Understanding Bayesianism: Fundamentals for Process Tracers
Supplemental Materials

Andrew Bennett, Georgetown University

Andrew Charman, University of California–Berkeley

Tasha Fairfield, London School of Economics, T.A.Fairfield@lse.ac.uk

Political Analysis, February 2021

Appendix A: Bayes' Rule and (non)Exclusive Hypotheses

We emphasized in Section 2 of our article that Bayes' rule holds universally. In fact, Bayes' rule itself, in its most basic form:

$$P(H \mid E\mathcal{I}) = \frac{P(H \mid \mathcal{I})\, P(E \mid H\mathcal{I})}{P(E \mid \mathcal{I})}\,, \tag{A1}$$

is simply a rearrangement of the product rule of probability:

$$P(AB \mid C) = P(A \mid BC)P(B \mid C) = P(B \mid AC)P(A \mid C)\,, \tag{A2}$$

as can be demonstrated by substituting H for proposition A, E for B, and background information $\mathcal{I}$ for C. Accordingly, any alteration of Bayes' rule will violate this fundamental law of probability and will inevitably produce logical contradictions. This reality follows from Cox's Theorem, which ensures that any purported extension or modification to probability theory will either end up simply recapitulating the original probability calculus, or will otherwise introduce inconsistencies with logic or common sense. In fact, Cox (1961) and Jaynes (2003) proved that Bayesian probability theory emerges as the uniquely consistent extension of deductive (Boolean) logic to situations where propositions are either true or false, but we do not know these truth values with certainty.[1] No modifications or extensions are required, or allowed, upon pain of logical inconsistency.

RAR nevertheless maintains that Bayes' rule "requires extensive modification when dealing with nonexclusive theories" and proceeds to devise non-standard rules for evaluating probabilities

[1] E.T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003; Richard Cox, *The Algebra of Probable Inference*, Baltimore: Johns Hopkins University Press, 1961.

in different contexts, depending on which of various "types" of non-exclusive hypotheses are considered.[2] To illustrate one example of how RAR's rules produce logical contradictions, as Cox's theorem dictates it must, consider the following proposed relationship for evaluating the likelihood of a binary clue K under the conjunction of two "congruent hypotheses" that "work together to bring about the outcome" (RAR:348):

$$P(K = 1 \,|\, H_1 \cap H_2) = 1 - P(K = 0 \,|\, H_1)\, P(K = 0 \,|\, H_2) \text{ (RAR equation 5),}$$

or in more compact notation:

$$P(K \,|\, H_1 H_2) \stackrel{?}{=} 1 - P(\overline{K} \,|\, H_1)\, P(\overline{K} \,|\, H_2). \tag{A3}$$

Applying the negation rule, $P(K \,|\, H_1 H_2) = 1 - P(\overline{K} \,|\, H_1 H_2)$, to the left-hand side and rearranging, we obtain:

$$P(\overline{K} \,|\, H_1 H_2) \stackrel{?}{=} P(\overline{K} \,|\, H_1)\, P(\overline{K} \,|\, H_2). \tag{A4}$$

This relationship is meant to hold when "K is expected with some positive probability under both H_1 and H_2" (RAR:354), which implies that $P(K \,|\, H_1) > 0$ and $P(K \,|\, H_2) > 0$. Now suppose that $0 < P(K \,|\, H_1) = 1 - P(\overline{K} \,|\, H_1) < 1$, and $0 < P(K \,|\, H_2) = 1 - P(\overline{K} \,|\, H_2) < 1$, i.e., none of the likelihoods is certain. Next, consider that "congruent hypotheses" include as a special case "inclusive hypotheses," where one hypothesis is an "extension of an existing theory" (RAR:350). This definition presumably encompasses the situation where a narrower theory H_1 is logically implied as a special case of a more general theory H_2, such that $P(\overline{K} \,|\, H_1 H_2) = P(\overline{K} \,|\, H_2)$, in which case the left-hand side of (A4) simply reduces to $P(\overline{K} \,|\, H_2)$, and we have:

$$P(\overline{K} \,|\, H_2) \stackrel{?}{=} P(\overline{K} \,|\, H_1)\, P(\overline{K} \,|\, H_2). \tag{A5}$$

Dividing by the strictly positive quantity $P(\overline{K} \,|\, H_2)$, and remembering that $P(\overline{K} \,|\, H_1) < 1$ by assumption, we find:

$$1 \stackrel{?}{=} P(\overline{K} \,|\, H_1) < 1, \tag{A6}$$

which is a contradiction, since 1 cannot be less than 1.

