' rule in particular, provide a rigorous framework for reasoning under conditions of uncertainty and incomplete information. In principle, Bayesian probability provides a unified approach for all scientific inquiry. From this perspective, we reexamine central debates on the logic of inference, research design, and analytic transparency in qualitative research.

Bayesianism is enjoying a revival across many fields, from astronomy and data science to macroeconomics and political polling, and it has much to offer qualitative social science. First and foremost, Bayesianism provides a solid methodological foundation for causal inference in case studies and qualitative research more broadly. In fact, we will argue that Bayesianism is the only sound approach for causal analysis with qualitative, non-stochastic data. Second, learning the basic principles of Bayesian reasoning can help leverage and improve intuition. Bayesinaism mirrors the way we naturally approach inference and implicitly underpins much of our common sense, but it also helps us avoid cognitive biases that can lead to sloppy reasoning. Even if scholars choose not to explicitly apply some of the more formal Bayesian techniques that we will introduce, learning and practicing these techniques can nevertheless help improve inferential judgments when conducting traditional, narrative-form case study analysis.

Third, Bayesianism facilitates consensus-building and promotes knowledge accumulation, by providing a clear framework for scrutinizing inferences and pinpointing sources of disagreement. Baysianism directs us to ask if disagreements stem from different initial views regarding which explanations are most plausible, given that everyone brings very different background knowledge to the discussion, and/or if scholars are assessing the inferential weight of the evidence differently, why their assessments diverge, and whether one line of reasoning about the evidence is more justifiable and compelling than another.

Finally, following Bayesian principles more explicitly enhances research transparency, which is a growing concern in light of the "replication crisis" in social science. We contend that the overarching concern in all scientific inquiry should be *reliability* of inference, which encompasses but extends beyond the notion of replication. Assessing reliability entails asking how much confidence we can justifiably hold in our conclusions. Bayesian probability is ideally suited to this task, because it provides the natural language for evaluating uncertainty.

1. PLACING OUR APPROACH IN PERSPECTIVE

Bayesian probability holds the potential to clarify and remap methodological debates within political science and to fully substantiate the importance of qualitative research vis-à-vis large-N statistical analysis and other quantitative research traditions. Despite recent innovations in process-tracing (e.g., Bennett 2015), applying Bayesian probability in qualitative research remains a frontier that has not been definitively addressed. Qualitative methodologists have not yet recognized the full potential and ramifications of Bayesian probability; in fact, many misconceptions persist in efforts to apply Bayesianism in case study research. Nor have quantitative social scientists who view Bayesian analysis as a powerful technical tool (e.g., Iversen 1984, Jackman 2009, Gelman et al. 2013) fully realized the broader implications of Bayesian probability for inference and research design, although we build here on pioneering works by Western and Jackman (1994) and Western (2001) that set out some key foundational principles. Humphreys and Jacobs (2015, forthcoming) break new ground by exploring implications for research designs that combine within-case clues and cross-case datasets within a Bayesian framework, yet the logical Bayesian perspective we espouse leads to distinct recommendations regarding how best to operationalize test strength, select cases, and iterate between theory building and theory testing.

Accordingly, we seek to contribute to three areas of methodology: process tracing, qualitative methods more broadly, and multi-method research. While Bayesian justifies many commonsense practices, our perspective also introduces ideas that go against some established ways of thinking in many realms of social science. However, in the spirit of *Rethinking Social Inquiry* (Brady and Collier 2010), we hope to broaden the scope of debate on how to make inference more rigorous and to help bridge the gap between qualitative and quantitative methods in a way that more firmly establishes the value of qualitative research.

1.1. Process Tracing

A growing movement within political science has identified Bayesian probability as the methodological foundation of process tracing, which Bennett and Checkel (2015:4) define broadly as "the use of evidence from within a case to make inferences about causal explanations of that case." As part of an initiative to establish process tracing as a rigorous method, the literature has moved from informal analogies to Bayesianism (McKeown 1999, Bennett 2008, Beach and Pedersen 2013) toward efforts to more formally apply Bayesian analysis in qualitative research (Rohlfing 2013, Bennett 2015, Humphreys and Jacobs 2015, Fairfield and Charman 2017). But whereas Bayesian statistical techniques have been successfully elaborated for large-N quantitative research, efforts to apply Bayesian probability in qualitative case research are still under development. Most efforts to formalize Bayesian process tracing have examined only a few illustrative pieces of evidence (Rohlfing 2013, Bennett 2015) and/or have included only highly simplified process-tracing clues (Humphreys and Jacobs 2015). Moreover, we have found that a number of important conceptual and technical points have been overlooked, incorrectly handled, or inadequately addressed within the literature that is innovating in this terrain.[2]

We aim to advance the process-tracing literature by providing a careful exposition of the foundations of Bayesian "probability as extended logic" from the physical sciences (Jeffreys 1939, Cox 1961, Jaynes 2003, Gregory 2005), which contrasts with the more "subjective" treatments of Bayesianism in most philosophy of science and medical testing literature that inform much of the existing work on Bayesian process-tracing.[3] We elaborate concrete guidelines to help scholars avoid potential pitfalls when endeavoring to apply Bayesian reasoning in process tracing, and we illustrate how to proceed when working with the kinds of complex, diverse, and nuanced real-world evidence that scholars gather during in-depth fieldwork and archival research.

1.2. Qualitative Methods

From a broader qualitative methods perspective, Bayesianism yields two major payoffs. First, it provides the rigorous foundation needed to definitively explicate and legitimate qualitative

[2] Problems of this sort are also pervasive in qualitative methods literature that invokes Bayesianism more informally (e.g., Beach and Pedersen 2016).
[3] See Section 3.1.

research. Earlier efforts to understand qualitative case studies and their relation to large-N quantitative research from a frequentist statistical perspective (e.g., King, Keohane, and Verba 1994) inevitably attributed a subordinate role to the former; there simply is no principled rationalization for small-N qualitative research within frequentism. If applied in strict accordance with its foundational tenets, frequentist techniques can only be used to analyze stochastic data, and large, independent samples are often considered critical for accurate inference. In contrast, Bayesianism narrows the divide between qualitative and quantitative research, because all inference in principle proceeds in the same manner, by applying Bayes' theorem and the other rules of probability theory. Furthermore, in the words of Pierre Simon Laplace (1814), Bayesian probability is essentially "common sense reduced to calculation." As such, it validates many intuitively-sensible practices that have long characterized qualitative research but are often discouraged by frequentist-oriented disciplinary norms, including non-random case selection and iteration between theory development and data analysis.

