PA

Understanding Bayesianism: Fundamentals for Process Tracers

Andrew Bennett[1], Andrew E. Charman[2] and Tasha Fairfield[3]

[1] Georgetown University, Washington, DC, USA. Email: bennetta@georgetown.edu
[2] University of California–Berkeley, Berkeley, CA, USA. Email: acharman@physics.berkeley.edu
[3] London School of Economics, London, UK. Email: T.A.Fairfield@lse.ac.uk

Abstract

Bayesian analysis has emerged as a rapidly expanding frontier in qualitative methods. Recent work in this journal has voiced various doubts regarding how to implement Bayesian process tracing and the costs versus benefits of this approach. In this response, we articulate a very different understanding of the state of the method and a much more positive view of what Bayesian reasoning can do to strengthen qualitative social science. Drawing on forthcoming research as well as our earlier work, we focus on clarifying issues involving mutual exclusivity of hypotheses, evidentiary import, adjudicating among more than two hypotheses, and the logic of iterative research, with the goal of elucidating how Bayesian analysis operates and pushing the field forward.

Keywords: Bayesian analysis, process tracing, Bayesian methods, qualitative research

1 Introduction

As scholars who have contributed to the development of Bayesian process tracing, we are pleased that this approach is drawing critical engagement. "Updating Bayesian(s)" (Zaks 2020, hereafter UB), raises issues that we have frequently encountered in the work of other methodologists as well, and in questions from students. We welcome the opportunity to clarify what in our current view is the correct way to apply Bayesian reasoning in qualitative research. Specifically, we focus on how to (1) conceptualize rival hypotheses, (2) assess evidentiary support, (3) adjudicate among multiple hypotheses, and (4) refine theory and proceed with inference when the evidence we discover suggests a new explanation. We conclude with a distinctly more positive view regarding the state of the field, highlighting that even a basic understanding of Bayesian principles can improve inference without need to apply the full mathematical formalism, while concurring that experiments on how scholars can become better Bayesians should play a role moving forward.

2 Mutually Exclusive Hypotheses

A central point of debate in process-tracing literature regards the role and the nature of rival hypotheses. We contend that as a matter of best practice in Bayesian research and more generally, inference entails comparing hypotheses that are *mutually exclusive*—a concept from deductive logic that means only one member from a set of propositions can be true. If we are not comparing mutually exclusive hypotheses—which we take to be synonymous with rival hypotheses, as is standard throughout nearly all scientific and statistical analysis—then we are not testing explanations that can meaningfully be treated as alternative possibilities, and it would make little sense to ask which one provides the best explanation.

However, social scientists often worry that mutual exclusivity precludes causal complexity. UB (p.10) and others (e.g., Beach & Pedersen 2016, 174; Rohlfing 2014, 629–631) equate mutual exclusivity with monocausal hypotheses and hence view exclusivity as an "often incorrect" or "demanding" assumption that oversimplifies the world. We agree that monocausal hypotheses are often inadequate for social science. However, mutual exclusivity of *hypotheses* is conceptually

Political Analysis (2022)
vol. 30: 298–305
DOI: 10.1017/pan.2021.23

Published
26 July 2021

Corresponding author
Tasha Fairfield

Edited by
Jeff Gill

distinct from exclusivity of their constituent independent variables, causal factors, or mechanisms. We can always construct a set of causally complex, but mutually exclusive hypotheses from a set of variables that need not operate in isolation (F&C, 2017, 366).

Consider a simple example, where suspect A and suspect B serve as possible causal factors. Explanations H_A = "A committed the crime alone," and H_B = "B committed the crime alone" are exclusive by construction—if H_A is true, then H_B cannot also be true, and vice versa. That one might pose a hypothesis H_C = "A and B colluded" does not render H_A and H_B nonexclusive, nor does the exclusivity of H_A and H_B prevent us from also considering H_C. Instead, we now have three mutually exclusive hypotheses constructed from two causal factors. No two of these hypotheses can simultaneously be true, and Bayesian analysis proceeds as usual, by asking which hypothesis makes the evidence more likely (Section 3). Stated differently, each hypothesis postulates a distinct possible world, in which a *different explanation* holds true. Our task is then to figure out which of these hypothetical worlds best approximates the one reality in which we actually live, by looking for evidence that discriminates among them.

