

Boosting Legal Probabilism

Marcello Di Bello and Rafal Urbaniak

1 The Book

1.1 Brief Description

In one or two paragraphs, describe the work, including its rationale, approach, and pedagogy. (This book is... It does... Its distinguishing features are...)

Can the evidence presented at trial be examined, weighed and assessed using probability theory? Can legal decision-making and standards of proof such as ‘preponderance of the evidence’ or ‘proof beyond a reasonable doubt’ be defined using the language of probability? Does the deployment of probability theory in assessing evidence and making decisions improve the accuracy and fairness of legal decision-making? Over the last fifty years, these questions have been debated in the literature in philosophy, law, forensic science and artificial intelligence. Legal probabilism is a research program whose supporters believe the answer to these questions is, by and large, affirmative. Others, however, hold a less optimistic view. ‘Boosting Legal Probabilism’ examines the most important objections to this research program and articulates a version of legal probabilism that is able to address many, if not all, of these objections.

We begin with the simple version of the theory. It comprises two key elements: first, Bayes’ theorem for assessing the weight of the evidence, and second, probability thresholds as decision criteria. This simple version has much to be recommended for, but also falls prey to several difficulties, including the problem of conjunction, puzzles of naked statistical evidence, the problem of priors. Some legal probabilists have attempted to dismiss these difficulties or downplay their significance. We confront them at face value and show they cannot be addressed within the confines of simple legal probabilism. We then develop a more sophisticated theory, what we call legal probabilism 1.02, which deploys Bayesian networks and takes advantage of seminal ideas in the literature in forensic science and artificial intelligence. ‘Boosting Legal Probabilism’ articulates the first comprehensive philosophical analysis of whether—and if so, to what extent—legal probabilism 1.02 can overcome the limitations of simple legal probabilism. We show that the more sophisticated version significantly improves on the simple version and rivals in explanatory power two competing accounts of judicial fact-finding: argumentation theory and relative plausibility. To add precision to the claims made in the book, the analytical argument is supplemented with an **R** code implementation.

‘Boosting Legal Probabilism’ is aimed at philosophers with an interest in legal epistemology and epistemology more generally. Many of the difficulties of legal probabilism resemble difficulties faced by Bayesianism in epistemology. The book will also draw attention outside philosophy from legal scholars who have championed applications of probability theory to evidence law as well as scholars who have resisted this trend. Another target audience includes computer scientists and psychologists interested in studying evidential reasoning and decision-making under uncertainty. Besides contributing to the literature about legal probabilism, the book aims to introduce unfamiliar readers to the rich interdisciplinary debate on the topic, often scattered throughout journals and books in philosophy, law, computer science, forensic science and psychology. So the book is aimed at scholars, advanced undergraduates and curious readers more generally. Some chapters present original research and require technical background in probability theory. Others are introductory, suitable for an advanced undergraduate course.

1.2 Outline

Part I - Legal Probabilism and Its Foes

The first part of the book will instill interest in legal probabilism among unfamiliar readers and refresh seasoned readers about the main points of contention. **Chapter 1** outlines simple legal probabilism which comprises a familiar repertoire: Bayes' theorem, likelihood ratios, probability thresholds, expected utility maximization. This repertoire has proven useful in several ways, especially in the assessment of explicitly quantitative evidence such as DNA matches and other expert evidence. At the same time, as **Chapter 2** shows, simple legal probabilism is liable to a host of conceptual difficulties: the conjunction problem, the problem of priors, paradoxes of naked statistical evidence. These difficulties are well-known. Others are less familiar: the problem of complexity, soft variables, the difficulty with corroboration.

The first two chapters provide the essential background for a deeper examination of legal probabilism and the development of its more sophisticated version. The remaining parts of the book cover two distinct topics: evidence assessment (Part II and Part III, Chapters 3 through 10) and decision-making (Part IV, Chapters 11 through ??). This distinction reflects the fact that legal probabilism is both a theory of evidence assessment (or evidence evaluation, evidence weighing) as well as a theory of decision-making at trial. These two topics are obviously intertwined in important ways, but are best kept separate for analytical clarity.

Part II - Evidence Assessment First Pass

The second part of the book discusses two formal tools used for the assessment of evidence from a probabilistic perspective: Bayes' theorem and likelihood ratios. We discuss how these tools can help to assess, weigh and evaluate evidence at trial as well as what their limitations are.