Second and relatedly, Bayesianism provides a simple, intuitive alternative to the wide range of inferential approaches advocated and debated within the qualitative methods literature. On the one hand, our framework makes it unnecessary to distinguish between approaches such as pattern matching (Campbell 1975), congruence analysis (George and Bennett 2005), process tracing (George and McKeown 1985), causal narrative (Sewell 1996), or the comparative sequential method (Falleti and Mahoney 2015). While these represent important initiatives to better understand the logic of qualitative research, a Bayesian perspective effectively subsumes them all, by revealing that inference always entails reasoning from evidence—whether within-case, cross-case, or a combination of both—to identify the best available explanation from a concrete set of alternatives. We emphasize that whereas Bayesianism in qualitative social science is strongly associated with within-case analysis, the same logic applies to comparative case studies that aim to assess theories with scope conditions that extend beyond a single case. On the other hand, our Bayesian framework elucidates fundamental problems with alternative approaches including QCA and fuzzy-set analysis (e.g., Ragin 1987 & 2000). When hypotheses depart from strict necessity or sufficiency, crisp set theory in effect introduces adhoc procedures to relax the rules of logic, instead of applying probabilistic reasoning. Fuzzy-set theory in contrast builds on vagueness as a central conceptual principle, whereas we argue that uncertainty—codified in Bayesian probability—is the appropriate framework for scientific inference.

1.3. Multi-Method Research

Our vision of Bayesian probability as a unified framework for scientific inference contrasts with the majority of literature on multi-method research, which at least implicitly operates within a frequentist framework (e.g., Lieberman 2005 & 2015, Gerring 2012, Weller and Barnes 2014, Seawright 2016, Goertz 2016). Authors maintain that combining different techniques—often statistical analysis and case studies—harnesses complementary sources of causal leverage. Statistical regression is generally conducted to establish a correlation, and case studies are included to illustrate causal mechanisms. However, many authors overlook the fact that the techniques they combine are associated with distinct epistemological principles and are designed to address different questions; the former aim to estimate population-level parameters, and the latter often aim to explain particular outcomes in specific cases (Mahoney 2010:141, Goertz and Mahoney 2012). It is difficult to see how a case study would strengthen an inference about the average causal effect in the eyes of a committed frequentist, or how the numerical value of a regression coefficient would bring much insight to understanding a particular case or set of cases. Seawright (2016:5) provides a cogent discussion of these problems, noting that "because qualitative and statistical approaches produce results that are different in kind, it is only possible to assess ... convergence very abstractly."

We contend that Bayesian probability is the only sound option for integrating qualitative and quantitative information on a more or less equal footing, because it is the only rigorous framework for which the same logic of inference applies across all types of data—whether stochastic or non-stochastic, experimental or observational, quantitative or qualitative. Efforts to bring both qualitative evidence and quantitative data to bear on a single research question without Bayesian probability (the approach pursued by Seawright 2016) inevitably prioritize either a quantitative method or a qualitative method as the main engine of inference.

Our perspective shares common ground with Humphreys and Jacobs (2015), who illustrate that multiple different goals, including estimating average effects, assessing case-level explanations, and comparing theories, can all be accommodated within a Bayesian model. However, we advocate moving away from the emphasis on average causal effects and distributions of unobservable causal types in a population—notions which remain rooted in frequentism—in favor of a simpler and more direct Bayesian approach in which causal hypotheses for explaining known outcomes of given cases are the primary propositions of interest. Here we agree with Mahoney (2015:202) that thinking about average causal effects is not very useful for many research agendas. We also caution that prospects for formally combining inferences from quantitative and qualitative research will be constrained by the inevitable difficulty of non-arbitrarily quantifying the inferential weight of evidence in social science contexts. Yet despite the challenges, we believe that both quantitative and qualitative components of research can benefit from applying Bayesian insights.

2. A GUIDE FOR READERS

This book aims to reach a broad readership. Qualitative research practitioners and scholars who include case studies within multi-method designs are of course central audiences. Our Bayesian framework and practical recommendations for inference and research design apply not just to single case studies, but also to small-N and medium-N comparative research, ranging from historical analysis to reconstruction of more contemporary policymaking processes, as well as studies that aim to combine qualitative information with quantitative data. Yet we also aim to foster greater understanding of Bayesian inference among scholars whose primary research involves large-N analysis, experimental designs, and/or formal theory. All research draws on insights from qualitative information—regardless of the core analytical approach employed—and we believe that Bayesian probability can serve as an important bridge between qualitative and quantitative methods.

With these diverse audiences in mind, we have sought to make the book accessible and engaging for readers with a wide range of technical backgrounds. For those who do not have any previous exposure to Bayesian inference or classical statistics, we stress that no mathematical skills are required beyond basic arithmetic and algebra. Indeed, no previous methodological training whatsoever is needed to follow our core arguments and recommendations for improving inference in qualitative research. At the same time, we include detailed mathematical foundations in order to fully substantiate our arguments and engage with readers who have more extensive technical training.[4] Both in the overview below and in the chapters that follow, we indicate which sections may be skipped by readers who are less interested in mathematical derivations and technical details. For readers who already have some familiarity with Bayesianism, we emphasize that

[4] On rare occasion, our technical treatments require familiarity with basic calculus (i.e., integration).

our approach diverges from most social science treatments of Bayesian inference—as such, we recommend proceeding through the book sequentially.

3. SCOPE OF THE BOOK

This section provides an overview of the coming chapters, introducing key Bayesian concepts and practical recommendations in plain language. We provide a fair amount of detail in order to convey a holistic sense of the book's scope and recommendations for qualitative research. Although some aspects of Bayesian reasoning require substantial practice to master, we hope that by the end of this introductory chapter, readers will be equipped with new insights and some easily-implemented practices for improving qualitative research.

Chapter 2, which together with this introduction forms Part I of the book, introduces the fundamentals of Bayesian probability. Part II, comprising Chapters 3–7, explains how to operationalize Bayesian analysis in qualitative research. Part III (Chapters 8–9) then places Bayesianism in methodological perspective. Part IV (Chapters 10–12) turns to Bayesian implications for research design.

3.1. Foundations

Because we live in a world of uncertainty, we intuitively reason in terms of probabilities. We think about how likely it is to rain in the afternoon, given the weather conditions we observe in the morning. If it does rain in the afternoon, we adjust our expectations about the likelihood of rain tomorrow. When we develop a cough, we decide if we should make a doctor appointment after assessing whether we are more likely to have contracted a common flu as opposed to a coronavirus, given our medical history, the severity of symptoms, and any other relevant knowledge such as news of a recent pandemic resurgence.[5] Similarly, in politics, we might contemplate the odds of democratic backsliding versus institutions constraining authoritarian tendencies in the United States. We update our beliefs about which outcome is more likely as we take new developments into account—Republican politicians' reluctance to acknowledge

[5] See Loredo (1990:90) for more examples of everyday probabilistic reasoning.