Working with non-exclusive hypotheses is technically permitted by the ordinary rules of probability, but it is awkward and inefficient. If one nevertheless opts for non-exclusive hypotheses, no

[2] These proposed rules are presented without justification other than "intuition" (RAR p.354) from Venn-like diagrams.

modifications or additions to ordinary probability theory are required; one must just carefully adhere to the dictates of the sum rule and the product rule. In particular, Bayes' rule (A1) still holds.[3] However, the prior $P(H \mid \mathcal{I})$ is more difficult to think about, because the probability assigned to one non-exclusive hypothesis does not necessarily come at the expense of another. And the marginal likelihood $P(E \mid \mathcal{I})$, which serves to normalize the overall probability, can no longer be expanded using the law of total probability, $P(E \mid \mathcal{I}) = \sum_{j=1}^{n} P(H_j \mathcal{I}) P(E \mid H_j \mathcal{I})$, which does not hold for non-exclusive hypotheses. Expanding the marginal likelihood instead requires the more complicated inclusion-exclusion rule (which can be derived from the sum and product rules):

$$\begin{aligned} P(E \mid \mathcal{I}) = &\sum_j P(H_j \mid \mathcal{I})\, P(E \mid H_j \mathcal{I}) - \sum_j \sum_{k>j} P(H_j H_k \mid \mathcal{I})\, P(E \mid H_j H_k \mathcal{I}) \\ &+ \sum_j \sum_{k>j} \sum_{\ell>k} P(H_j H_k H_\ell \mid \mathcal{I})\, P(E \mid H_j H_k H_\ell \mathcal{I}) - \ldots \\ &+ (-1)^{n-1} P(H_1 \cdots H_n \mid \mathcal{I})\, P(E \mid H_1 \cdots H_n \mathcal{I}) \,, \end{aligned} \tag{A7}$$

involving probabilities of conjunctions of multiple hypotheses, as well as likelihoods conditioned on multiple hypotheses.

The best way to deal with the complications of non-exclusive hypotheses is to avoid them. As argued here and elsewhere (Fairfield & Charman 2017 and forthcoming), constructing mutually exclusive hypotheses does not require a strong or onerous modeling assumption, nor does it limit in any way the causal complexity that can be captured. Instead, ensuring that our hypotheses are exclusive usually just requires a bit of attention to wording, along with an awareness that exclusive hypotheses can reference non-disjoint causal variables (Appendix B).

Far from a limitation of Bayesian analysis, as UB (p.10) frames it, the requirement to work with a set of carefully constructed, mutually exclusive alternatives is actually a strength of the Bayesian approach that forces us to be clear about the hypothesis we are proposing and how it diverges from alternatives. UB (p.10) is correct that many examples can be found where scholars use hypotheses that "implicitly take the form 'this matters, too,'" but as we emphasized in Section 2, such hypotheses are too vague—we can and should aim to do better. For how are other scholars to interpret a claim that "X_i matters for Y"? Does this mean that

[3] Bayes' rule in the odds-ratio form also still holds, contrary to UB's claim (p. 10) that "Mathematically, when two nonexclusive hypotheses are treated as though exclusivity holds, nearly every term in the equation is inaccurate." What is true is that the odds-ratio form of Bayes' rule may not be convenient, or informative, if applied directly to non-exclusive hypotheses.

we are seeking to improve a leading theory from the literature by including X_i? If so, how salient is the role of X_i relative to other causal factors in the theory? Is X_i a minor contributor that entails an incremental revision of the existing theory? Or is X_i to be taken as a more important contributing factor than the variables in the leading theory, such the "X_i matters" claim constitutes a major departure from existing explanations? Equally important, especially from a process-tracing perspective, how does this variable operate alongside or in conjunction with other causal factors? What is the mechanism or causal process? Does X_i operate independently of other contributing causes, or through some kind of interaction? Assertions that "X_i matters" can have a role in early stages of theory development. But until we begin to answer these kinds of questions about how X_i operates and how salient it is relative to other causes, we have taken only a preliminary step toward theory development, and we cannot meaningfully test our hypothesis against alternatives—recall that hypotheses must be specific enough to be able to "mentally inhabit" the corresponding world and evaluate likelihoods for the evidence. Stated in more technical terms, $H = X_i$ *matters* is too vague a proposition upon which to condition probabilities, such that $P(E \mid H\,\mathcal{I})$ will be ill-defined, or at least impossible to assess even approximately.