Chapter 3 begins with Bayes' theorem and surveys many of its applications, for example, as a tool to avoid reasoning fallacies such as the prosecutor's fallacy and the base rate fallacy. At the same time, its applications are also limited. As discussed in **Chapter 4**, court cases often require fact-finders to weigh several pieces of evidence, sometimes conflicting and susceptible to different interpretations. The hypotheses that the fact-finders are asked to evaluate in light of the evidence are structured stories or explanations constituted by several sub-propositions. This level of complexity can hardly be modeled by successive discrete applications of Bayes' theorem. A more sophisticated machinery for evidence assessment is needed.

Chapter 5 describes a formal tool that is, in some important way, different from Bayes' theorem and that many legal probabilists have found useful: likelihood ratios. Bayes' theorem requires one to assess the prior probabilities of the hypothesis of interest. The problem is that assessing prior probabilities is notoriously difficult. Likelihood ratios, instead, offer a way to evaluate the evidence presented at trial without the need of assessing prior probabilities. We illustrate the applications of likelihood ratios focusing on the debate about cold-hit DNA matches. The chapter also examines the weaknesses of this approach. The choice of the competing hypotheses compared in the likelihood ratio is often a source of confusion, manipulation and subjective judgment. Another problem is that, like simple applications of Bayes' theorem, likelihood ratios are still unable to model complex bodies of evidence. A further problem, as critics have alleged, is that likelihood ratios cannot model the notion of evidential relevance.

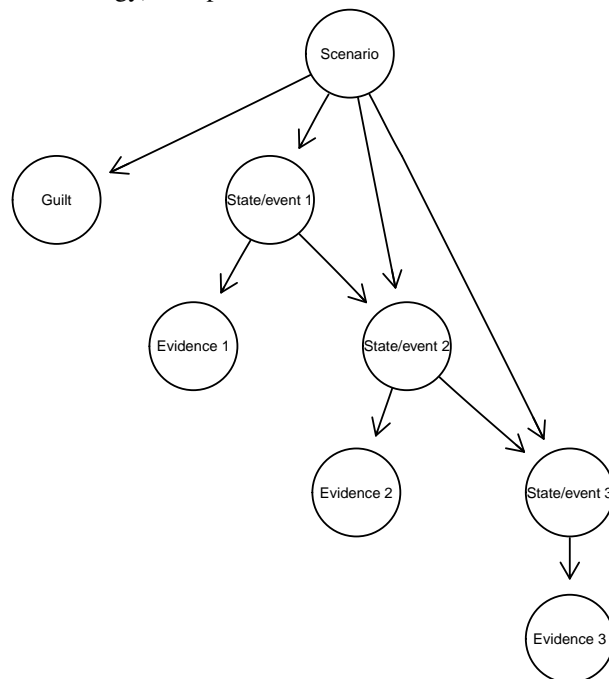
Part III - Evidence Assessment More and Better

Chapters 3 and 4 show that we need to move past simple legal probabilism. Chapter 5 shows that likelihood ratios, while useful in many ways, are still an unsatisfactory approach overall. The journey toward legal probabilism 1.02 is accomplished in Part III, Chapters 6 through 10. Our analysis is informed by the following working hypothesis. An accusation of liability must be substantiated by providing a well-specified account – a story or narrative – of the alleged illegal act committed by the defendant. This account consists of several moving parts, each supported by different items of evidence. The defense may respond by attacking the supporting evidence, the internal consistency of the account, or by offering an alternative account. Judges and jurors are tasked with assessing how well the claim that the defendant is civilly or criminally liable is supported by the evidence and how well it stands up against criticism. It is this complex

dynamics that we aim to formalize in the following chapters.

We focus on two aspects of the evaluation of trial evidence which simple legal probabilism is unable to model: first, judges and jurors often think holistically about the evidence, say in terms of coherent stories or explanations, without assessing the evidence by discrete applications of Bayes' theorem or likelihood ratios; second, different pieces of evidence interact in complex relationships, such as undercutting, rebutting, converging, corroborating evidence. We show that Bayesian networks constitute the formal machinery necessary for developing a more sophisticated legal probabilism that is able to capture these phenomena. The key idea is that the coherence of a story as well as conflicts between pieces of evidence can be modeled formally by corresponding properties of, operations on, and relations between Bayesian networks.