President Biden's victory in 2020, or the status of efforts to restrict voter rights.

While probabilistic thinking is commonplace in daily life, we do not always reason correctly about probabilities, as demonstrated by a large body of literature in cognitive psychology (e.g., Kahneman and Tversky 1974). Consider the following famous example from Kahneman and Tversky:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
> Which alternative is more probable?
> Linda is a bank teller.
> Linda is a bank teller and is active in the feminist movement.[6]

When presented with this question, the majority of respondents selected the second possibility as more likely. However, that answer violates the mathematics of probability theory—the probability of a conjunction (both a bank teller and a feminist activist) cannot be greater than the probability of either proposition alone (bank teller, or feminist activist).

As a second example, suppose there are two parrots in an enclosure at the zoo, one red and one blue, and the keeper tells us the blue parrot is female. What is the probability that both parrots are female? Now suppose instead that the keeper says at least one parrot is female (but is not sure whether both are female). When asked, most people think it is equally likely for both parrots to be female in these two scenarios.[7] However, this reasoning is incorrect, because (as we will explain below) it does not take into account that we must evaluate a *conditional probability* given the information we have been told, which differs across the two scenarios.

Beyond these two examples where untrained intuition tends to lead us astray, numerous cognitive biases can undermine our reasoning. One salient example is confirmation bias, where we interpret information in a way that is partial to our preexisting beliefs or desires (e.g., attachment to a particular hypothesis that we hope to be or want to prove true). Another common

---

[6] See Kahneman (2011:156).

[7] Out of roughly 45 students at the Syracuse Institute for Qualitative and Multi-Method Research in 2019, 71% chose one half as their response for the first scenario, and 51% chose one half as their response for the second scenario. Only 18% selected the correct response in the second scenario, where we know that at least one of the parrots is female.

example is anchoring bias, where the information we first encounter wields a disproportionate influence on our reasoning, such that we do not reevaluate our views as much as we should in light of additional evidence. On other occasions, availability bias might lead us to overweight whatever information we have learned most recently, rather than carefully considering all of the evidence we have collected. Familiarity with Bayesian probability can counteract these cognitive biases, thereby improving our reasoning and helping us make better judgments in a wide range of both real-world contexts and research endeavors.[8]

Logical Bayesianism, the school of Bayesianism that we advocate as the foundation for inference, seeks to represent the rational degree of belief we should hold in propositions given the information we possess, independently of subjective opinions, predilections, or aspirations. In ordinary Boolean logic, the truth-values of all propositions are known with certainty. Bayesian probability is an "extension of logic" in that it provides a prescription for how to reason when we have incomplete knowledge and are thus uncertain about the truth of propositions. When degrees of belief assume limiting values of zero (impossibility) or one (certainty), Bayesian probability automatically reduces to ordinary logic.

A central tenet of logical Bayesianism is that probabilities should encode knowledge in a unique, consistent manner. Incorporating information in different but logically equivalent ways (e.g., learning the same pieces of information in different orders) must produce identical probabilities, and rational individuals possessing the same information must assign the same probabilities. Physicists including Cox (1961) and Jaynes (2003) have proved mathematically that these consistency requirements lead directly to all of the standard rules of probability theory.

The consistency requirements of logical Bayesianism are more demanding than requirements imposed in approaches that draw on the "psychological" or "subjective" school of Bayesianism, which is common in philosophy of science literature and many Bayesian statistics textbooks for social science (e.g., Jackman 2009). In this latter, personalistic approach, rationality requires probabilities to be *coherent*, which means that utility-maximizing agents must decline "Dutch Book" bets (e.g., de Finetti 1937, Ramsey 1926, Savage 1954), where loss is certain. Coherence in turn implies the basic rules for manipulating probabilities. But as long as probabilities sat-

[8] There is of course a large literature that examines to what extent humans are innately Bayesian in their reasoning (e.g., Griffiths et al. 2011, Marcus and Davis 2013, Ullman and Tenenbaum 2020). Our concern in this book is distinct—we seek to teach social scientists how to be *better* Bayesians.

isfy these rules, they can be based on pure opinion or whim—whatever happens to motivate an individual to hold some particular subjective degree of belief. Accordingly, within subjective Bayesianism, individuals who possess the same information need not assign identical probabilities. In our view, this feature of subjective Baysesianism makes it untenable as a foundation for inference and knowledge accumulation.[9]

We develop the fundamentals of logical Bayesianism and probability as extended logic in Chapter 2, which reviews Boolean logic, elaborates in more detail the fundamental consistency requirements, or "desiderata" for reasoning under uncertainty, and explicates the basic rules of probability that emerge from them—including of course Bayes' rule. Appendix B of Chapter 2 presents the Cox-Jaynes derivation of the sum rule and product rule of probability from the consistency requirements described above; we recommend this appendix to readers who are interested in gaining a greater appreciation of the deep foundational justification for logical Bayesianism as a framework for scientific inquiry.

The second half of Chapter 2 provides worked examples to show how carefully applying the laws of Bayesian probability can help us correctly solve problems where untrained intuition might lead us astray. A central theme running throughout these problems is that inferences depend critically on the details of the knowledge we possess, and seemingly small changes in that knowledge can lead to very different conclusions. To briefly illustrate this central point, we revisit the two-parrot puzzler mentioned above, which is easy to understand without mathematical equations.

In the first version of the puzzler, we know that the blue parrot is female, but we have no information about the sex of the red parrot. This state of knowledge limits the color-sex combinations to only two possibilities: either (1) the blue parrot is female and the red parrot is male, or (2) the blue parrot is female and the red parrot is also female.[10] In the absence of any further information, we should assume that these two possibilities are equally likely, and

[9] In discussions about the foundations of probability, "subjective" and objective" are used in many distinct senses, which invites confusion. Unfortunately, these adjectives seem too entrenched to avoid. Some authors use "subjective probabilities" to refer to any Bayesian probabilities, in contrast to probabilities based on relative frequencies or propensities, which are supposed to be "objective" physical properties of some stochastic process (see Chapter 7). As explained above, we reserve the term "subjective Bayesianism" for a particular variant of Bayesianism that differs from logical Bayesianism; the latter approach is "objective" in the sense that two observers with identical information must rationally assign the same probabilities. Following Jaynes (2003), we use the term "objective Bayesianism" as a synonym for logical Bayesianism.