Appendix B: Constructing Mutually Exclusive Hypotheses From Contributing Causes

Readers may wish to see explicit political science illustrations to better understand how we can construct a set of mutually exclusive hypotheses that share causal variables. We offer two such examples below; the first draws on Slater's (2009) research on democratic mobilization, and the second reworks UB's (Section 4.2) greed, grievance, and rebellion example.

1. Communal Elites, Stolen Elections, and Democratic Mobilization

In this example, we begin with two causal factors—autonomous communal elites (X_1) and stolen elections (X_2)—that authors have theorized to be salient for eliciting democratic mobilization against dictatorship (Y). Hypotheses H_A and H_B below each focus on one or the other of the two causal factors, whereas H_C, H_D, and H_E invoke both of the causal factors in distinct ways (Fairfield & Charman, forthcoming). In essence, each of these hypotheses tells a different story about how democratic mobilization comes about. They are all mutually exclusive, in that no two of these stories can simultaneously be true.

> H_A = Autonomous communal elites are critical agents for eliciting democratic mobilization (Slater 2009).[4] Emotive appeals that these actors make, invoking politically salient nationalist and religious sentiments and solidarities, are the central factor that sparks and sustains collective action against dictatorship. When autonomous communal elites oppose the regime, democratic mobilization is probable; if these actors are absent, mobilization is unlikely.
>
> H_B = Stolen elections are the central factor that elicits democratic mobilization, by revealing the regime's lack of legitimacy, quashing expectations of change, and creating a focal event to catalyze collective action.
>
> H_C = Autonomous communal elites and stolen elections are jointly critical factors for eliciting democratic mobilization. The combination of communal elites opposing the regime and stolen elections provokes outrage, solves collective action problems, and catalyzes protest. In the absence of either (or both) of these factors, democratic mobilization is unlikely—neither causal factor on its own is enough to catalyze and sustain collective action.
>
> H_D = Autonomous communal elites are critical agents for eliciting democratic mobilization. In addition, stolen elections contribute to democratic mobilization, by giving communal elites extra motivation to oppose the regime and by making citizens even more likely to respond to their appeals. When autonomous communal elites oppose the regime, democratic mobilization is probable. When stolen elections also occur, the probability of mobilization increases, and protests are likely to be even more massive. In contrast, democratic mobilization is unlikely when autonomous communal elites are absent.
>
> H_E = Autonomous communal elites and stolen elections both contribute independently to eliciting democratic mobilization. The probability and expected scale of mobilization increases with the net effect of these two factors.

We could of course propose additional hypotheses that posit even more complicated causal sto-

[4] Dan Slater, "Revolutions, Crackdowns, and Quiescence: Communal Elites and Democratic Mobilization in Southeast Asia," *American Journal of Sociology* 115(1):203-254, 2009.

ries involving communal elites and stolen elections, but a central recommendation that emerges from the Bayesian framework (i.e., Occam penalties, Appendix D below) is that we should begin by assessing simpler hypotheses before proposing highly complex and intricate explanations. For that reason, a natural approach to the problem at hand would be to begin by comparing H_A and H_B, and only consider hypotheses H_C, H_D, or H_E in a later round of inference if the evidence suggests that they might significantly improve explanatory leverage.

Here we would point out that while some scholars are of the opinion that qualitative research by and large does not consider monocausal hypotheses, there are actually many such examples, including Slater (2009), which we draw on here. That is, Slater does not simply suggest that autonomous communal elites are "a cause of" democratic mobilization; he argues that they are in fact the primary cause of democratic mobilization. In accord with the Bayesian guidelines that we recommend, Slater (2009:206, 207 footnote 7, 226-27) explicitly pits his communal elites explanation, which corresponds to H_A above, against the stolen elections hypothesis H_B and several other rivals that each invoke a different central causal factor (e.g., economic decline).