Chapter 6 offers a crash course on Bayesian networks with a focus on the assessment of legal evidence. A Bayesian network comprises a directed acyclic graph (called a DAG) that represents relations of dependence between variables, along with conditional probability tables corresponding to these relations. In the last decade, the literature in artificial intelligence and forensic science has made significant progress in modeling holistic notions such as the coherence of a story and argument-based notions such as conflicts between pieces of evidence. Chapter 6 surveys this literature focusing on the work of Charlotte Vlek and Norman Fenton. Vlek, together with Bart Verheij and Henry Prakken, proposed to model the coherence of a story by adding a node in the Bayesian network, call it a 'story node'. The story node has other nodes as its children corresponding to the events that make up the story. In turn, these events are linked to their supporting evidence. An example of a Bayesian network with a story node (or scenario node to use Vlek's terminology) is depicted below:



Since the story node unifies the different parts of a story, changes in the probabilities of these parts can be used to model the notion of coherence. The stronger the (positive) probabilistic dependence between the different parts, the more coherent the story. To model conflicts of evidence, a Bayesian network can be built that comprises two competing stories, say, one story was put forward by the prosecution and another by the defense, each supported by their own evidence. The network would specify that these stories are incompatible and cannot be true concurrently. Another approach to model conflicting evidence and competing stories was developed by Norman Fenton and his research group. Separate stories are represented by separate Bayesian networks, and Bayesian model comparison is then used for assessing the comparative evidential support of the competing stories.

Chapter 7 focuses on the story node approach as an account of coherence. The chapter

contains a critical argument followed by a positive proposal. We show that adding a story node by fiat – without any good reason for supposing that the different parts of the story are connected other than being part of one story – introduces unnecessary probabilistic dependencies between the elements of a story. In addition, the story node approach is overly simplistic as an account of coherence and fails to engage with the rich philosophical literature on the topic. After the critical argument, the chapter articulates a more adequate probabilistic account of coherence. Instead of adding a story node, we show that it is more appropriate to assess the dependency between the different parts of a story on a case-by-case basis and build the Bayesian network accordingly. We then define a formal notion of ‘story coherence’ that reflects properties of the Bayesian network used to model the evidence. We show that our formal notion of coherence addresses the objections against probabilistic accounts of coherence in the philosophical literature. [M: ANYTHING ELSE TO DESCRIBE THE POSITIVE ACCOUNT?]

Chapter 8 focuses on conflicts between pieces of evidence. Neither Vlek’s story node approach nor Fenton’s model comparison approach adequately capture how pieces of evidence and competing stories may conflict with one another. It is too simplistic to posit that the complex adversarial dialectic that takes place in a trial could be modeled by averaging different Bayesian networks (Fenton) or postulating relationships of incompatibility between different story nodes (Vlek). We need an account of more fine-grained notions, such as undercutting and rebutting evidence, and more generally we need an account of how cross-examination operates at trial. What cross-examination often accomplishes is not so much the creation of an alternative story, but rather, the reinterpretation of an existing story by supplying additional information. We show that this process of reinterpretation can be represented formally as the refinement of an existing Bayesian network. Conflicts between pieces of evidence such as undercutting and rebutting can be modeled by drawing additional arrows between evidence nodes and hypothesis nodes.

The reverse of the phenomenon of conflicting evidence is that of converging evidence, in particular, the fact that one piece of evidence corroborates another. Corroboration has been the focus of extensive scholarly debate often independently of the debates within legal probabilism. **Chapter 9** surveys the literature on corroboration and the main difficulties that have been levelled against proposed probabilistic accounts. The chapter then formalizes a notion of corroboration using Bayesian networks that overcomes most of the difficulties of existing accounts. [M: ANYTHING ELSE TO DESCRIBE THE POSITIVE ACCOUNT?]

Chapter 10 compares legal probabilism 1.02 to competing accounts of judicial fact-finding – in particular, argumentation theory and relative plausibility – and draws some general morals. Argumentation theory is well suited to model conflicts between evidence, but cannot easily model the fact that evidence may conflict more or less strongly with other evidence. Unlike argumentation theory, legal probabilism 1.02 offers an account of evidential support, conflict and convergence that captures how these relations come in degrees of strength. The other competing theory we consider, relative plausibility, is often criticized because the defense lawyer need not always present a full-fledged alternative story. Without settling this controversy, we note that legal probabilism 1.02 is flexible enough to model competing stories (in agreement with relative plausibility) or model conflicts without the need to construct a full-fledged alternative story (as critics of relative plausibility prefer).