[10] Here we set aside the rare possibility of a gynandromorph—a half-male, half-female bird. www.nytimes.com/2019/02/09/science/cardinal-sex-gender.html

the probability that both are female is accordingly one half.[11] In the second version of the problem, we are told that at least one of the parrots is female, but we do not know which of the two is definitely female. This distinct state of knowledge leaves open three possibilities: (1) the blue parrot is female and the red one is male; (2) the blue parrot is male and the red one is female; or (3) the blue parrot is female and the red one is also female. The probability that both parrots are female is then one third—not one half! Our different states of knowledge in the two versions of the problem make all the difference.

As a note to readers, if this overview of Chapter 2 satisfies your interest in the foundations of probability as extended logic, you are comfortable with the idea that probabilities depend on the information you know, and your goal is to learn the basics of how to apply Bayesian reasoning in qualitative research as quickly as possible while being subjugated to as little math as necessary, you may skip Chapter 2 and proceed directly to Chapter 3. However, many efforts to apply and/or critique the use of Bayesian reasoning in qualitative research espouse misconceptions that stem from an inadequate understanding of the basic foundations of probability theory, so we would encourage readers to consider spending at least a little time with Chapter 2.

3.2. Operationalizing Bayesian Reasoning in Qualitative Research

Bayesian probability provides a mathematical framework for how we should revise our confidence in hypotheses, given our prior knowledge and the evidence we discover during our research. Bayesian inference involves comparing hypotheses and figuring out which one provides the best explanation for the phenomenon under investigation. The end goal is to assess how much confidence we can justifiably hold in our conclusions based on the full body of information we bring to bear on the problem.

Chapter 3 illustrates how to apply Bayesian probability to qualitative research in a "heuristic" manner, where we aim to follow Bayesian principles in our reasoning process without applying

[11] In many bird species, males are more colorful than females. However, sexual dimorphism is not very apparent in most parrot species, so we have no reason to believe that sex and plumage color are linked. If readers (erroneously) applied background assumptions regarding sexual dimorphism in avians in general, they might arrive at different probabilities for our two-parrot problem. There are however some remarkable exceptions. For example, in the Australian parrot species *Eclectus roratus*, males are predominately intense green, with yellow beaks and some red and blue markings, while females are a striking combination of red, maroon, and blue with black beaks.

the full mathematical apparatus of Bayesian analysis. We overview below three core aspects of heuristic Bayesian reasoning.

The first task is to devise a set of hypotheses to compare. These hypotheses must be *mutually exclusive*, which we take as synonymous with "rival," meaning that only one can be true—if a given hypothesis is a correct characterization of how the world works, the other hypotheses must necessarily be incorrect. Much confusion surrounds this point in the literature; it tends to raise concerns that Bayesianism precludes consideration of multiple contributing causal factors or complex causal hypotheses. This is not the case. We can construct rival hypotheses that incorporate as many causal factors as we deem salient, operating through any configurations and interactions that we judge plausible. The only requirement is to ensure that alternative hypotheses are clearly posed as, and/or explicitly asserted to be, mutually exclusive propositions. This task is usually straightforward; it is always possible to take a set of causal *variables* that need not operate exclusively and create from them a set of causal *hypotheses* that are mutually exclusive in logical terms.

As a simple example, suppose we come home and discover that the cookie jar, which was previously full, is now empty. We want to figure out who is responsible, and the only likely suspects are the 5-year old and the dog. We could of course consider single-culprit hypotheses: H_1 = *The child alone consumed the cookies*, or H_2 =*The dog alone consumed the cookies.* Here we have carefully worded these hypotheses to clearly communicate that only one or the other can be true. If we consider it plausible that both the child and the dog contributed to the outcome, we can pose hypotheses such as: H_3 = *The dog and the child colluded in acquiring and consuming cookies,* or H_4 = *The dog and the child both ate cookies but did so independently*, where we explicitly assert that H_3 and H_4 are to be understood as mutually exclusive alternatives. If we think these hypotheses do not adequately encompass the range of feasible explanations, we could pose additional, more complex alternatives, such as: H_5 = *The dog and child began by eating cookies independently, and later cooperated to extract the remaining cookies from the jar.* Each of these five hypotheses tells a distinct causal story to explain the empty jar, such that no two of them can simultaneously be true.

The next step in Bayesian reasoning is to assess *prior odds* that express how much initial confidence we hold in a given hypothesis compared to rivals, based on our background knowledge. Logical Bayesianism aims to assign probabilities that are objectively grounded in the information we posses—which should reflect an assessment of the cumulative state of knowledge in our field that serves as the starting point for our research. In practice however, given the complexity of social science, it is impossible to systematically incorporate every component of our background knowledge into our prior beliefs. Accordingly, we should provide readers with a thoughtful but reasonably concise justification for why we prefer one hypothesis over another, highlighting the elements of our background information that matter most for our judgments. A second option is to start with "indifference" priors, which usually means even prior odds (i.e., all hypotheses are equally likely). This approach is reasonable because different scholars inevitably possess different background information and may thus hold very different initial views about the plausibility of hypotheses. Placing equal prior probabilities on the hypotheses shifts attention to the inferential import of the evidence, with the motivation that even if we disagree on priors, we can strive to concur regarding which hypothesis becomes *more* plausible in light of the evidence presented—and if the evidence is strong enough, scholars with different priors may end up with posteriors that clearly favor the same hypothesis. Regardless of which approach we adopt, we can also conduct sensitivity analysis to assess how much our conclusions would change if we start with different assignments for the prior odds.

Here we should address a common argument from skeptics that prior probabilities are a weakness of Bayesianism, on the basis that they are arbitrary or subjective. Our response is fourfold. First, all research involves assumptions, but these assumptions often remain hidden or unstated. Logical Bayesianism encourages us to make our assumptions explicit, which includes openly identifying the elements of our own background knowledge that motive our prior views.[12] Second, if the evidence turns out to be strong, priors will not make much of a difference to our conclusions; many of the worked examples in this book will illustrate this point. And even when prior odds do matter significantly for the posterior odds, simply building some consensus on which hypothesis is favored by the evidence in a given work of scholarship can serve as a significant step toward knowledge accumulation. Third, rational reasoning requires specifying prior odds—there is no way to sidestep this reality. Just as Newton's laws require an input of initial conditions in order predict the trajectory of a falling apple, we cannot arrive at posterior odds without prior odds. And finally, logical Bayesianism does dictate objective criteria for

[12] Berger and Berry's influential 1998 article on "Statistical Analysis and the Illusion of Objectivity" in *American Scientist* develops these points in favor of Bayesian statistics vs. the frequentist alternative.

deriving prior probabilities for many important inferential problems,[13] although these types of problems tend to arise more commonly in quantitative analysis.