The same distinction between hypotheses and variables applies in quantitative research. When fitting data, $Y = 2X_1 + 5X_2$, $Y = 4X_1 + 3X_2$, and $Y = 2X_1 + 7X_1X_2$ are mutually exclusive hypotheses that invoke the same independent variables (X_1 and X_2) in different ways—at most one of these hypotheses can be the correct model for the data-generation process. Importantly, it is neither a particular variable of interest, nor its regression coefficient, that constitutes a hypothesis; instead, the full functional relationship serves as the hypothesis. Whereas scholars sometimes present arguments of the form "X_i matters for Y," or "X_i is a cause of Y," these kinds of propositions are not adequately specified for hypothesis comparison. Whether we are working with quantitative models, or qualitative explanations, the corresponding hypotheses should aim to articulate how and to what extent X_i matters, in relation to any other causal factors deemed important for explaining the outcome (Appendix A expounds this point).

As a best practice, we recommend trying to use language that forces hypotheses to be exclusive, while also directly asserting that the hypotheses should be read as mutually exclusive. This is the approach adopted in F&C (2019, 159)'s state-building example, where each hypothesis begins by postulating: "X_i is the central factor hindering institutional development..." (X_i = resource wealth for H_R, vs. absence of warfare for H_W). This wording should clearly convey that H_R and H_W cannot both be true, but in case any doubt remains, we also explicitly remind readers that these hypotheses are "assumed mutually exclusive." While this phrase appears to have created some confusion, it is nothing more than a formalism for conveying that the theories articulated are intended as rivals and are to be understood as such. We emphasize again that asserting these two explanations to be exclusive does not preclude considering a more complex hypothesis $H_{R/W}$ postulating that both resource wealth and absence of warfare contribute significantly to institutional weakness in some specified manner, nor does the possibility that $H_{R/W}$ might provide a better explanation than H_R or H_W "violate the mutual exclusivity assumption" (UB, 10). We can always expand the hypothesis set to include hypotheses like $H_{R/W}$; H_R and H_W will remain exclusive by construction.

As an alternative to specifying a concrete set of mutually exclusive rivals, some have proposed comparing H directly to its logical negation, $\overline{H}$. Earlier work on Bayesian process tracing, including some of our own (Bennett 2015) took this approach. However, subsequent work has clarified that we cannot compare H versus $\overline{H}$ except in very simple situations, because $\overline{H}$ will generally be an ill-specified proposition that includes an enormous range of different possibilities (F&C, 2017, 366, 375). For instance, $\overline{H_A}$ in the crime example amounts to the assertion that one of 7-billion-some individuals on the planet other than suspect A (or collusions among them in various combinations) committed the crime. It will be essentially impossible to "inhabit the world" of $\overline{H_A}$ and ask how likely this hypothesis makes the evidence, because $\overline{H_A}$ is actually an amalgamation of many very different alternative worlds, each of which could assign a different likelihood to the evidence. Moreover, the more specific the working hypothesis is, the more possibilities are embodied in $\overline{H}$.[1] To set up a well-defined inferential problem, we must specify which alternatives we consider plausible. We are free to provisionally restrict our attention to a reasonable number of alternatives—if another hypothesis comes to mind later on, we proceed as discussed in Section 5.

Another proposed alternative to working with mutually-exclusive hypotheses is the "Relationships Among Rivals" (RAR) framework developed in Zaks (2017) and advocated in UB (p. 10), which is presented as an "expansion to Bayes' rule to accommodate nonexclusivity." The premise behind RAR is that Bayes' rule is not valid unless the hypotheses are exclusive, so it must be revised to apply more broadly. But this is not the case—Bayes' rule holds universally. If despite advice to the contrary, one were to work with nonexclusive hypotheses, Bayes' rule becomes far more awkward to use, but it remains perfectly valid. We must simply apply the standard rules of probability to each term (Appendix A, equation A7). In fact, Cox (1961)'s Theorem dictates that any effort to modify Bayes' rule that alters the laws of probability, as RAR does, will introduce mathematical inconsistencies. We demonstrate how the RAR framework generates mathematical contradictions in Appendix A.