While legal probabilism 1.02 can still be challenged because of its questionable empirical adequacy – judges and jurors hardly follow probability theory – we emphasize how it offers a rich account of evidential support beyond what critics have allged. Susan Haack, for example, has criticized legal probabilism for its monodimensional view of evidential support. This is true of simple legal probabilism in which evidential support is modeled by the posterior probability of a hypothesis given the evidence, or by a likelihood ratio. Legal probabilism 1.02 offers a richer account in which evidential support also depends on the degree of specificity and coherence of the story put forward, and the extent to which the supporting evidence withstands objections. These features of evidential support – specificity and coherence, as well as resistance to objections – can serve to formulate decision criteria that are not mere unidimensional thresholds. Decision-making at trial is the topic of the next part of the book.

Part III

The final part of the book examines trial-decision making, specifically, to what extent standards of proof such as ‘preponderance of the evidence’ and ‘proof beyond a reasonable doubt’ can be understood through the lenses of probability theory. This is perhaps where most of the philosophical debate of the last decades has focused. There has been a spur of research arguing that legal probabilism is unfit to model standards of proof. But this research often holds an overly narrow view of legal probabilism. We show that legal probabilism 1.02, as we have developed it in the earlier parts of the book, provides the adequate groundwork for theorizing about standards of proof.

Chapter 11

Chapter 12 tackles the puzzle of naked statistical evidence. This is a topic of enormous scholarly attention. A book on legal probabilism would not be complete without addressing it. Our solution to the puzzle consists of two observations. First is the observation that an accusation of liability should be substantiated in a well-specified account – or story, narrative – whose moving parts are each supported by adequate evidence. In cases of naked statistical evidence, the probability of liability is very high, but the specificity of the accompanying narrative is suspiciously low. This explains people’s intuitive resistance against a finding of liability. The second observation is that the supporting evidence should typically be ‘causally grounded’. This grounding is achieved during cross-examination by eliciting additional information about the relation between the evidence and the alleged facts – for example, information about the visibility conditions for an eyewitness testimony; the academic credential of an expert witness; the chain of custody of a document. However, statistical evidence blocks cross-examination because no undercutting evidence could in principle be brought against it, and thus cannot be causally grounded. Importantly, both these observations – lack of specificity of the accusatory narrative and lack of causal grounding – can be modeled using Bayesian networks. So our solution to the puzzle of naked statistical evidence is housed within the confines of legal probabilism, albeit not the narrow version its critics have in mind.

Chapter 13 tackles another central problem for legal probabilism, the difficulty with conjunction.

Chapter 14

Table of contents

- I Legal probabilism 1.01 and its foes
 - 1 The emergence of legal probabilism
 - 1.1 Famous cases
 - 1.2 Probabilistic evidence
 - 1.3 Trial by mathematics
 - 1.4 Some history
 - 2 A skeptical perspective
 - 2.1 The difficulty about conjunction
 - 2.2 The problem of priors
 - 2.3 Naked statistical evidence
 - 2.4 The complexity problem
 - 2.5 Soft variables
 - 2.6 Corroboration
 - 2.7 The reference class problem
 - 2.8 Non-probabilistic theories
- II Evidence assessment First Pass
 - 3 Bayes’ Theorem and the usual fallacies
 - 3.1 Assuming independence
 - 3.2 The prosecutor’s fallacy
 - 3.3 Base rate fallacy
 - 3.4 Defense attorney’s fallacy
 - 3.5 Uniqueness fallacy