Evaluating the inferential strength of the evidence is the third and most important aspect of Bayesian reasoning. Using informal language, this task requires assessing how well the evidence fits with a given hypothesis compared to rivals. Bayes' rule tells us precisely what it means for the evidence to "fit better" with one hypothesis versus a rival—that hypothesis must make the evidence more expected, or equivalently, less surprising. In Bayesian terminology, we evaluate *likelihood ratios*. We must imagine a hypothetical world where a given hypothesis is true and think about what kinds of things (e.g., clues, events, outcomes, and scenarios) might transpire in that world. We ask whether the evidence at hand would seem plausible and natural in that world, or if it would have to be regarded as an unusual fluke or part of an unlikely chain of occurrences. We then "mentally inhabit the world" of a rival hypothesis (Hunter 1984), applying the same thought process described above, and we ask which of these hypotheses makes the evidence seem more expected. Assessing likelihood ratios is the critical step that tells us how to update our prior views about the relative plausibility of our hypotheses. According to Bayes' rule, we gain more confidence in a particular explanation to the extent that it makes the evidence we obtained more plausible than a specified rival.

Figure 2 provides an overview of heuristic Bayesian reasoning. If there are more than two explanations that we consider plausible from the outset, we conduct pairwise comparisons following the steps described in Figure 2. If at any point in the process we decide to revise our hypotheses or devise a new one, we must simply return to step one, articulate the new or revised hypotheses that we now wish to compare, and then proceed though each subsequent step in turn.

While Bayesian reasoning is intuitive in many regards, assessing likelihood ratios—the key inferential step—requires using intuition in a somewhat different way from how one might be accustomed to thinking about evidence. Becoming skilled at "mentally inhabiting the world" of each hypothesis and assessing which one makes the evidence more plausible takes practice. To that end, Chapter 3 and subsequent chapters provide multiple examples that apply this thought process to real-world case study research. We introduce various recommendations for

[13] See Chapter 2, Appendix C on the principle of indifference.

1. Consider a pair of rival hypotheses
2. Assess your prior odds on the hypotheses
   - How plausible is one hypothesis compared to the rival in light of your background knowledge?
3. Evaluate likelihood ratios for the evidence
   - Mentally inhabit the world of each hypothesis in turn
   - Ask which hypothesis makes the evidence that you obtained more expected
4. Update your prior odds to obtain the posterior odds
   - According to Bayes' rule, we gain confidence in the hypothesis that makes the evidence more expected

FIG. 2 **Basic Steps in Heuristic Bayesian Reasoning.**

evaluating likelihood ratios; for example, we explain how to handle "testimonial evidence" from potentially fallible or biased human sources, and how to reason about additional evidence in light of previously-analyzed evidence.

Heuristic Bayesian reasoning can do much for qualitative research, and by the end of Chapter 3, readers will have a set of guidelines that can substantially improve inference, even if they choose to proceed no further through the book. However, we encourage even math-averse readers to continue to the next chapter, which presents an approach that can be more powerful than heuristic Bayesian reasoning.

Chapter 4 introduces "explicit" Bayesian analysis, which entails quantifying prior odds and likelihood ratios, and then deriving posterior odds by directly applying Bayes' rule. This approach allows us to communicate degrees of belief more precisely than is possible with ordinary language, which in turn helps pinpoint sources of disagreement between scholars more effectively. Explicit Bayesian analysis also allows us to draw inferences from multiple pieces of evidence more systematically, which is especially valuable when all of the evidence does not weigh in favor of the same hypothesis. At the same time, we must keep in mind that it is impossible to unambiguously quantify probabilities when working with complex, nuanced, qualitative information. Some amount of subjectivity will necessarily intrude into our probability assessments.

Our central recommendation for explicit Bayesian analysis is to work with a logarithmic scale, which greatly simplifies computations and yields an easily remembered, linear form of Bayes' rule: the *posterior log-odds* in favor of one hypothesis relative to a rival equals the sum of the *prior log-odds* and the *weight of evidence*, an intuitive concept promoted by Jack Good and Alan Turing that is simply proportional to the logarithm of the likelihood ratio. This approach affords a second advantage: it helps leverage intuition, because our sensory perception (e.g., sound, sight, touch) works on a logarithmic scale, not a linear scale. We can accordingly quantify prior log-odds and weights of evidence in decibels (dB), using an analogy to sound levels. When quantifying the weight of evidence, we ask how loudly the clue at hand is speaking in favor of H_j vs. H_k. Do the facts whisper, or shout in favor of one hypothesis over another? Likewise, when quantifying prior log-odds, we ask how loudly one hypothesis speaks to us compared to the alternatives. For readers who are less familiar with logarithms, Chapter 4 reviews the basic features of this function, but it is important to stress that using decibels in practice requires no working knowledge about the mathematics of logarithms.

The log-odds form of Bayes' rule allows us to conduct inference by imagining a "Bayesian balance" where each side corresponds to a different hypothesis. We place our prior-log odds, measured in decibels, on the side of the scale that corresponds to the hypothesis we initially prefer. We then place the weight of evidence for each clue, also measured in decibels, on the side of the balance corresponding to the hypothesis that the clue in question favors. The deflection between the two sides of the scale represents our posterior log-odds, which express in decibels how much confidence we hold in one hypothesis relative to the rival in light of the evidence and our background knowledge. For example, in Figure 3, the prior log-odds tip the scale 7 dB in favor of H_1, representing a fairly weak preference for H_1 over H_2 in light of our background knowledge. One piece of evidence also weakly supports H_1 over H_2 (5 dB), but a second piece of evidence strongly favors H_2 over H_1 (25 dB). Overall, the balance tilts toward H_2, yielding posterior log-odds of 13 dB in favor of H_2 vs. H_1.

The balance analogy and the weight of evidence concept help illustrate that the amount of evidence or number of clues is not what matters for Bayesian inference. As the renowned 17th century logician and mathematician Gottfried Wilhelm Leibniz remarked, evidence is not to be counted but weighed. Updating depends only on the overall, net weight of evidence aggregated across all of the clues we discover. In some instances, a large quantity of evidence may not allow us to make strong conclusions about which hypothesis is correct; each clue might provide so little inferential weight that the net total does not accumulates to a significant value, or different clues might weigh in favor of different hypotheses such that net weight of evidence

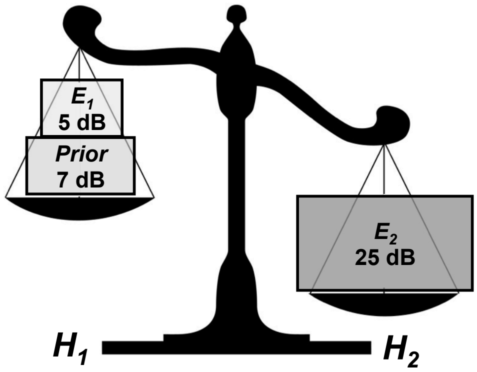

FIG. 3 **Bayesian Balance:** Inference with prior log-odds and weights of evidence in decibels (dB).

is small. Yet in other situations, we might find just one or two highly informative pieces of evidence that substantially boost our confidence in a given explanation relative to rivals.