In sum, the best approach is to craft well-posed, mutually exclusive rivals. This imperative in no way restricts the extent or the nature of the causal complexity we choose to invoke in our explanations. We must simply articulate the causal processes and interactions we have in mind and how each hypothesis differs from rivals. We can always ensure exclusivity for any number of complex hypotheses through careful construction, regardless of how many variables they invoke. In most situations, this task simply requires a bit of thought, common sense, and careful wording (Appendix B provides examples).

3 Evidentiary Import

A related debate regards the nature of evidentiary import, or probative value—namely, how much support the evidence itself lends to a hypothesis over rivals. Perhaps because scholars do not always distinguish between nonexclusive causal factors that could simultaneously contribute to an outcome, and rival hypotheses that posit distinct explanations invoking one or more of those causal factors, evidentiary import is often portrayed in ways that diverge from Bayesian principles. A common example is asserting that evidence can bear on the plausibility of a hypothesis without implications for the rivals under consideration (e.g., Beach and Pedersen 2016, 175; RAR, 351; Rohlfing 2014). UB (p. 11) extends this line of thinking, concluding that: "If the evidence supports one hypothesis but is unrelated to the other, this equation [Bayes' rule] introduces disconfirmation bias—in which neutral evidence acts as an undue penalty against the unrelated theory."

Recall that the odds ratio form of Bayes' rule equates posterior odds (how plausible one hypothesis is relative to a rival, given evidence E and background information $\mathcal{I}$) to the prior odds multiplied by the likelihood ratio:

$$\frac{P(H_i \,|\, E\,\mathcal{I})}{P(H_j \,|\, E\,\mathcal{I})} = \frac{P(H_i \,|\, \mathcal{I})}{P(H_j \,|\, \mathcal{I})}\,\frac{P(E \,|\, H_i\,\mathcal{I})}{P(E \,|\, H_j\,\mathcal{I})}. \tag{1}$$

The likelihood ratio determines the inferential import of the evidence. Accordingly, if H_i makes E more likely than does H_j, then E supports H_i over H_j, and vice versa. Within a Bayesian framework, "neutral evidence" could only mean *uninformative evidence* that is *equally likely* under H_i and under H_j, such that the likelihood ratio equals 1 and there is no updating. "Neutral evidence" thus

1 For example, if we replace H_A with a more specific hypothesis H'_A = "A committed the crime alone using a candlestick," then beyond a myriad of hypotheses involving other suspects, $\overline{H'_A}$ also includes all possibilities in which A committed the crime alone, but using a different weapon.

penalizes neither hypothesis, contrary to the assertion that Bayes' rule can introduce bias. Rather, Bayes' theorem is the uniquely consistent and unbiased way to update probabilities.[2]

The crucial point is that hypotheses must always be compared, and evidentiary support is always relative to a specified pair of rival hypotheses. Talk of "finding separate, but disconfirming evidence for the alternative hypothesis" (UB, 11; RAR, 345), which we also observe in broader literature on process tracing (e.g., Ricks and Liu 2018),[3] overlooks this fundamental Bayesian principle. The question we must ask is not whether the evidence fits with a particular hypothesis, but whether and to what degree the evidence fits *better* with that hypothesis as compared to rivals. We cannot even begin to ascertain whether E supports H_i before identifying a rival hypothesis—E may seem like just what we expect under H_i, but E could be even *more* likely under H_j, in which case E undermines H_i relative to H_j. And however unlikely under H_i, E supports H_i over H_j if that evidence is even *less* plausible under H_j. In practice, we must mentally inhabit the world of *each hypothesis* in turn and ask how likely the evidence would be, even if E on its face seems "unrelated" to the causal explanation that one of the hypotheses proposes.[4]

4 Comparing Multiple Hypotheses

When working with multiple hypotheses, we use Equation (1) to conduct pairwise comparisons. We briefly illustrate this process and how evidence can adjudicate among rival explanations that share some of the same causal factors by considering another crime example:

H_A = *A was solely responsible for the murder.*
H_B = *B was solely responsible for the murder.*
H_C = *B lured the victim to the pre-arranged crime scene, where A, who was lying in wait, committed the murder.*

Assuming no relevant background knowledge about the suspects, we take all prior odds ratios to equal one. Suppose E_1 = *credit card receipts place B out of town on the night of the crime.* It should be intuitively clear that E_1 is uninformative with respect to H_A versus H_C, but this evidence weighs very strongly in favor of H_A over H_B and in favor of H_C over H_B. Notice that while B's absence may not seem "relevant" to H_A, or may not on its face seem to support H_A, E_1 nevertheless favors H_A over H_B, because this evidence is far more likely if H_A is true than if H_B is true.[5] Now consider E_2 = *a voice message was recovered in which B asks the victim to deliver documents to an office near the crime scene shortly before the murder occurred.* This evidence strongly favors H_C over H_A, because while quite plausible under the collusion hypothesis, E_2 would be coincidental and hence highly unlikely if A orchestrated the murder singlehandedly. We view E_2 as moderately favoring H_C over H_B—if acting alone, it seems more unlikely that B would be careless enough to leave an incriminating message.[6] Overall, the joint evidence E_1E_2 strongly supports H_C over H_A (thanks to E_2) and very strongly supports H_C over H_B (thanks to both E_1 and E_2). Starting with equal prior odds, the posterior odds identify H_C as the clear front-runner, followed by H_A and then H_B.

As social scientists, our task may be harder than solving crimes. It can be challenging to find evidence that strongly distinguishes among complex theories that invoke common causal factors in distinct ways, and we may not end up with evidence that adjudicates among them as decisively as we would like. But the principle remains the same, regardless of how many hypotheses we

2 Appendix D addresses issues regarding UB's discussion of confirmation bias that follow from these points.
3 They recommend scholars first engage in "finding evidence for [the] primary hypothesis," and then in a separate step proceed to "finding evidence for [the] rival hypothesis" (with the goal of ruling out alternatives).
4 F&C, forthcoming, provide examples drawing on exercises from our American Political Science Association and Syracuse Institute for Qualitative and Multi-Method Research workshops.
5 E_1 is not impossible under H_B—B could have somehow faked the receipts—but we consider that task difficult and hence consider this scenario unlikely.
6 Under H_B, B ought to be more circumspect, knowing that the alibi is false and might unravel.

are comparing or how complex they may be: assess likelihood ratios to evaluate how strongly the evidence favors one explanation over another. Our posterior odds then convey how much confidence we hold in each hypothesis relative to rivals, given the limited knowledge we possess. Bayesianism is a prescription for rational reasoning under incomplete information, and the goal is to give a well-justified assessment of how much uncertainty surrounds our conclusions.

As the number of hypotheses under consideration grows, UB (p. 5) worries that Bayesian analysis becomes "orders of magnitude more complex" due to a "combinatorics problem." However, for n hypotheses, there are only $(n-1)$ independent likelihood ratios. To see why, suppose we have assessed likelihood ratios for one hypothesis, say H_1, relative to each of its $(n-1)$ rivals, $H_2, H_3, \ldots, H_n$. We can then readily calculate likelihood ratios for all other pairs:

$$\frac{P(E \mid H_j \mathcal{I})}{P(E \mid H_k \mathcal{I})} = \frac{P(E \mid H_j \mathcal{I})/P(E \mid H_1 \mathcal{I})}{P(E \mid H_k \mathcal{I})/P(E \mid H_1 \mathcal{I})}. \tag{2}$$

Accordingly, if the number of hypotheses increases from n to $(n+1)$, we need only analyze *one* additional likelihood ratio for each piece of evidence; there is no combinatoric explosion. We must simply compare H_{n+1} to some reference hypothesis (e.g., H_1). This reference hypothesis H_1 can be any hypothesis that is convenient; in practice, we often take it to be the explanation that the author of the study proposes. Likelihood ratios for all other pairs of hypotheses can then be immediately obtained from equation (2). For examples that compare four or five rival hypotheses, as well as guidelines for aggregating evidence into a manageable number of parcels, see F&C (2017 and forthcoming).[7]

5 Iterative Research

Rather than proceeding linearly from hypothesizing to testing, qualitative research commonly iterates among theory development, data collection, and data analysis. Bayesianism provides a methodological foundation for this practice, as first recognized by McKeown (1999). Below, we reprise how iterative research works and then clarify some important points about temporality and inference that underpin the Bayesian approach.