- 3.6 Case studies
- 4 Complications and caveats
 - 4.1 Complex hypotheses and complex bodies of evidence
 - 4.2 Source, activity and offense level hypotheses
 - 4.3 Where do the numbers come from?
 - 4.4 Modeling corroboration
 - 4.5 Stories, explanations and coherence
- 5 Likelihood Ratios and Relevance
 - 5.1 Likelihood ratio as a measure of evidence strength
 - 5.2 The risk of false positive and its impact
 - 5.3 Hypothesis choice
 - 5.4 Levels of hypotheses and the two-stain problem
 - 5.5 Relevance and the small-town murder scenario
 - 5.6 The cold-hit confusion
 - 5.7 Likelihood ratio and cold-hit DNA matches
- III Evidence assessment More and Better
 - 6 Bayesian Networks
 - 6.1 Bayesian networks to the rescue
 - 6.2 Legal evidence idioms
 - 6.3 Scenario idioms
 - 6.4 Modeling relevance
 - 6.5 Case study: Sally Clark
 - 6.6 DNA evidence
 - 7 Coherence
 - 7.1 Existing probabilistic coherence measures
 - 7.2 An array of counterexamples
 - 7.3 Coherence of structured narrations with Bayesian networks
 - 7.4 Application to legal cases
 - 8 Conflicts
 - 8.1 Argumentation theory
 - 8.2 Undercutting and rebutting evidence
 - 8.3 Cross-examination
 - 8.4 Conflicting evidence in Bayesian networks
 - 9 Corroboration
 - 9.1 Boole's formula and Cohen's challenge
 - 9.2 Modeling substantial rise in case of agreement
 - 9.3 Ekelöf's corroboration measure and evidentiary mechanisms
 - 9.4 General approach with multiple false stories and multiple witnesses
 - 10 Towards Legal Probabilism 1.02
 - 10.1 Outperforming competing accounts
 - 10.2 Empirical adequacy
 - 10.3 Specificity and coherence
 - 10.4 Resistance against objections
 - 10.5 Bayesian network implementation
- IV Standards of proof
 - 11 Are Standards of Proof Thresholds?
 - 11.1 Legal background
 - 11.2 Probabilistic thresholds
 - 11.3 Theoretical challenges
 - 11.4 The comparative strategy
 - 11.5 The likelihood strategy

- 11.6 Probabilistic thresholds revised
 - 11.7 Bayesian networks and probabilistic standard of proof
- 12 Naked statistical evidence
 - 12.1 Forty years of hypotheticals
 - 12.2 Specific narratives
 - 12.3 Cross-examination and causal grounding
 - 12.4 Bayesian networks and naked statistical evidence
 - 12.5 Are cold-hit DNA matches naked statistics?
- 13 The Difficulty with Conjunction
 - 13.1 The problem
 - 13.2 The likelihood strategy
 - 13.3 The comparative strategy
 - 13.4 The holistic strategy
 - 13.5 Complex bodies of evidence and structured narratives
- 14 Other accounts
 - 14.1 Baconian probability
 - 14.2 Sensitivity
 - 14.3 Normic Support
 - 14.4 Foundherentism
 - 14.5 Relevant alternatives
 - 14.6 Knowledge
 - 14.7 Relative Plausibility
 - 14.8 Arguments
- V Accuracy and Fairness
 - 15 The functions of the proof standards
 - 15.1 Conceptual desiderata
 - 15.2 Protecting defendants
 - 15.3 Error reduction and error distribution/allocation
 - 15.4 Dispute resolution and public deference
 - 15.5 Justification and answerability
 - 16 Accuracy and the risk of error
 - 16.1 Minimizing expected costs
 - 16.2 Minimizing expected errors
 - 16.3 Expected v. actual errors
 - 16.4 Competing accounts of the risk of error
 - 16.5 Bayesian networks and the risk of error
 - 17 Fairness in trial decisions
 - 17.1 Procedural v. substantive fairness
 - 17.2 Competing measures of substantive fairness
 - 17.3 Bayesian networks and fairness
 - 18 Conclusions

1.3 Outstanding Features of the Book

- (First) comprehensive sustained philosophical discussion of legal probabilism.
- Multi-faceted in its incorporation of insights from various discussions present in legal, philosophical, and forensic research.
- With a practical accent, due to the implementation of the conceptual points by means of bayesian networks and R programming language.

what else?

1.4 Apparatus

a. Will the book include photographs, line drawings, cases, questions, problems, glossaries, bibliography, references, appendices, etc.?

Yes, the book will contain various plots, either of Bayesian networks, or some other data visualisations generated by `ggplot2`. The book also will contain bibliography.

b. If the book is a text, do you plan to provide supplementary material to accompany it? (Teacher's manual, study guide, solutions, answers, workbook, anthology, or other material.)