Chapter 5 provides more examples of how to apply explicit Bayesian analysis, with a focus on studies that include multiple cases. Analysis with more than one case proceeds in the same manner as analysis with multiple pieces of evidence from a single case, by aggregating the inferential weight of all of the evidence from all of the cases studied. We further address the question of how to generalize a hypothesis, by broadening its scope and examining evidence from cases beyond the narrower original scope of the study.

Chapter 6 revisits hypotheses and priors; we present some additional technical details and more advanced topics that we draw on in subsequent chapters. For readers who harbor further questions about mutually exclusive hypotheses that were not resolved in Chapter 3, we begin with an extended discussion of this issue that includes precise mathematical guidelines for elaborating exclusive hypotheses from an initial set of non-exclusive causal propositions. We then introduce the concept of Occam factors, a feature of logical Bayesianism that allows us to concretely assess tradeoffs between parsimony and explanatory power.

In essence, Occam factors are a mathematical representation of Occam's razor that is built into Bayes' rule. To give a brief preview, a complex hypothesis incurs an Occam penalty relative to simpler rivals via its prior odds. If the more complex hypothesis is actually the better explanation, its posterior odds should win out thanks to the improved inferential leverage it provides compared to the simpler alternatives. More precisely, the likelihood ratio for the evidence will overwhelm the initial Occam penalty in the prior odds. Bayesianism thus penalizes (overly) complex hypotheses *if* they do not provide enough additional explanatory leverage compared to simpler alternatives—in accord with Einstein's dictum that things should be a simple as possible, but no simpler.

Chapter 7 returns to an important use of explicit Bayesian analysis: scrutinizing inference in qualitative research. We break down this process into three steps: assessing whether the logic of inference follows Bayesian principles, scrutinizing the weight of evidence, and conducting sensitivity analysis on priors and/or weights of evidence. We examine several published narrative-form case studies, including examples that do a good job of intuitively approximating Bayesian reasoning, as well as examples that deviate substantially from Bayesian principles.

We identify two common practices in qualitative research that depart from Bayesianism: providing evidence that "traces the mechanism" of the working hypothesis but neglecting to ask how consistent that evidence would be with rival hypotheses, and deploying one set of observations to substantiate the working hypothesis while using a different set of observations to rule out rival hypotheses. The problem with the first practice is that Bayesian updating depends on likelihood *ratios*. Evidence that appears consistent with a given hypothesis does not necessarily support it, because the evidence might be equally or even more plausible in the world of a rival hypothesis. Likewise, evidence that seems inconsistent with a given explanation does not necessarily undermine it, because the evidence could be even less plausible under a rival. The problem with the second practice is that all hypotheses must be compared in light of all salient evidence—we should not use different subsets of the evidence to evaluate different hypotheses. Stated differently, inference does not proceed by sorting evidence into observations that support the working hypothesis and observations that undermine the rivals. In published research, it may be reasonable to dismiss a rival in light of a few highly decisive pieces of evidence without having to explicitly analyze all of the other information that we bring to bear elsewhere in our analysis, but we should at least use our intuition to assess whether the rest of the evidence might counteract the inferential weight of these few key observations that on their own cast substantial doubt on the rival.

With an eye toward avoiding these pitfalls, we recommend the following organizational approach for writing Bayesian-informed case narratives. Begin by stating all of the hypotheses under consideration. Then present all of the salient evidence in an initial narrative that inhabits the world of the hypothesis deemed most plausible (presumably the author's argument). After the main narrative, use either heuristic or explicit Bayesian analysis to assess how strongly the evidence favors the leading hypothesis over the rivals. If we decide to leave a few anomalous pieces of evidence out of the narrative, we can introduce them and assess to what extent they weaken the overall inference in this subsequent hypothesis-comparison section.

This organization has a number of advantages. First, arguing against rival hypotheses only after presenting the core case evidence helps keep us from casting an argument aside too hastily; if an author fails to consider countervailing evidence from the case narrative when comparing hypotheses, readers may be more likely to catch the problem than if the author summarily dismisses rivals from the outset. Second, the Bayesian hypothesis comparison forces us to consider how well the evidence fits with rivals, rather than focussing only on whether the evidence is consistent with the working hypothesis. Finally, this organizational structure preserves readability, by providing a case narrative before turning to more technical inferential analysis. Case narratives are a highly effect way to communicate information and analysis, and we stress that they should retain a prominent place in qualitative research even while we introduce Bayesian techniques for improving inference.

We emphasize that the goal from a Bayesian perspective should be to honestly assess the uncertainty associated with inferences. We rarely end up with all of the evidence we would like to be able to establish our argument as the best explanation beyond reasonable doubt. But as Bennett and Checkel (2015:30) observe, "conclusive process tracing is good, but not all good process tracing is conclusive." Accordingly, we should explicitly address those pieces of evidence that run most counter to the overall inference. If the evidence is not decisive, we should be forthright about the limitations and suggest avenues for future research that might help resolve the doubts.

3.3. Bayesianism in Methodological Perspective

Chapters 8 and 9 take a step back to place our Bayesian approach in comparative perspective. We contrast logical Bayesianism with other inferential approaches commonly used in the social sciences and develop our argument that Bayesianism provides a universal framework for inquiry that transcends commonly emphasized boundaries between different research traditions. Whereas Part II of the book emphasizes the practicalities of applying Bayesian reasoning in qualitative research, these two chapters that form Part III engage in some epistemological discussions that aim to shift the terms of debate regarding the nature of qualitative research and its potential contribution to social inquiry. In addition to engaging with methodologists and multi-method scholars, we hope these chapters will help arm qualitative research practitioners with some resources to push back against currents of thought within their fields that might diminish the value of their work.

Chapter 8 compares Bayesianism to frequentism, the most widely used framework in quantitative social science that forms the basis for classical statistics, and by analogy or emulation has influenced much of the qualitative methods literature as well. Bayesian and frequentist approaches to inference differ first and foremost in how they conceptualize probability. Whereas Bayesianism treats probabilities as rational degrees of belief in logical propositions given partial or imperfect information, frequentism defines probability as a limiting proportion, or relative frequency, in an infinite series of random trials or repeated experiments. These contrasting definitions of probability give rise to very different approaches to inference. While Bayesianism directly applies the rules of probability to evaluate the plausibility of rival explanations in light of all available information, frequentism disallows probability statements about hypotheses, and inference proceeds much more indirectly. Generally speaking, frequentist hypothesis testing rejects or fails to reject a single null hypothesis by asking how likely we would be to observe data as or more extreme than what we actually obtained if the null hypothesis is correct, under infinite imagined repetitions of the experiment.