Note also that each of the hypotheses above articulates a process that aims to clarify "how" and "why" the causal factors matter for the outcome. Bayesians, like process tracers more generally, should aim to craft hypotheses that are more specific than simply positing "a causal arrow" leading from X to Y. Some might prefer to see even more detail when articulating causal processes (Beach and Pedersen 2019). And of course, Slater (2009) provides more context and discussion in developing his communal elites hypothesis than we have captured in our summary statement, H_A. Nevertheless, the above hypotheses are reasonably well-articulated rivals—they include enough detail for us to "mentally inhabit the world" that each describes and assess how likely various concrete pieces of evidence that Slater uncovers would be in each of the respective worlds (Fairfield & Charman, forthcoming: Chapters 5 & 6).

2. Greed, Grievance, and Rebellion

We emphasize again that constructing mutually exclusive hypotheses is simply a matter of good housekeeping, to "make our work much neater," as one introductory Bayesian textbook puts it.[5] The author proceeds to note: "People sometimes carelessly list sets of alternatives which violate these conditions, but the situation can usually be remedied," (Schmitt 1969:12). Let us put this basic housekeeping principle to work and show how carefully defining the hypothesis space immediately resolves the concerns that UB poses in Section 4.2 (p.10-11).

UB's example is set up to compare "two hypotheses that could jointly be true: whether greed or grievance motivates participation in rebellion."[6] UB asserts that "literature is unclear on how to proceed" in these instances, when hypotheses take the form of "this matters, too." To reiterate the Bayesian guidance: avoid working with such hypotheses. Instead, construct mutually exclusive alternatives that articulate how the particular causes of interest matter along with any other causes deemed important for constructing a compelling explanation.

In UB's example, as a starting point we would do well to include simple possibilities of the form: H_1 = *Greed is the primary factor motivating rebellion*, and H_2 = *Grievance is the primary factor motivating rebellion*, which are exclusive by our own deliberate construction. We might find that one or the other of these hypotheses does a very good job of explaining the evidence, even if the real world is somewhat more complex. Importantly, one should also specify some mechanism or concrete causal logic that explains how and why grievance matters in H_1, and similarly for H_2, but we leave that exercise to readers.

We can of course include more complex hypotheses that invoke both greed and grievance in our set of mutually exclusive explanations. But here too we need to articulate well-specified alternatives. Simply claiming that "both matter" is too vague a hypothesis to make concrete evidentiary predictions, such that $P(E \mid H\mathcal{I})$ cannot be assessed. We have to explain how greed and grievance together produce the outcome, and there are many different possibilities one could envision. Examples could include something along the following lines:

> H_3 = Greed and grievance both make independent contributions to motivating rebellion. The probability of rebellion increases with the net effect of either or both factors, once enough group members have been pushed past a threshold of

[5] Samuel Schmitt, *Measuring Uncertainty: An Elementary Introduction to Bayesian Statistics.* Addison-Wesley, 1969.

[6] Here we would point out that a standard reading of "greed or grievance" would interpret "or" as an exclusive disjunction, suggesting two mutually exclusive hypotheses, but in stating that these hypotheses can operate jointly, UB clearly has in mind something along the lines of H_1 = "greed is a cause of rebellion" and H_2 ="grievance is a cause of rebellion."

indifference.

H_4 = The combination of greed among group leaders and grievance among group members is critical for rebellion. When group members experience high levels of grievance, leaders who are motivated by greed are able to capitalize on discontent, foment rebellion, and use the power it grants them to pursue their own personalistic agenda. In the absence of widespread grievance, leaders are not able to foment rebellion, while leaders motivated by factors other than greed find it difficult to take the drastic steps necessary to translate grievance into rebellion.

Note that all four hypotheses, H_1, H_2, H_3, and H_4 are to be understood as mutually exclusive alternatives. We emphasize again that asking if greed matters and if grievance matters may be a reasonable starting point at very early stages of theorizing, but hypotheses must be better developed and better articulated before they can be tested.