The book will be accompanied by an online-only appendix detailing the use of the R code in the book and the source code we used.

1.5 Competition

a. Consider the existing books in this field and discuss specifically their strengths and weaknesses. Spell out how your book will be similar to, as well as different from, competing works.

Three types: BNs in the law, Philosophy & law, Statistics in law and forensics

- “Bayesian Networks and Probabilistic Inference in Forensic Science” by Taroni, Aitken, Garbolino and Biedermann.
- “Risk Assessment and Decision Analysis with Bayesian Networks” by Fenton and Neil.
- “Bayesian Networks With Examples in R” by Marco Scutari and Jean-Baptiste Denis.
- Alex Stein, foundations of evidence law
- Nance, Burdens of proof
- Schauer, Profiles, ...
- Ho, Philosophy of evidence law
- Robertson, Vignaux
- Lucy Dawid,
- Statistics for Lawyers etc.

b. Consider what aspects of topical coverage are similar to or different from the competition. What topics have been left out of competing books and what topics have been left out of yours?

c. Please discuss each competing book in a separate paragraph. (If possible, please provide us with the publisher and date of publication as well.) This information will provide the reviewers and the publisher a frame of reference for evaluating your material. Remember, you are writing for reviewers and not for publication, so be as frank as possible regarding your competition. Give credit where credit is due, and show how you can do it better.

For now, let's list competition, and discuss key differences

2 Market Considerations

2.1 The Primary Market

1. What is the major market for the book? (Scholarly/professional, text, reference, trade?)
2. If this is a text, for what course is the book intended? Is the book a core text or a supplement? What type of student takes this course? What is the level? (Major or non-major; freshman, senior, graduate?) Do you offer this course yourself? If so, how many times have you given it? Is your text class-tested?
3. If the market is scholarly/professional, reference, or trade, how may it best be reached? (Direct mail, relevant journals, professional associations, libraries, book or music stores?) For what type of reader is your book intended?

3 Status of the Work

1. Do you have a timetable for completing the book?
 - a. What portion or percentage of the material is now complete?

- b. When do you expect to have a complete manuscript?
- 2. What do you estimate to be the size of the completed book?
 - a. Double spaced typewritten pages normally reduce about one-third when set in type; e.g., 300 typewritten pages make about 200 printed pages. There are about 450 words on a printed page.
 - b. Approximately how many photographs do you plan to include?
 - c. Approximately how many line drawings (charts, graphs, diagrams, etc.) will you need?
 - d. Do you plan to include material requiring permission (text, music, lyrics, illustrations)? To what extent? Have you started the permissions request process?
- 3. Do you plan to class-test the material in your own or other sections of the course? (Any material distributed to students should be protected by copyright notice on the material.)

4 Sample Chapters

Select one or two chapters of the manuscript that are an integral part of the book. They should be those you consider the best-written ones, and do not have to be in sequence. For example, you might submit chapters 3, 7, and 14 of a 20-chapter book, so long as these chapters represent the content and reflect your writing style and pedagogy in the best possible light. It is also advisable to submit any chapter that is particularly innovative or unique. Sample chapters should contain rough sketches, charts, hand-written musical examples or xerox reproductions, and description of photographs to be included. The material need not be in final form, although it should be carefully prepared and represent your best work. In your preparation, emphasis should be on readability. Please do not bind your manuscript, as we will have to unbind it in order to make photocopies for reviewers. Also be sure all pages are numbered either consecutively or double-numbered by chapter.

5 Reviews

If we are interested in your project, we will commission outside reviewers to read and evaluate your proposal. We will, of course, obtain the best available reviewers to consider your work. If you wish to suggest the names of experts in your field whom you believe to be ideally suited to evaluate your proposal, you may provide their names, titles, and email addresses. While we are unlikely to approach these scholars to act as reviewers themselves, we may ask them for their suggestions for peer readers. Naturally, we do not reveal the names of reviewers without their permission.

6 Author Background

Please include a current CV or brief biography of your writing, teaching, and/or educational background and experience. Be sure to list any books that you have previously published, and any other information about yourself on why you are qualified to write this book.

7 Response Time

Please allow at least 6-10 weeks for the manuscript proposal evaluation and review process. We will contact you as soon as we have had a chance to thoroughly examine your manuscript proposal. Thank you for your interest in Oxford University Press. We look forward to reading your materials.