The Bayesian approach affords many advantages. Perhaps most importantly for qualitative research, Bayesian inference does not require a stochastic data-generation model or reference to a population or ensemble of possible observations. Bayesianism is therefore well-suited for handling qualitative and historical evidence and for drawing inferences in single case studies, whereas strictly speaking, frequentism is appropriate only for data that can be naturally treated as a random sample. For these reasons, any scholar who is willing to consider that qualitative evidence may be valuable for inference has already moved beyond the bounds of frequentism toward a Bayesian perspective. More broadly, Bayesianism provides inferences that are generally easier to interpret and more relevant to the immediate question at hand; namely, what can we learn about the truth of our hypotheses from the actual data we observe. In contrast, frequentism provides results that can only be interpreted with respect to the infinite long-run and invokes hypothetical data that might have been but were *not* obtained. This feature of frequentist hypothesis testing can lead to hidden subjectivity, where unobservable thoughts in the experimenter's mind can affect identification of random variables and probability assignments (Loredo 1990:93). Reliance on hypothetical data also necessitates rigidly pre-specified research designs, whereas Bayesian inference, which uses only the actual data in hand, allows much greater flexibility to amend the research design as data collection and analysis proceed.

It is worth highlighting here a few common frequentist-oriented tendencies that run counter to Bayesian principles and can lead to misunderstandings when seeking to apply Bayesian inference in qualitative research. These include (i) taking probabilities to always be proportions or base rates in a population of cases, as per the frequentist definition of probability, rather than degrees of belief given the information we possess; (ii) aiming to assess the truth of a single hypothesis in light of the evidence, as per frequentist null hypothesis testing, without directly comparing how well the evidence fits with rival hypotheses; (iii) dichotomizing evidentiary support into confirming vs. disconfirming categories, rather than asking to what degree the evidence supports a particular hypothesis over rivals, and (iv) associating each potentially contributing independent variable with a separate hypothesis, rather than articulating a comprehensive hypothesis that explains how all variables deemed salient bring about the outcome. The final tendency may arise from a frequentist regression perspective that focusses on the causal contribution of each independent variable in turn. We will expound on all of these issues throughout the book.

Chapter 9 explicates our view that Bayesian probability provides a single, unified framework of inference that makes it unnecessary to introduce various types of multi-method research, to triangulate between distinct methodologies, or to distinguish between different sources of analytic leverage associated with correlational data versus within-case evidence. In fact, from a Bayesian perspective, many commonly-drawn distinctions in social science methodology are overstated.

We develop this perspective with an emphasis on quantitative and cross-case vs. qualitative and within-case research, as well as experimental vs. observational data. We argue that within a Bayesian framework, while different types of evidence may be better suited to particular kinds of questions, any and all relevant evidence, whether quantitative or qualitative, cross-case or within-case, experimental or observational, can and should inform our probabilities. In principle, all evidence contributes to inference in the same manner via Bayes' rule, although the details of the analysis may look different depending on the data in hand. With regard to experimental research, we argue that random assignment is no silver bullet; all research, whether experimental or observational, is vulnerable to systematic error and bias and cannot be completely immunized from the possibility of confounding variables.

A logical Bayesian perspective both simplifies our understanding of inference in small-N qualitative research and places scholarship within that tradition on more equal footing with respect to other kinds of research. Our remapping of the methodological terrain identifies a single fundamental underlying dimension that cuts across conventional distinctions between different kinds of data—the extent to which likelihood ratios can be objectively and unambiguously specified. This dimension tends to separate hard sciences like physics from the social sciences, simply because the latter study far more complex and inherently noisier systems.

In accord with this perspective, we recommend applying Bayesian analysis across all components of research, as an alternative to more conventional mix-and-match multi-method research. In principle the posterior odds from a Bayesian regression analysis of quantitative data would simply become the prior odds for a Bayesian analysis of qualitative evidence, or vice versa, and the final posterior odds then reflect all empirical information leveraged in the project. But if quantitative Bayesian analysis proves too complicated or lies beyond the skill set that scholars posses or wish to develop, we nevertheless encourage multi-method practitioners to apply Bayesian reasoning in their case studies in order to make better inferences from qualitative evidence.

3.4. Bayesian Implications for Research Design

The final part of the book turns to research design. Chapter 10 explores the Bayesian foundations of iterative research, where prior knowledge informs hypotheses and data gathering strategies, evidence inspires new or refined hypotheses along the way, and there is continual feedback between theory and data. Qualitative research almost always follows this kind of trajectory, yet prevailing norms hold instead that research should progress linearly, from (1) articulating hypotheses, to (2) testing those pre-specified hypotheses on new data—namely, data not known to the investigator when the hypotheses were devised—during a second, separate stage of research.

We argue that the mismatch between standard methodological prescriptions and actual practice in qualitative research stems from different logics of inference. The linear, deductive model arises from the dominant frequentist perspective, where confirmation bias and ad-hoc hypothesizing (i.e., constructing "just so stories," or over-fitting to the details of the evidence in hand) can be serious problems unless a firewall is maintained between theory building and theory testing. In contrast, iterative research is grounded in a Bayesian logic that fits naturally with how we intuitively move back and forth between theory and data. Within logical Bayesianism, "new" evidence that is learned after devising hypotheses has no special status relative to "old" evidence that was known while formulating hypotheses when it comes to adjudicating between rival explanations, and dichotomies between inductive vs. deductive and theory-testing vs. theory-building stages of research are artificial and unnecessary.

Our perspective has important implications for debates on research transparency. In contrast to suggestions that qualitative scholarship should keep track of and report what evidence was known before inventing a hypothesis and what evidence was discovered subsequently, we contend that such information about relative timing is irrelevant for conducting and scrutinizing inference. Instead, conscientiously applying Bayesian reasoning in itself helps safeguard against confirmation bias and ad-hoc hypothesizing. Bayesianism precludes a common form of confirmation bias—subconsciously focussing only on a favored working hypothesis—by forcing us to consider rivals when assessing the weight of evidence. Bayesian reasoning restrains ad-hoc hypothesizing via Occam factors (Chapter 6) that penalize overly complex hypotheses if they do not adequately outperform simpler alternatives. Our practical Bayesian-inspired advice for guarding against ad-hoc hypotheses includes (a) treating inductively-inspired hypotheses with healthy skepticism, (b) starting with reasonably simple theories and adding complexity incrementally as justified by the data, (c) scrutinizing whether all of the causal factors in a hypothesis actually improve explanatory leverage compared to simpler rivals, and (d) asking if the hypothesis might apply more broadly. We further stress that a *post-hoc* hypothesis (meaning devised after the evidence) is not necessarily an *ad-hoc* hypothesis (meaning that overly complex or arbitrary features have been added to fit the data).