Once the hypothesis space is set up properly, Bayesian analysis proceeds as usual—we assess prior odds and evaluate likelihood ratios for evidence under pairs of rival hypotheses. None of the various quantities UB (p.11) proposes, involving differences of probabilities or probabilities of conjunctions of hypotheses, need to be computed (nor would these be the correct calculations if the hypotheses were nonexclusive).[7] No assumptions of any kind must be made about the evidence to "return valid results." And Bayes' rule, if properly applied, never introduces "bias." We must simply remember that a given piece of evidence supports a hypothesis H_j over a rival H_k to the extent that H_j makes that evidence more expected than does H_k. Asking if E is "relevant" to a hypothesis (UB p.11) is not the right question. We must instead ask which hypothesis makes E more expected (or equivalently, less surprising). A given hypothesis H_j may not make sharp predictions about whether we should observe E, but if a rival H_k predicts that evidence with a higher likelihood, then E weighs in favor of H_k over H_j.[8] To reiterate, we cannot even begin to ask whether evidence E supports hypothesis H_j before we have identified a rival for comparison. Indeed, evidence that supports H_j over H_k might actually *undermine* H_j

---

[7] In the "greed vs. grievance" example, UB actually seems to be working in a very indirect way with the *exclusive* hypotheses $H_\$ H_{\frownie}$, $H_\$ \overline{H_{\frownie}}$, and $H_{\frownie} \overline{H_\$}$, with prior probabilities $P(H_\$ H_{\frownie} \,|\, \mathcal{I}) = P(H_\$ \,|\, \mathcal{I})\, P(H_{\frownie} \,|\, H_\$ \mathcal{I})$, $P(H_\$ \overline{H_{\frownie}} \,|\, \mathcal{I}) = P(H_\$ \,|\, \mathcal{I}) - P(H_\$ H_{\frownie} \,|\, \mathcal{I})$, and $P(H_{\frownie} \overline{H_\$} \,|\, \mathcal{I}) = P(H_{\frownie} \,|\, \mathcal{I}) - P(H_\$ H_{\frownie} \,|\, \mathcal{I})$, respectively. But then the likelihood ratios in UB's example should also be conditioned on these conjoined hypotheses, not the original pair $\{H_\$, H_{\frownie}\}$ if regarded as non-exclusive.

[8] If the evidence is equally likely under a pair of hypotheses, it is simply uninformative with respect to that pair of hypotheses, and we should proceed to look for other evidence that does discriminate between these hypotheses.

relative to a different rival, H_m (see Fairfield & Charman 2017: 374-5 for empirical examples). Moreover, "evidence consistent with a theory can actually lower its posterior and evidence that does not fit a theory can raise its posterior" (Bennett 2015:291)—depending on how likely that evidence is under the rival hypothesis in question.

Appendix C: Odds Ratios vs. Posterior Probabilities

Concerns about a combinatoric explosion when comparing multiple hypotheses appear to stem in part from UB's (Appendix p.2) notion that the odds-ratio form of Bayes' rule (equation 1 in our letter) is somehow inadequate, because: "the probabilities in the numerator and denominator of the posterior cannot be isolated for each individual hypothesis." Under the standard practice of working with mutually exclusive and exhaustive (MEE) hypotheses, the posterior odds-ratios uniquely determine the posterior probabilities, and vice-versa. In the instance UB considers, there are three unknowns (posterior probabilities for H_M, H_R, and H_C) that can easily be solved for using the two independent odds ratios that have been evaluated—e.g., $P(H_M \,|\, E\mathcal{I})/P(H_R \,|\, E\mathcal{I})$ and $P(H_M \,|\, E\mathcal{I})/P(H_C \,|\, E\mathcal{I})$)—along with the normalization constraint $P(H_M \,|\, E\mathcal{I}) + P(H_R \,|\, E\mathcal{I}) + P(H_C \,|\, E\mathcal{I}) = 1$, which follows from extending the MEE assumption that UB makes for two hypotheses to include the third hypothesis. More generally, when working with n MEE hypotheses, $H_1, H_2, \ldots, H_n$, posterior probabilities can be obtained directly from the $(n-1)$ independent odds ratios as follows:

$$P(H_k \,|\, E\mathcal{I}) = \frac{\frac{P(H_k \,|\, E\mathcal{I})}{P(H_\ell \,|\, E\mathcal{I})}}{1 + \sum\limits_{j \neq \ell} \frac{P(H_j \,|\, E\mathcal{I})}{P(H_\ell \,|\, E\mathcal{I})}} \,. \tag{C1}$$

Appendix D: Additional Points on Bayes' Rule and Iterative Research

(a) As discussed in Section 5, if the evidence inspires a new explanation, we need to expand the hypothesis set and return to the original background information in order to assign prior odds for the new pairs. UB's online appendix instead contemplates three distinct alternatives for analyzing a newly-conceived hypothesis H_C relative to a main hypothesis H_M, after having previously compared H_M to a rival H_R in light of evidence E_1. We consider each in turn in order to clarify how Bayesian analysis operates:

- "*One option is to analyze how E_1 affects H_M relative to H_C by beginning that analysis with equal odds placed on the two hypotheses. The choice to revert to equal odds seems strange because we already know E_1 supports H_M. Though, it could help avoid bias from double-counting evidence*," (UB Appendix p.2).