The central question is how to proceed when we discover evidence that inspires a hypothesis that we had not previously contemplated. The steps are: (a) initiate a revised inferential problem with an expanded set of hypotheses that includes the newly inspired alternative; (b) revisit the initial background information ($\mathcal{I}$) to assess prior odds for the new hypothesis; (c) evaluate likelihood ratios involving the new hypothesis for all evidence in hand; (d) combine prior odds with likelihood ratios to arrive at posterior odds given our full present state of knowledge ($E\,\mathcal{I}$); and potentially (e) collect additional evidence.

F&C (2019, 161–63) illustrate this process with an example that begins with two hypotheses, H_R and H_W. A first piece of evidence E_1 favors H_R over H_W. A second piece of evidence E_2 inspires a new alternative, H_{LRA}, which then requires going back to the beginning and setting up a new inferential problem that seeks the best explanation among three (rather than two) hypotheses. After revising the hypothesis set, F&C (2019) reassign equal prior odds in light of the original background knowledge alone. They proceed to incorporate E_1, and then E_2, by evaluating likelihood ratios for each piece of evidence under H_{LRA} versus H_R and H_{LRA} versus H_W (F&C 2019, 162, equation 5). If desired, likelihood ratios for H_R versus H_W can be calculated via Equation (2) above. Combined with the prior odds, these likelihood ratios yield updated posterior odds in light of $E_1 E_2$. UB (p. 4 & Appendix p. 1) seems to have overlooked the step in which F&C (2019) incorporate E_1, suggesting that after assigning equal priors, they ignore E_1 and update

7 Appendix C addresses UB's notion that Equation (1) is somehow inadequate, which may have contributed to concerns about a combinatoric explosion.

based on E_2 alone. UB's core critique that the literature "lack[s] the necessary guidelines for executing a systematic, iterative process" (p. 5) and "exhibits contradictory practices (e.g., not using updated probabilities when analyzing subsequent pieces of evidence)" (p. 15) accordingly rests on a misreading, which we hope is now resolved.

More broadly, the key to understanding iterative research lies in distinguishing among different aspects of temporality and how they do, or do not matter for inference:

(1) temporal claims made by a hypothesis as part of the causal explanation it proposes,
(2) temporal information contained in the evidence,
(3) the sequence in which we analyze the evidence, and
(4) the timing of when one learns the evidence relative to when a hypothesis was devised.

Among these, (1) and (2) certainly matter for inference—what we learn about how events unfolded may weigh strongly in favor of one hypothesis relative to rivals that predict different sequencings of events. In contrast, (3) is irrelevant for the mathematics of Bayesian analysis. The joint likelihood of two pieces of evidence can be written in any of the following equivalent ways: $P(E_1 E_2 \mid H \mathcal{I}) = P(E_2 E_1 \mid H \mathcal{I}) = P(E_1 \mid E_2 H \mathcal{I}) P(E_2 \mid H \mathcal{I}) = P(E_2 \mid E_1 H \mathcal{I}) P(E_1 \mid H \mathcal{I})$, so we are free to analyze evidence in any convenient order. Additionally, note that we need not present evidence in a case narrative in the same order that we analyzed the inferential import of each piece of evidence—narratives generally follow the sequential causal story that the evidence suggests, independently of the order in which the evidence was analyzed or learned.[8] F&C (2019) further contend that (4) should play no role in inference. Whereas the linear-deductive model requires that hypotheses be tested exclusively with "new evidence" learned after hypotheses were devised, Bayesianism mandates that *all* evidence should be taken into account when adjudicating between hypotheses, regardless of whether we learned that evidence before, or after formulating hypotheses. This latter point in particular underpins the Bayesian approach to iterative research that F&C (2019) elaborate.