Chapter 11 turns to Bayesian interpretations of test strength, which will serve an important role in our discussion of case selection in the following chapter. We critique recent initiatives that draw on the concepts of sensitivity and specificity to place Van Evera's (1997) four process-tracing test types within a fully probabilistic Bayesian framework. These earlier approaches map test type onto a probative value space defined by the likelihoods of finding a clue under one or the other of two rival hypotheses (Humphreys and Jacobs 2015). We present three preferable approaches that instead classify test strength in terms of (1) *weights of evidence* (the logarithm of the likelihood ratio, Chapter 4), (2) *relative entropies*, and (3) *expected information gain.* These options more effectively highlight the connection between test strength and *relative* likelihoods, while also employing logarithmic scales to better measure how much we learn from a given test, in accord with standard practice in Bayesian analysis and information theory. Weights of evidence characterize post-data learning, while relative entropies and expected information gain may be of interest when making decisions about what data to gather. Relative entropies average the weight of evidence over the possible clue outcomes we might encounter, weighted by their respective likelihoods. Expected information gain in addition averages the relative entropies over our prior uncertainty regarding which hypothesis is correct.

Finally, Chapter 12 elaborates our Bayesian perspective on case selection. The literature on this topic has produced a plethora of case selection strategies but little consensus on which strategies we should follow in what contexts, and why. We present a much simpler view of case selection that draws on information theory. The core principle is to identify cases that will be highly informative for inventing, testing, and/or revising hypotheses.

At early stages of research, any information-rich case will be useful for devising hypotheses. Once we have articulated hypotheses to compare, the goal becomes selecting cases that are anticipated to serve as strong tests—namely, cases that can be expected to provide a large weight of evidence in favor of the best hypothesis. The idea is to choose cases that will help hone in on the best explanation as efficiently as possible. In formal Bayesian terms, we aim to maximize a generalized expression for expected information gain, a concept introduced in the previous chapter for the specialized context of a binary clue and just two rival hypotheses.

In practice, calculating expected information gain would be prohibitively difficult. However, the mathematical properties of expected information gain tell us that from the outset, we can justifiably expect to learn from virtually *any* relevant case we choose. So despite the fact that practical realities will generally prevent us from identifying optimal cases ahead of time, we can still expect that on average, the (possibly sub-optimal) cases that we do study will bring us closer to the best explanation. Accordingly, from a Bayesian perspective we ought to worry less about case selection than most of the literature would suggest. As such, we advocate prioritizing practical considerations as needed, and allocating a greater share of limited time and resources to data collection rather than efforts to identify truly optimal cases ahead of time.

Our broader case selection recommendations include retiring the notion of most-likely and least-likely cases; we conclude after careful scrutiny that these terms have been used in too many different ways, and that the various intuitions authors have in mind by and large do not map onto Bayesian principles. Instead, expected information gain provides a cogent prospective measure of test strength; cases become more "crucial" as expected information gain increases. We further stress that there is no need to begin the selection process by enumerating every possible case. Bayesianism need not entail population-based inference; our goal is not to choose a representative sample for estimating average causal effects, but rather to update the relative odds on explanatory hypotheses that are meant to apply in clearly-specified contexts. Scholars should nevertheless provide a cogent rationale for case selection to ensure that hard tests have not been deliberately avoided, to allow other scholars to scrutinize claims about the argument's scope, and to facilitate follow-up research.

4. A SPOTLIGHT ON LOW-TECH BEST PRACTICES

We conclude this introduction to Bayesian reasoning by overviewing a few of our central, less technically-demanding guidelines for improving inference in qualitative research. Some of these recommendations may lead scholars to think differently about evidence and inference. Others correspond closely to common sense, yet our Bayesian perspective gives these practices solid methodological foundations that are not widely appreciated.

1. Work with mutually exclusive hypotheses, and articulate them as clearly as possible so that each corresponding world is defined well enough for us to reason about what kinds of evidence would be more expected, or less expected, if that hypothesis is correct. Bear in mind that it is always possible to construct mutually exclusive hypotheses that may share some causal factors in common, and constructing mutually exclusive rivals does not limit the extent or nature of the causal complexity we choose to theorize.

2. Remember that inferential weight derives from the relative fit of the evidence with rival hypotheses. We cannot assess how strongly the evidence supports a hypothesis without specifying a rival. Accordingly, we should not ask whether the evidence supports a

hypothesis, but rather whether the evidence supports that hypothesis more than it corroborates rivals. Within a Bayesian framework, inference does not come from producing evidence that is consistent with a hypothesis, but instead from showing that the evidence is *less likely* under a rival.

3. Be sure to "mentally inhabit the world" of *each* hypothesis in turn. Even if the evidence does not seem to have anything to do with one of the hypotheses under consideration, we must still ask how surprising or expected the evidence would be in that world.

4. "Tracing" causal processes is often useful. But it is neither sufficient nor necessary for inference. We must always ask whether the evidence is more expected, or less expected in the world of a rival hypothesis. And *any* evidence that is more expected in the world of one hypothesis relative to a rival allows us to update our view about which of those hypotheses provides the better explanation—whether that evidence speaks directly to one of the theorized causal mechanisms or not, and however incomplete that evidence is.

5. Describe the evidence in enough detail for other scholars to independently assess its inferential weight. Remember that small details can sometimes make a big difference for the likelihood of the evidence in the world of a particular hypothesis.

6. The biases and instrumental incentives we attribute to a source (e.g., an informant, or the author of a primary document) may vary depending on the hypothesis under consideration. For that reason, the evidence to be evaluated should include information about the source and the context in which that source made the claim or statement of interest.

7. When writing case narratives, remember that in principle hypotheses must be compared in light of the total body of evidence. We may justifiably focus on the most decisive pieces of evidence, but we should check whether weaker pieces of evidence might collectively produce a consequential counterbalancing inferential weight.

8. Instead of trying to prove that a hypothesis is correct, assess how much uncertainty surrounds the inference: in light of the evidence and our background information, how confident are we that the leading hypothesis provides the best explanation?

We end with a simple message: it is perfectly possible to conduct systematic causal inference when working with observational, non-stochastic, qualitative data. Logical Bayesianism provides an intuitive and rigorous framework to that end. Although qualitative social science can never hope to be free from all subjectivity, Bayesianism can help narrow disagreements and bring us closer to the truth.