Recall that the likelihood ratio in Bayesian analysis, $P(E \mid H_j \mathcal{I})/P(E \mid H_k \mathcal{I})$, tells us that evidence supports a given hypothesis H_j to the extent that it is less likely under a rival H_k. The fact that E_1 supports H_M over H_R, that is, $P(E_1 \mid H_M \mathcal{I}) > P(E_1 \mid H_R \mathcal{I})$, tells us nothing about how E_1 bears on H_M vs. H_C, because we have yet to assess how $P(E_1 \mid H_C \mathcal{I})$ compares to $P(E_1 \mid H_M \mathcal{I})$. Therefore, we cannot assert that "*we already know E_1 supports H_M*"—at the present stage of analysis, we know only that E_1 supports H_M over H_R. Furthermore, what we happen to know at a given moment in time may differ from what we have or have not already explicitly incorporated as conditioning information in formal probability statements. It is also important to stress that when applied correctly, Bayesian updating never double-counts evidence (see Fairfield & Charman 2019: 160 and Appendix A).

- "*A second option is to analyze how E_1 affects H_M relative to H_C by beginning that analysis with a slightly higher prior on H_M. This choice seems consonant with respect to updating our confidence in H_M, but problematic in that it double-counts the effect of E_1*," (UB Appendix p.2).

Our same points about the nature of evidentiary support apply here as well. Moreover, whatever we know about the likelihood ratio $P(E_1 \mid H_M \mathcal{I})/P(E_1 \mid H_R \mathcal{I})$ tells us nothing about the prior odds on H_M vs. H_C. The prior odds, $P(H_M \mid \mathcal{I})/P(H_C \mid \mathcal{I})$, excludes E_1 from the conditioning information and is therefore entirely independent of E_1.

- "*A third option is to analyze H_M relative to H_C by examining a different piece of evidence entirely, E_2. This choice is also difficult to justify because then researchers must make the case for why they chose a given piece of evidence for one analysis, but not another*," (UB Appendix p.2).

Logical Bayesianism and the principles of rationality upon which it is based never entail choosing to analyze some pieces of information while ignoring other known pieces of evidence that are relevant. Rather, "Information is never intentionally disregarded in logical Bayesianism; any subsequent stage of research following the inspiration of a hypothesis must take all previously-obtained evidence into account through the prior probability on that hypothesis," (F&C-2019:163). In sum, to proceed consistently, scholars should follow the steps explicated in Section 5 when seeking to include a new explanation in the hypothesis set.

(b) On questions of timing and sequencing, UB (p.7) asks: "Even if the order in which a researcher analyzes her evidence proves entirely irrelevant to the probabilities she assigns and conclusions she draws, why encourage her to disregard the sequence in her narrative? If we do not structure our write-ups to broadly map onto the sequence of iterative updating, what is the alternative way of structuring them?" The standard way to structure a case narrative is to follow the sequential causal story that the evidence suggests, independently of the order in which the evidence was analyzed or learned. The order of events can itself be relevant for adjudicating among alternative explanations, but this is conceptually separate from the order in which we receive or analyze evidence or the order in which we present a case narrative to make it understandable. Just as a director might film the final scene of a movie before shooting earlier scenes, we might have initially analyzed evidence in a different order than we present that evidence in a published narrative. Moreover, none of the literature that UB critiques encourages scholars to disregard causal sequences in narratives.

(c) Confirmation bias is a common pitfall associated with iterative research that goes back and forth between theory revision, data collection, and data analysis. When applied correctly, Bayesian reasoning automatically precludes a common form of confirmation bias—namely, focusing only on a single hypothesis without considering alternatives—because the key inferential step necessarily involves evaluating likelihood *ratios*. Instead of asking how expected the evidence would be if the working hypothesis is true, we must ask whether the evidence would be more expected or less expected under that hypothesis *as compared to a rival.* UB's (p.12) assertion that Bayesianism does not help to control this variant of confirmation bias appears to result from the same issues regarding the nature of evidentiary support and mutually exclusive hypotheses that we discussed in Sections 2 and 3.