We hope these distinctions help clarify that a Bayesian approach does not posit the "global irrelevance of timing" (UB, p. 7), nor does it "disregard the role of new evidence." F&C (2019, 154) instead argue that "new evidence has no special status relative to old evidence for testing hypotheses," in the sense that "old evidence" (known before hypothesis formulation) and "new evidence" (learned after hypothesis formulation) must both inform the posterior odds. New evidence is of course valuable for additional testing. But in this view, it does not have any extra import over and above the inferential weight associated with its likelihood ratio simply because of the relative timing of when we learn the information (F&C 2019, 162–163).

On this point, there are longstanding debates in the philosophy of science about giving special value to "use-novel" evidence that was not used to build a hypothesis. One intuition for valuing "use-novelty" is psychological: "novel" evidence may be less subject to confirmation bias. Another line of reasoning holds that evidence which helped generate a hypothesis might seem incapable of undermining confidence in that hypothesis. Referencing objective Bayesianism, F&C (2019) take on the latter view by arguing that even if evidence was used to build a hypothesis, its inferential value derives from its logical relationship to competing hypotheses, which in principle should be no different whether the evidence was evaluated before or after the hypotheses were constructed. Recognizing that it may be impossible to achieve this standard in practice due to cognitive biases, F&C (2020 and forthcoming) recommend various robustness checks on likelihood-ratio estimates. The point is that we must first understand what is rational in logical terms. Then if psychological biases lead scholars to depart from principles of rationality, we should devise and experiment with procedures to help them become better Bayesians.

8 See Appendix D for further elaboration.

Further to these points, UB's call for a "demonstrat[ion] that researchers are capable of arriving at both consistent and accurate conclusions irrespective of the sequence in which they see evidence" misses a key advantage of Bayesianism. We can use the mathematical requirement that evidence analyzed in different but logically equivalent ways must yield the same posterior odds to conduct consistency checks on our reasoning. As fallible humans, we may very well arrive at different conclusions when analyzing evidence in different orders—in which case we should reevaluate our reasoning and seek to identify and resolve the inconsistencies. For examples of how and when to use these consistency checks, see F&C (2020 and forthcoming). Additional queries from UB about timing and sequencing are addressed in Appendix D.

6 State of the Field

We are convinced that Bayesianism has much to offer qualitative research. It provides a rigorous foundation for inference to best explanation as well as a coherent framework for consensus building and knowledge accumulation. While Bayesian analysis can be labor intensive, in many contexts, careful assessment of likelihood ratios for just a few key pieces of evidence can significantly improve analytic transparency without imposing undue burdens. Moreover, scholars need not apply the full mathematical apparatus of Bayesian probability to reap benefits—even a basic conceptual understanding of essential Bayesian principles can improve inference.[9] Our most fundamental recommendations include working with mutually exclusive hypotheses that articulate distinct possible explanations, and remembering to ask not whether the evidence supports a hypothesis, but whether the evidence supports that hypothesis more than it corroborates rivals. We have seen many instances where authors neglect to ask whether their evidence might be equally or more consistent with a rival hypothesis, leading them to overstate the extent to which the evidence supports their argument.

Furthermore, Bayesianism helps leverage and improve intuition. Among the case-study works widely lauded as exemplars, we contend that the most compelling succeed precisely because their authors are excellent intuitive Bayesians. Yet we know from cognitive psychology that intuition often fails. Training in Bayesian analysis provides tools that can help us identify some of our own cognitive limitations and reason more rationally. Beyond making us better data analyzers, Bayesianism can also improve data collection. Simply recognizing that inferential support comes from differential likelihoods under rival hypotheses helps us identify the kinds of evidence to seek that will most effectively adjudicate among rival explanations.

We also emphasize that the state of the field has quickly progressed well beyond that depicted in UB. Fairfield and Charman's forthcoming book more extensively addresses the issues discussed above as well as additional queries that UB and others have posed—regarding for example, writing Bayesian-inspired case narratives without compromising word limits or readability, and judging when to rely on heuristic insights from Bayesianism versus quantifying probabilities and explicitly applying Bayes' rule. The book provides concrete guidelines for constructing exclusive hypotheses that are neither too simple nor overly complex, assessing the inferential weight of evidence, conducting iterative research and case selection, and using the mathematical structure of Bayesian analysis to identify and counteract cognitive biases, along with worked examples that test multiple hypotheses and identify departures from Bayesian reasoning in published case studies. Our workshops at Syracuse's Institute for Qualitative and Multi-Method Research (IQMR) and short courses at the American Political Science Association's Annual Meetings feature instruction on many of these points. Through these venues, we look forward to continuing the dialog about how Bayesianism can inform and improve qualitative research.

9 See "A Spotlight on Low-Tech Best Practices," in F&C (forthcoming, Chapter 1): https://tashafairfield.wixsite.com/home/bayes-book

PA

Finally, regarding the future of Bayesian process tracing, we agree that there is a useful role for experimental research to evaluate, for example, whether training in Bayesian reasoning helps scholars achieve greater intersubjective agreement on the weight of evidence (Bennett 2015, 290), and whether Bayesian analysis improves case-study inferences. These are agendas that we have been actively exploring. In addition, research on crowd-sourcing (Surowiecki 2004) and structured discussions among groups trained in Bayesian analysis (Tetlock and Gardner 2015) indicates that these techniques can improve estimates and forecasts. Experiments that apply these techniques to case study analysis might also prove fruitful.

Supplementary Material

For supplementary material accompanying this paper, please visit https://dx.doi.org/10.1017/pan.2021.23. Supplementary material is also available at https://tashafairfield.wixsite.com/home/bayes-articles.

References

Beach, D., and R. Pedersen. 2016. *Causal Case Study Methods*. Ann Arbor: University of Michigan Press.

Bennett, A. 2015. "Disciplining Our Conjectures: Systematizing Process Tracing with Bayesian Analysis." In *Process Tracing in the Social Sciences*, edited by A. Bennett, and J. Checkel, 276–298. Cambridge: Cambridge University Press.

Cox, R. 1961. *The Algebra of Probable Inference*. Baltimore: Johns Hopkins University Press.

Fairfield, T., and A. E. Charman (F&C). 2017. "Explicit Bayesian Analysis for Process Tracing." *Political Analysis* 25(3):363–380.

Fairfield, T., and A. E. Charman (F&C). 2019. "The Bayesian Foundations of Iterative Research in Qualitative Social Science." *Perspectives on Politics* 17(1):154–167.

Fairfield, T., and A. E. Charman (F&C). 2020. "Reliability of Inference: Analogs of Replication in Qualitative Research." In *The Production of Knowledge*, edited by C. Elman, J. Gerring, and J. Mahoney, 301–333. Cambridge: Cambridge University Press. URL: https://tashafairfield.wixsite.com/home/bayes-articles.

Fairfield, T., and A. E. Charman (F&C). Forthcoming. *Social Inquiry and Bayesian Inference: Rethinking Qualitative Research.* Cambridge: Cambridge University Press. Chapter 1. https://tashafairfield.wixsite.com/home/bayes-book.

McKeown, T. 1999. "Case Studies and the Statistical Worldview." *International Organization* 53(1):161–190.

Ricks, J., and A. Liu. 2018. "Process-Tracing Research Designs: A Practical Guide." *PS: Political Science & Politics* 51(4):842–846.

Rohlfing, I. 2014. "Comparative Hypothesis Testing Via Process Tracing." *Sociological Methods & Research* 43(4):606–642.

Surowiecki, J. 2004. *The Wisdom of Crowds*. New York, NY: Doubleday.

Tetlock, P., and D. Gardner. 2015. *Superforecasting: The Art and Science of Prediction*. New York, NY: Crown Books.

Zaks, S. (RAR). 2017. "Relationships among Rivals." *Political Analysis* 25(3):344–362.

Zaks, S. (UB). 2020. "Updating Bayesian(s): A Critical Evaluation of Bayesian Process Tracing." *Political Analysis* 29(1):58–74.