(d) Occam factors are an intrinsic feature of Bayesianism that mediates the tradeoff between parsimony and accuracy (Western 2001),[9] in accord with Einstein's dictum that things should

[9] Bruce Western, 2001, "Bayesian Thinking about Macrosociology," *American Journal of Sociology* 107(2):353-

be as simple as possible, but no simpler. In essence, Occam factors are a mathematical representation of Occam's Razor that is built into Bayes' rule (Jaynes 2003, Jefferys 2003).[10] However, UB (p.13) proposes that Occam's Razor is not an appropriate principle for social science: "Occam's Razor is infrequently well-suited to social phenomena—and until someone demonstrates that the simpler explanation tends to be the right one where politics is concerned, penalizing marginal complexity is unsubstantiated." Here UB seems to overlook the fact that Bayesianism penalizes complex explanations only "if they do not provide enough additional explanatory power relative to simpler rivals," (Fairfield & Charman 2019:161 and forthcoming: Chapter 6). How does this work? A complex hypothesis incurs an Occam penalty relative to simpler rivals via its prior odds (see Fairfield & Charman 2019:161 for an intuitive illustration). If the more complex hypothesis is actually the "right explanation," its posterior odds should win out thanks to the improved inferential leverage that 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 (see Fairfield & Charman 2019: Appendix C for a mathematical illustration).

UB (p.13) asks for "pragmatic guidelines for evaluating what constitutes enough explanatory power" and suggests that there is somehow an "inherent contradiction" here within Bayesian analysis. On the latter point, we stress again that Bayes' rule is universally valid and never produces contradictions if properly applied. On the former point, the pragmatic guideline is simply to apply Bayes' rule, and examine the posterior odds to see whether the more complex hypothesis comes out ahead of the rivals. As discussed in Fairfield & Charman (2019), while Occam factors arise automatically in quantitative Bayesian model comparison, there are no cookie-cutter rules for assessing Occam factors in qualitative research, and we must use our judgment when assessing the relative complexity of our hypotheses (see also Western 2001:375). Nevertheless, readily applicable practical guidelines include starting with reasonably simple theories and adding complexity incrementally as justified by the data, scrutinizing whether all of the causal factors in a hypothesis actually improve explanatory leverage compared to simpler rivals, and asking if the hypothesis might apply more broadly. If a given hypothesis invokes many more causal factors or very elaborate conjunctions of causal factors, good practice entails penalizing its prior relative to the rivals. If an author fails to treat an especially complex hypothesis with adequate prior skepticism, other scholars should take notice and call attention

78.

[10] William Jefferys, 2003, "Bayes' Theorem," *Journal of Scientific Exploration* 17(3:537-42).

to the problem.

(e) UB (p.12) worries that: "Beyond the argument that using explicit probabilities forces us to justify our choices and allows other scholars to evaluate them in the review process, the approach lacks any mechanism to enforce the recommendations." Here UB seems to want to hold Bayesianism to an unattainable standard. No method—Bayesian, frequentist, or otherwise—contains in and of itself any mechanism to enforce the recommendations it prescribes. The point is that Bayesianism provides guidelines and consistency checks to help scholars improve their reasoning and better communicate their judgments, while also helping other members of the research community to more effectively scrutinize their reasoning.

UB's (p.9, 16) call for a moratorium on teaching Bayesian process tracing until social scientists' ability to "reliably implement" the method has been definitively demonstrated likewise proposes a standard to which no other methodology has been or should be held. As an alternative, UB (p.14) simply recommends good data gathering, analysis, and writing—but does not offer any principled guidance by which their efficacy should be evaluated.

On these points, we might end with a salient quotation from E.T. Jaynes (1983:250-51):

> It is as true in probability theory as in carpentry that introduction of more powerful tools brings with it the obligation to exercise a higher level of understanding and judgment in using them. If you give a carpenter a fancy new power tool, he [*sic*] may use it to turn out more precise work in greater quantity; or he may just cut off his thumb with it. It depends on the carpenter.[11]

That is, training is required before scholars can effectively implement Bayesian reasoning, but that fact certainly does not justify withholding the training and the tool.

---

[11] E.T. Jaynes, 1983, "Where Do We Stand on Maximum Entropy?" in *Papers on Probability, Statistics, and Statistical Physics,* Kluwer Academic Publishers.